Heliyon 7 (2021) e08315

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Predictive analysis of the value of information flow on the shop floor of developing countries using artificial neural network based deep learning

André Marie Mbakop [a,b,c,*], Florent Biyeme [a,b,c], Joseph Voufo [a,b,c], Jean Raymond Lucien Meva'a [a,b,c]

[a] *National Advanced School of Engineering of Yaoundé (NASEY), Cameroon*
[b] *Laboratory of Civil and Mechanical Engineering, Yaoundé, Cameroon*
[c] *Department of Industrial and Mechanical Engineering, Yaoundé, Cameroon*

HIGHLIGHTS

- Mathematical modeling of the information flow on the shop floor.
- Information flow analysis model with ANN, PSO-ANN, and GA-ANN for regression as deep learning.
- Predictive analysis system of the value of information flow (VIF) for performance improvement.

ARTICLE INFO

Keywords:
Deep learning
Characteristics of information flow
Developing countries
Performance
PSO-ANN
Value of information flow

ABSTRACT

To facilitate the continuous improvement of performance and the management of information flow (MIF) for production and manufacturing purposes on the shop floor of developing countries, there is a need to characterize information flow that will be shared during the process. MIF provides a key performance shop floor metric called the value of information flow (VIF). Previous methods have been used to analyze VIF in developed countries. However, these methods are sometimes limited when applied to developing countries where the shop floor is disorganized. It then renders the MIF with the imported software inefficient because of the gap between the user environments. Taking Cameroon as a case study, this study proposes a new method of modeling and analyzing the information flow and its value based on the characteristics of information flow (CIF) for developing countries. In addition, a predictive analysis of the VIF based on CIF using an artificial neural network (ANN) on one hand and optimized ANN with particle swarm optimizer (PSO) and genetic algorithms (GA) on the other is performed. The ANN model of regression developed has the following performance: coefficient of determination: 0.99 and mean squared error (MSE): 0.00043. For the PSO-ANN, the MSE decreased to 0.00011, and this model result was similar to that of the deep learning model used for regression. The GA-ANN model results were not as satisfactory as those of the PSO-ANN model. A predictive system to analyze VIF is proposed for managers of companies in developing countries.

1. Introduction

Increases in productivity and leadership markets have always been the main target of companies. For production, one requires raw materials, machines, process implementation, and all necessary information to run the production. Production requirements impose good management of production flow, and flow management focuses on end-to-end planning of materials and information flows from suppliers to customers (Wagner et al., 2003). In production companies, material flows are formed of raw materials that are used in manufacturing processes to deliver finished products and sometimes waste from production errors that may occur. In contrast, information flow includes all the inquiries, data sharing that comes before the processed materials, raw material for processes, operators considered as machines or humans, and details on the finished product to be delivered to the customer on a deadline. The high level of technologies in developed countries has made their companies face competition in the international market because of the good management of information flow (MIF) in the shop floor by automated devices. In some cases, these devices replace the human operator's work in the planning and control of material flows, which require good

* Corresponding author.
E-mail address: mbakop.andre@yahoo.com (A.M. Mbakop).

https://doi.org/10.1016/j.heliyon.2021.e08315
Received 27 May 2021; Received in revised form 29 July 2021; Accepted 29 October 2021

management of numerous information flows (Richter et al., 2017; Ferrero Bermejo et al., 2019; Lu et al., 2020; UNIDO, 2021). These numerous information flows are most often related to appropriate software such as enterprise requirement planning (ERP) systems. However, in some small and medium-sized enterprises (SMEs), there are constraints and complexities in flow management (information flow) that impede the production performance of the shop floor (SF) (Drigo et al., 2020). In the context of SMEs in some developing countries, there is a fuzzy circulation of information flow on the SF due to poor MIF (Durugbo et al., 2011; Durugbo and Erkoyuncu, 2014). The software that has been designed to manage static information in well-organized production companies cannot perform the same tasks on the SF of developing countries where information flow is dynamic and does not sometimes respect specific criteria of flow. This situation is presented in Figure 1 by Mbakop et al. (2021).

In manufacturing processes, material removal processes, woodworking, and welding processes, information is shared through the medium depending on how the SF is configured or highly developed. Understanding how information is shared allows for the characterization of information flow. A characterized information flow makes the MIF easier for traditional and semi-developed enterprises. Static or dynamic information flow will have value and can add value to a production process (Tomanek et al., 2020). To improve the performance of a SF for manufactured products, some researchers have focused their work on the VIF because information flow in the manufacturing SF is at the center of all production processes, and it influences other performance criteria such as product quality, operator competence, and production due time (Jabur and Dawood, 2015; Sütöova and Seginakova, 2018). Related works on the analysis of the value of information flow (VIF) have mostly focused on the quality and dimension (medium) of information flow using methods of information process integration, and the value-added heat map. In these studies, dimensions such as paper information and visual and audio information could not provide a major VIF to increase the performance of the SF. Visual management is currently a major tool for information flow management for the SF. Their results showed that automated and digital (Internet of Things) were the best information-sharing media that had suitable VIF (4, 5, 0.8, and 1). Using most of the CIF, the above methods may be limited because of the large dataset produced by the CIF on the SF. A large dataset, the complexity of information flow on the SF, and the dynamic change of information flow occurring in the production chain demand a machine learning model to carry out predictive analysis of the VIF. Pandey et al. (2020) and Aldahoul et al. (2021) discussed the performance of various machine learning algorithms for regression, such as decision trees (DT), multiple linear regression (MLR), and support vector machine (SVM), compared to artificial neural networks (ANNs). In this paper, we propose an artificial neural network (ANN) model for the analysis (regression) of the VIF, and then, we use artificial algorithms related to complex problems such as the particle swarm optimizer (PSO) and genetic algorithms (GA) to improve the performance of ANNs. This analysis is based on most of the CIF in the SF having specific descriptions as those of developing countries, by considering two scenarios. Many scenarios can occur in the SF environment of developing countries, such as lack of information concerning the raw materials (it concerns the type of information flow) when the production process is to be launched. Another one can be, the influence of the information given by a production supervisor to an operator machine considered as the direction of information. However, in this paper, the focus will be on the impact of the presence and absence of information disruptions (parameters of information flow) on the management of dynamic information flow when using PSO-ANN. These scenarios are as follows: The first considers all the CIF. The second scenario considers the parameters of information flow as excluded from the CIF because some information flow parameters of CIF are consequences of disruptions in information flow sharing. In some developing countries, small- and medium-sized manufacturing enterprises are still in this stage of production with semi-automated techniques because of the lack of financial resources and the lack of technology transfer (Paul, 2020; Akpan et al., 2020). Considering how small- and medium-sized companies in Cameroon struggle to move from a semi-automated process to an automated process because of the lack of financial resources (Ngoungo, 2012)

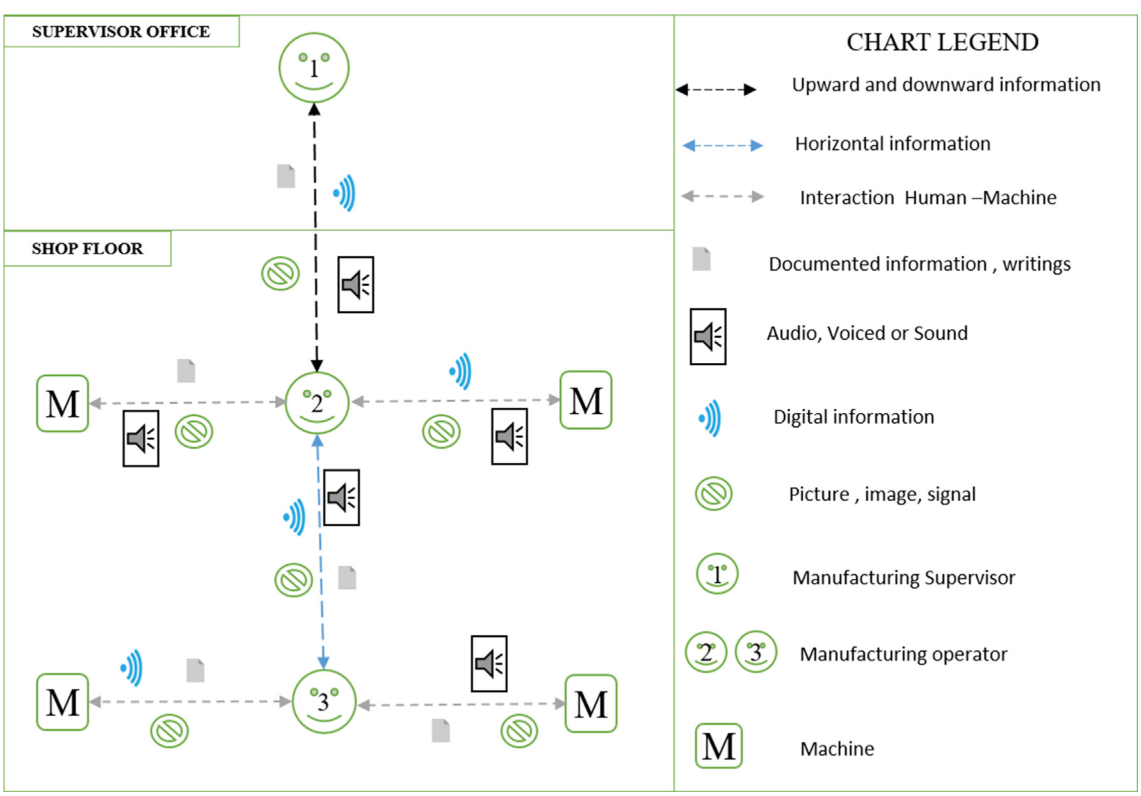


Figure 1. Information flow sharing in small- and medium-sized manufacturing companies of developing countries.

and technology transfer, small- and medium-sized manufacturing companies in Cameroon will be taken as a case study. This paper presents a general review of the CIF and related methods used to analyze the VIF. Then, information flow modeling is presented after an audit data survey. The performance of the PSO-ANN and deep learning models used for regression will be compared, and the PSO-ANN model for regression will be used to analyze the VIF. Finally, a VIF analysis system will be proposed for managers of developing countries.

2. Characteristics of information flow

The characterization of information flow aims to know in detail what makes an information flow not useful in a production operation. The CIF globally presented by Mbakop et al. (2021) is shown in Figure 2. The characterization of the information flow in the SF is based on the following:

- The types of information flow describe different types of information that are directly or indirectly related to the process fabrication of a product.
- The dimension of information flow is the medium (paper, visual, audio, electronic not real time (ENRT), electronic real time (ERT), and digital) used to transfer information.
- The direction of information flow can be presented as axes of circulation of information flow depending on the level of decision-making in the production company.
- The information flow parameter concerns all information disruptions, complexities, and delays in the production process.
- The quality of information flow represents all the attributes of a shared information flow integrated into industrial processes.

A proper explanation of each sub-characteristic according to the manufacturing or production environment is presented in Table 1. The CIF allows us to easily comprehend the value of information.

3. Review methods on the analysis of the VIF in SF operations

The VIF on the SF has long been of great interest in the research domain, and can be defined as the weight that information can have on industrial or manufacturing processes. Therefore, information flow can stop the start of production operation; information flow can stop a production operation that is going on; information flow can speed up the production operation; and finally, it can have a weighted impact on the product quality and manufacturing due time. Considering that not many studies have focused on these in the literature to the best of our knowledge, we will present some methods used to analyze the VIF in a SF based on some CIF.

3.1. The value-added heat map

Tomanek and Schröder (2017) proposed an innovative method, the value-added heat map (VAHM), which is a visualization tool that

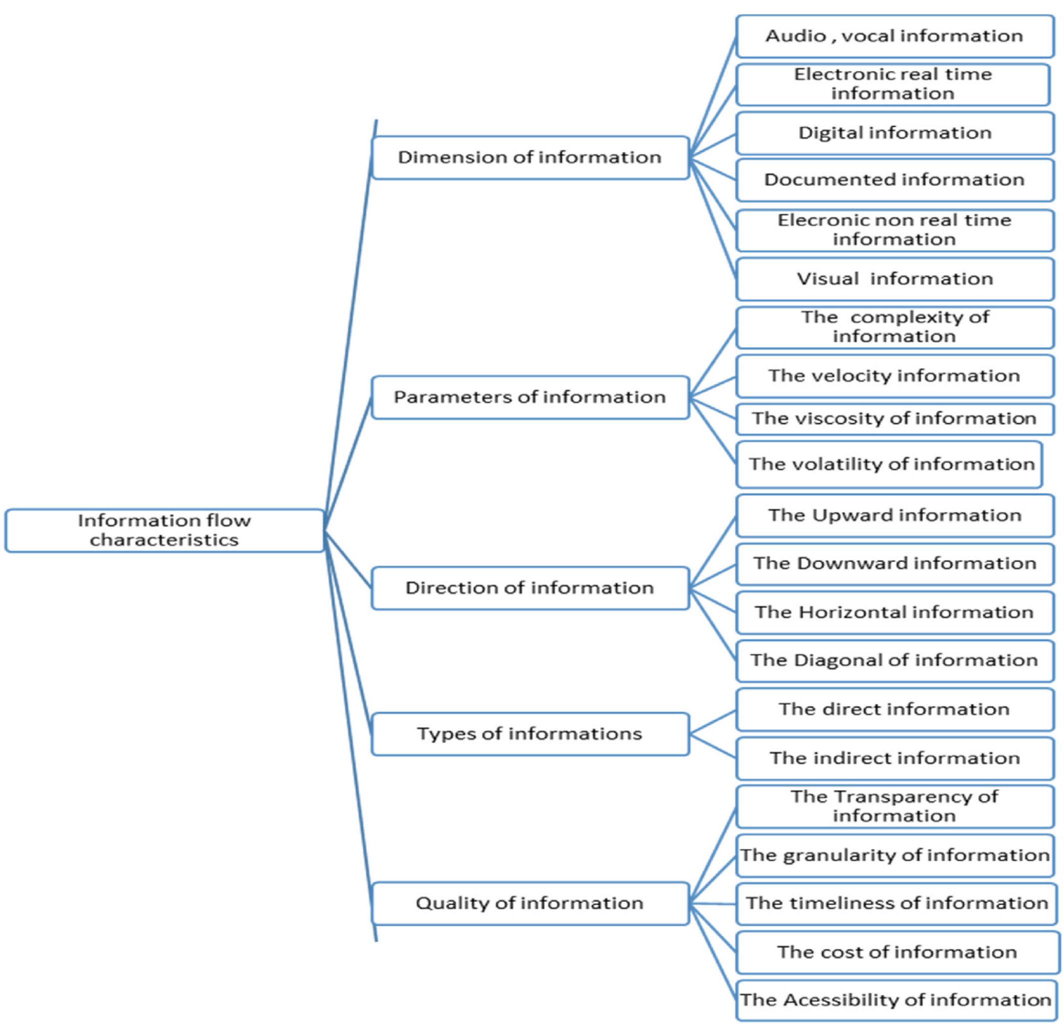


Figure 2. General information flow characteristics.

Table 1. Information flow characteristics described in the production or manufacturing environment (Author's source).

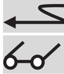
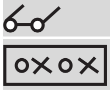
Characteristics of information flow	Sub-characteristics of information flow	Description in the production or manufacturing domain
Type of information	Direct information	Information related directly to the process; e.g., information about raw materials, process transformation, and customer order.
	Indirect information	Information that may stop the production chain; information that indirectly influences the production. Machine breakdown or unavailability, incompetent, absent, or untrained workers, void information discussed in the shop floor, and inadequate machine and tool settings.
Dimension of information	Documented or Paper information	Manufacturing order, the document of raw materials, change of product order, customer order, work order, spare parts document, datasheet, and history of breakdown document.
	Audio, vocal information	Alarm for maintenance intervention, call, word to word information, recorded audio information.
	Visual information	Images, pictograms, signboard, live information on the screen, total productive management information on board, production key performance indicator on board.
	None real-time electronic information	Information sent through email, fax that necessitates a printer, logic automated with cable.
	Real-time electronic information	Automated machine, sharing information through Grace sequences, or programming languages, or computer aided manufacturing (CAM) and computer aided design (CAD) sending information through a cable.
	Digital information	Interconnected things that share information, absence of electronic cables, information is shared wirelessly.
Direction of information	Upward information	The information that is sent by personnel (a machine, computer) of a lower decision level to personnel (a machine, computer) of a higher decision level.
	Downward information	The information that is sent by personnel (a machine, computer) of a higher decision level to personnel (a machine, computer) of a lower decision level.
	Horizontal information	The information that is sent by personnel (a machine, computer) of the same decision level to another personnel (a machine, computer) of the same decision level.
	Diagonal information	The information that is sent between managers of different companies.
Parameters of information	Velocity	The speed with which information is sent depends on the support or medium of information. From operation 1 to operation 2.
	Viscosity	The noise of the environment, the low connectivity, the incomplete information that can be found in information for the execution of operations.
	Complexity	The understandability of information, referring to the language used for information flowing for a production operation to be executed.
	Volatility	The loss of information when executing an operation or when moving from task 1 to task 2 in a manufacturing process, the loss of data.
Quality of information	Transparency	The ease with which information that is passing from one job post to another can be understood.
	Accessibility	The availability, the reliability, and the ease to use of information in an operation or a task process.
	Timeliness	Just in time information for the realization of a task process or operation.
	Granularity	Information passed from one task 1 in a process to another task 2 has to be detailed.

indicates the level of value creation concerning production-relevant factors. This method is based on a thermal heat map camera. The VAHM uses different colors to show where the information flow has a great value using a scale. For the scale to describe the VIF, they based their work on the impact that the dimension (medium) of a shared information flow can have on the production operation tasks occurring on the SF. The obtained results are presented in Table 2 and Figure 3, which illustrate the information flow in circulation in the SF and the VAHM, respectively.

3.2. Quantification of the information flow

VIF analysis can be performed by quantifying or determining the amount of information shared. To determine the amount of information or the capacity of the shared information flow in a medium system, Mourtzis et al. (2019) quantified the information flow using Shannon's theory on information for a transition from a traditional SF to an Industry 4.0 SF based on the research work of Mourtzis et al. (2017). Shannon's theory considers the entropy of information flow found in the dimension of information commonly mentioned as the medium of information flow. This theory includes the quantity of information flow transferred through a medium. The results showed that when the entropy of the information flow is null, the information transmission capacity is 0.

Table 2. Value stream analysis symbols for the visualization of information flow (Tomanek et al., 2016).

Symbols for the information flow	Meaning
	Manual information flow
	Electronic information flow
	Electronic information flow
	Leveled production planning
	Route of a Kanban card

3.3. Digitalization degree of information flow

Tomanek et al. (2020) used the digitalization degree of information flow to analyze the VIF based on the hypothesis of the results obtained by Tomanek and Schroder (2017). The VIF was considered as the degree of digitalization on the SF. They found that the more the SF was digitalized, the higher the VIF was. The digitalization degree that they computed depends on the amount of information that can flow through a medium according to the given information flow scale by Tomanek and Schroder (2017).

3.4. Information quality value stream mapping (IQVSM)

The IAVSM is a recent method developed by Busert and Fay (2020) based on value stream mapping (VSM) and used to capture the information flow that will serve for production planning and control to improve the performance of the SF. In the developed method, they considered information quality based on granularity, timelines

Categorization	Value Added Level	Dimension of information Flow	Scale
No Added Value	0	Insufficient, incorrect, or unnecessary exchange of information	Effort
Limited Added Value	1	Written exchange of information (e.g., paper document, fax, and e-mail)	
	2	Verbal or visual exchange of information	
	3	Electronic exchange of information not in real time (e.g., by spreadsheet application)	
	4	Real-time electronic exchange of information (e.g., by system application)	
Maximum Added Value	5	Real-time digital exchange of information (e.g., by the Internet of Things and Services)	

Figure 3. VAHM evaluation scale for information flow (Tomanek and Schröder, 2017).

(frequency), and accessibility of information to improve the value that an information flow may have on the SF. To accomplish their objectives, they carried out interviews from which they received information concerning the SF to build up scenarios for their analysis. To harmonize the information quality, they quantified the tolerance of the quality of information; the lower the tolerance, the higher the value of information. They also remarked that the percentage tolerance of information quality is very important in planning control and SF management. This type of analysis is related to the value of information.

3.5. SF management technique

Based on simple shop-floor management related to lean management, Mathiasen and Hass (2021) focused on the digitalization of the SF to improve its performance by reducing superfluous work that occurs due to lack of information, obsolete information, inappropriate information, and poor quality information illustrated with low accessibility. To improve the performance of the SF based on the reduction of superfluous works on the SF, they focused on improving collaboration with the accessibility of information quality on all shop floors.

The SF management method has been used to improve communication in the shop for production activities by Benyahya and Macurova (2021). The developed method of shop floor management consists of interviewing the workers about how communication is occurs to improve the flow of productivity. They found that in some companies, communication between employees during meetings and brainstorming was noisy due to machine noises, and this may have a poor influence on operators' or workers' performances in the company. A similar situation occurs when the parameters of information flow are used. To directly solve the issues related to noise, they used a microphone during meetings to facilitate the flow of information.

From the above-presented methods, it is evident that the VAHM considers the dimensions of information flow to analyze the VIF, and the authors did not consider other characteristics such as quality, direction, and parameters. Certainly, the dimension of the information scale (from 1 to 5, Figure 3) can be a function of the quality, parameters, and direction of information, but this was not mentioned in their work. Considering all the CIF to analyze the VIF in a production process will not be an easy task if we use the VAHM. Likewise, for the IQVSM, the quality of information flow and the dimension based on the digital were portrayed as major CIF and as those that have to be improved on the SF when production operations are launched. However, this method does not integrate the fact that disruptions of information flow found in the most digitalized SF can reduce the VIF. The quantification method used did not consider the direction of information or the quality of information. Therefore, quantification of information flow may not be accurate enough to analyze the CIF by considering the 20 variables that we have as inputs. Using lean management techniques to improve the SF is a good idea when we focus on materials related to information flow; however, this will be more efficient if lean management was used with another method that could consider all the CIF of information flow. Knowing that the above-proposed methods from the literature are limited to analyzing the VIF based on the CIF, it is therefore important to focus on another method that integrates all the CIF for predictive analysis of the VIF in the SF where production activities are being carried on; thus, artificial intelligence is adapted for complex systems with high input variables.

4. ANN, PSO-ANN, and GA-ANN: major analysis tools for complex systems

A general view of machine learning algorithms based on supervised learning for prediction presents an ANN (a multi-layer design based on the human brain's biological neural process used to solve difficult problems, where its goal is to converge towards mathematical models) as a predictive model that can analyze a large amount of data related to the input parameters. ANNs are adaptive to complex problems and systems even when the network typology changes; it can find complex relations among variables with good accuracy (Ferrero Bermejo et al., 2019; Jalaee et al., 2019; Habtamu and Megersa, 2021). Following studies focusing on the comparative analysis (Pandey et al., 2020; Aldahoul et al., 2021) of machine learning algorithms for regression such as linear regression (LR), DT, support vector machine (SVM), and ANN, we find that ANN is more efficient than the above algorithms. The complexity of the analysis of the information flow value will pass through an ANN model that will be built. The development of the ANN model depends on three fundamental aspects: the inputs and activation functions, the network architecture, and the weight of each input connection (Kankar et al., 2011; Qiokat and Khan, 2014; Osisanwo et al., 2017; Thomas et al., 2017; Lingitz et al., 2018; Cinar et al., 2020; Hosseinzadeh et al., 2020; Göppert et al., 2021). In this work, the ANN model is formed of the inputs, which are the characteristics of the information flow, and the hidden layers are activated by the rectified linear units (ReLu) function presented by Eq. (1) because the data are between 0 and 1, and the output (VIF) will be conditioned by the linear function.

$$A(x) = \max(0, x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{1}$$

Sachdev (2020) proposed that the number of hidden layers must be a maximum of three. According to Sheela and Deepa (2013), the number of hidden nodes is given by Eq. (2);

$$N_h = \sqrt{N_i N_o} \tag{2}$$

with N_i being the number of neurons of the inputs and N_o the number of neurons of the output.

From the ANN model, the loss function can be evaluated from the error criteria describing the performance of the ANN: the mean square error (MSE), root mean square error (RMSE), mean relative error (MRE), mean

$$MSE = \sum_{i=1}^{N} \frac{\left(y_i' - y_i\right)^2}{N} \tag{3}$$

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{\left(y_i' - y_i\right)^2}{N}} \tag{4}$$

$$MAE = \sum_{i=1}^{N} \left| \frac{\left(y_i' - y_i\right)}{N} \right| \tag{5}$$

absolute error (MAE), Huber loss function (L_δ), and mean absolute percentage error (MAPE), which were calculated to assess the performance of the proposed model.

$$MAPE = \frac{100\%}{N} \sum_{i=1}^{N} \left| \frac{\left(y_i' - y_i\right)}{y_i'} \right| \tag{6}$$

$$L_\delta \begin{cases} \frac{1}{2}\left(y_i' - y_i\right)^2, \text{ if } \left|y_i' - y_i\right| \leq \delta \\ \delta\left|y_i' - y_i\right| - \frac{1}{2}\delta^2, \text{ Otherwise.} \end{cases} \tag{7}$$

As the ANN model is for regression, it is important to know the coefficient of determination (R^2), which is given by:

$$R^2 = \frac{\sum_{i=1}^{N}\left(y_i' - \overline{y_i'}\right)\left(y_i - \overline{y_i}\right)}{\sum_{i=1}^{N}\left(y_i' - \overline{y_i'}\right)^2 \sum_{i=1}^{N}\left(y_i - \overline{y_i}\right)^2} \tag{8}$$

where y_i is the predicted output, y_i' is the actual output, and N is the number of samples. For ANN, we will use python Tensorflow for simulation.

PSO and GAs have been used to train ANNs for other purposes in the engineering sector (Le et al., 2019; Pant and Chatterjee, 2020; Shariati et al., 2020). However, they have not yet been employed for the analysis of the VIF, and their performances depend on the type of analysis they are used for. PSO is an artificial algorithm inspired by the behavior of swarm particles, and was developed by Eberhart (1995); it is also known as a metaheuristic technique to solve optimization problems (Tian et al., 2018; Haidar et al., 2021). The PSO technique considers a population with a random position in an environment with a velocity; for each particle belonging to the population, the model computes the position and the new velocity of the particles, and ends the computation when the criteria are attained, which is the best position and the fitness of particles.

GA, which is an evolutionary algorithm, is also a heuristic method inspired by Charles Darwin's theory of natural evolution and is used to solve the problem of optimization. The GA technique considers a sample of the population with individuals defined by their genes and consequently their chromosomes. Chromosomes of these individuals through crossover can mutate until they reach the non-evaluative level with a given generation; this evaluation is based on the fitness, which is considered as the reverse of the cost function or the loss function (Deshwal et al., 2020; Bardeji et al., 2020). The performances of ANN trained with gradient descendant and PSO-ANN and GA-ANN will be evaluated on datasets 1 and 2, and a comparative analysis will be focused on these datasets.

5. Proposed methodology

The proposed methodology to analyze the value of the information flow is based on the following chart (see Figure 4).

We will carry out an audit to obtain the value of the weight of each characteristic of information flow in some companies in Cameroon, which will help in the construction of our dataset. We propose a model for calculating the VIF based on the weighted characteristics of the information. From the observed inefficiencies of the model and the weighted value obtained from small- and medium-sized manufacturing companies, a suitable analysis of VIF will be conducted using ANN, PSO-ANN, and GA-ANN. This analysis is based on the structure of the ANN that will be built by means of the two scenarios that will be obtained from datasets 1 and 2. The details of the methodology are as follows.

5.1. Proposed model of the VIF for audit result analysis

From the definition of the CIF that constitutes an information flow and considering the weight attributed to the sub-characteristics of information flow in the audit, W is defined as the weight attributed to each sub-characteristic. These sub-characteristics are modeled as presented in Eqs. (9), (10), (11), (12), and (13).

The type of information flow:

$$TY = \frac{W(t_{ij})}{W_{TYT}} \begin{cases} W(t_{ij}) \text{ the weight of each component of the type of information} \\ W_{TYT} \text{ the total weight of the type of information} \end{cases} \tag{9}$$

The dimension of information:

$$DI = \frac{W(x_{ij})}{W_{DIm}} \begin{cases} W(x_{ij}) \text{ the weight of the dimension of information} \\ W_{DIm} \text{ the maximum weight of the dimension of information} \end{cases} \tag{10}$$

The direction of information:

$$DRI = \frac{W\left(y_{ij}\right)}{W_{DRIm}} \begin{cases} W\left(y_{ij}\right) \text{ the weight of the dimension of information} \\ W_{DRIm} \text{ the maximum weight of the dimension of information} \end{cases} \tag{11}$$

The parameters of information flow:

$$PrI = \frac{\sum_{j=1}^{4} W\left(p_{ij}\right)}{W_{PrIT}} \begin{cases} W\left(p_{ij}\right) \text{ the weight of each component of the parameters of information} \\ W_{PrIT} \text{ the total weight of the parameter of information flow} \end{cases} \tag{12}$$

The quality of information:

$$QI = \frac{\sum_{j=1}^{4} W\left(q_{ij}\right)}{W_{QIT}} \begin{cases} W\left(q_{ij}\right) \text{ the weight of each component of the quality of information} \\ W_{QIT} \text{ the total weight of the quality of information} \end{cases} \tag{13}$$

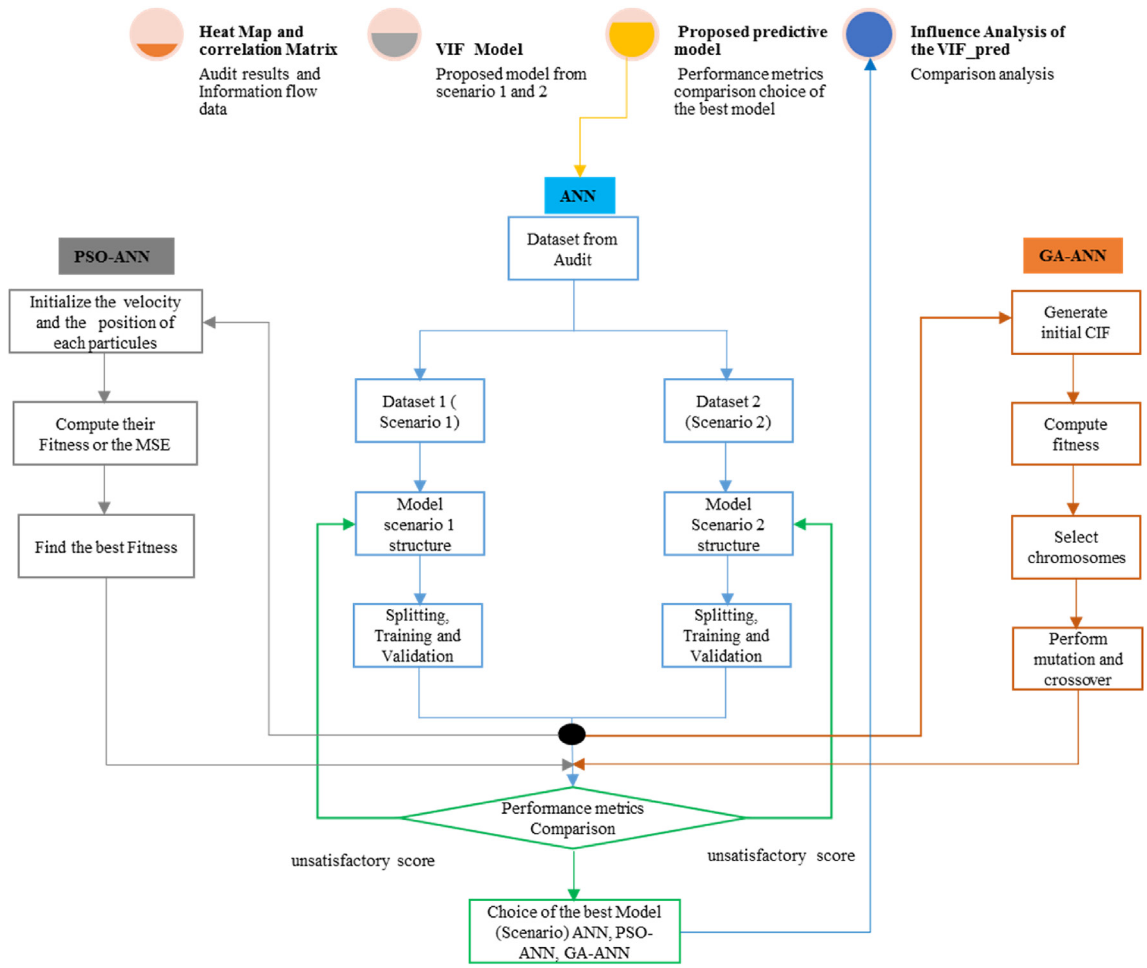


Figure 4. Proposed methodology for the analysis of the VIF using the ANN, PSO-ANN, and GA-ANN models.

If N is the number of characteristics (five in this case) of information flow containing the information flow, the VIF that we want to obtain for every information flow arriving in the system is described by Eq. (14).

$$VIF = \frac{TY + DI + DRI + PrI + QI}{N} \tag{14}$$

The parameters of information flow, such as complexity, velocity, volatility, and viscosity, were not considered in previous studies because they do not optimize the circulation of information flow and because volatility and viscosity have a negative impact on the velocity of information, although complexity remains neutral. Maximizing this information is important. For this case,

The optimum VIF can be given by maximizing Eq. (15):

$$\begin{cases} TY \leq 1 \\ DI \leq 1 \\ DRI \leq 1 \\ PrI \leq 1 \\ \dfrac{TY + DI + DRI + QI}{5} \leq 1 \end{cases} \tag{15}$$

Therefore, the CIF that minimizes or maximizes the VIF in SF operations must be examined. Examining the CIF that maximizes or minimizes VIF is a challenging task. Therefore, we will focus on machine learning methods that will help analyze this model and give us the characteristics that maximize or minimize the VIF by considering the information received from the audit.

5.2. Audit results for small- and medium-sized companies in Cameroon

After analyzing the questionnaires given to the workers in companies, we found that the type of information flow for the direct case can define the information concerning the raw materials of the product to be manufactured. It can also imply the technical design and various descriptions of the drawing and the quality materials of the raw material to be processed. For the indirect case, we can obtain information resulting from machines and operators. Therefore, we come to the following first hypothesis: **Hypothesis H_1**, the production processes of raw material may start when the information about the raw material (direct information) is available to the operator (machine, human) and can be 0 or 1, 0 when it does not exist, and 1 when it does. Indirect information can also be valued if there is information related to maintenance, canceled orders by a customer, production errors, shortage of energy, and availability of operators (human or machine). The more indirect information increases, the more the production processes can be stopped, but it cannot have a major influence on the VIF. Similarly, it can be binary as direct information.

Concerning the dimension of information flow, in this case, the scale values of Tomanek and Schröder (2017) were considered. The second hypothesis is presented as **Hypothesis H_2:** information flow can only come from one dimension and not more than two to avoid disruptions and conflicts at the level of the receiver. When information is given at the same time on two supports, it may be conflictual, and if it is conflictual, it will not add any value to the process. From the direction of information coming from the audit response, the dimension of information flow influences the value of the information flow when information is shared with operators or machines.

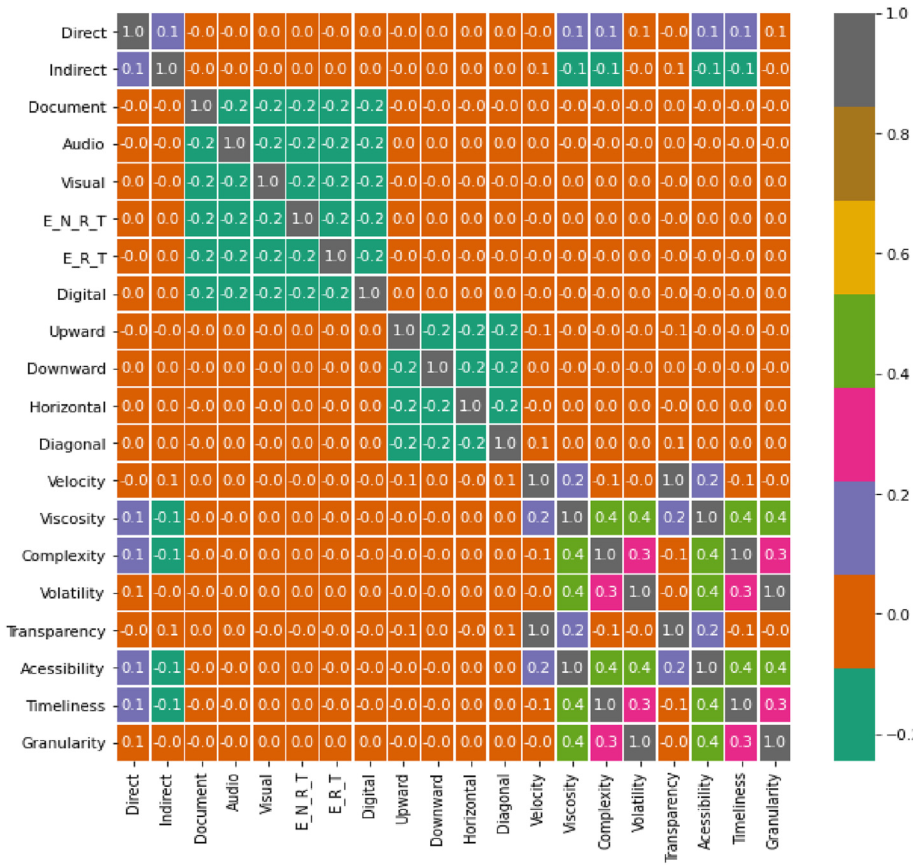


Figure 5. First scenario correlation matrix of information flow.

The third hypothesis can be presented as **Hypothesis H_3**: information flow can only come from one direction and not from more than two directions to avoid conflict in decision-making by human operators or machines. According to the importance of information from one level direction to another, the weighted value of the direction of information flow varies from 0 to 1. Likewise, the information from the manager to the operator (the machine) cannot certainly have the same weight as the one going from the operator to the manager. The information flow direction can vary from 0 to 1.

The parameters of information flow, which are velocity, viscosity, complexity, and volatility, vary from 0 to 1 according to the level of influence of each of the characteristics. Thus, **Hypothesis H_4** states that an information flow value will be low when volatility and viscosity appear in the flow, even if the complexity depends on what or who receives the information.

Concerning the quality of information flow, all sub-characteristics vary from 1 to 0. According to the answers received from the audit, the transparency and granularity of information flow depend on the personnel appreciation of operators, so its weighted value can vary from 0 to 1. However, for machines, it retains a binary value of 0 or 1. The timeliness and accessibility of information flow can also vary from 0 to 1 according to the appreciation of operators when the information is accessible and arrives timely. However, for machines, we have binary values. The hypothesis that arises from this is **Hypothesis H_5**, which states that VIF cannot exist if an information flow is not accessible or not timely (value nearly equal to 1 for operators and 1 for machines), and granularity and transparency values must be at least superior to 0.5. It should be noted that the parameters of information flow have the same evolution as the quality of information flow.

From hypotheses **H_1, H_2, H_3, H_4, and H_5**, the first scenario of the analysis arises wherein the information flow is influenced negatively by its parameters. This is the case of poorly industrialized countries, and from **H_1, H_2, H_3, and H_5**, the second scenario of analysis arises wherein the problems of volatility, complexity, and viscosity of information have been solved in the company of developing countries. Using these five hypotheses, scenarios 1 and 2 can be adopted, and the corresponding datasets 1 and 2 of scenarios 1 and 2, respectively, can be obtained from the data collected from the audit.

The dataset obtained from the audit response received according to the first and second scenarios is the cause of the correlation matrix heat map, as presented in Figures 5 and 6, respectively. The heat map and correlation matrix present the dependencies that exist between the sub-CIF after an audit.

It can be observed that from the parameters of information flow such as viscosity, complexity, and volatility which have a negative impact on the circulation of information flow have a correlation coefficient of 0.4 with the quality of information, and their impact on information flow is not negligible. The parameters and the quality of information flow also have the same correlation with the direction and dimension of information flow, which is 0, meaning that the parameters and quality of information flow have a stable or constant and equal evolution when the direction and dimension of information flow are concerned.

From the correlation matrix and the hypothesis mentioned above, we mathematically define a proposed model of information flow characteristics that will help us to train the ANN, PSO-ANN, and GA-ANN.

5.3. Information flow modeling from the obtained dataset

The information flow is a matrix where M_{IF} (X_k, $1 \leq k \leq 4$), which represents that all the four components of the CIF, and all the CIFs are not dependent on each other. Each of X_k matrices has a t_{ij}, x_{ij}, y_{ij}, q_{ij} component.

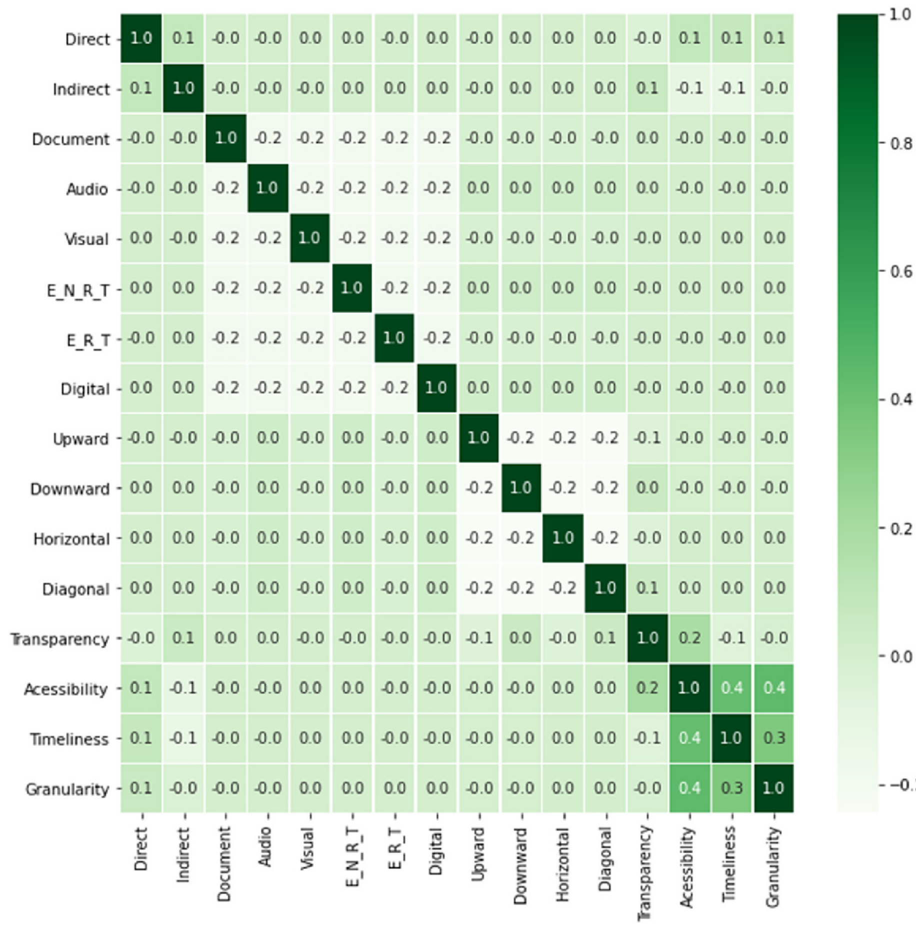


Figure 6. Second scenario correlation matrix of information flow.

$$\text{For each, } X_k = \begin{cases} X_1\ ,\ 1 \le Dimension \le 6 \\ X_2\ ,\ 1 \le Direction\ \le 4 \\ X_3,\ 1 \le Quality \le 4 \\ X_4\ ,\ 1 \le Type \le 2 \end{cases} \tag{16}$$

The type of information has as component, $X_1 = (Direct\ ,\ Indirect) = (t_{11}, t_{12})$ (17).

The dimension of information has as components,

$$X_2 = (Document\ ,\ Audio,\ Visual\ ,\ ENRT,\ ERT,\ Digital) = \begin{bmatrix} x_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & x_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & x_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & x_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & x_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & x_{66} \end{bmatrix} \tag{18}$$

The direction of information flow has as components,

$$X_3 = (\ Upward\ , Downward\ , Horizontal\ ,\ Diagonal) = \begin{bmatrix} y_{11} & 0 & 0 & 0 \\ 0 & y_{22} & 0 & 0 \\ 0 & 0 & y_{33} & 0 \\ 0 & 0 & 0 & y_{44} \end{bmatrix} \tag{19}$$

The parameters of information flow have as components,

$$X_4 = (\ Velocity\ , Viscosity\ , Complexity\ , Volatility\ ) = (\ p_{11}\ ,\ p_{12}\ , p_{13}\ ,\ p_{14}) \tag{20}$$

For quality, we will not consider the fifth sub-characteristic, which is the cost of information flow in this study, because information quality contains four components.

Quality of information flow has as components,

$$X_4 = (Transparency,\ Accessibility,\ Timeliness,\ Granularity) = (\ q_{11}\ ,\ q_{12}\ , q_{13}\ ,\ q_{14}\ ) \tag{21}$$

A shared information flow on the SF for product manufacturing or delivery has the following matrix, which is given by Eq. (22).

$$M_{IF} = (X_1,\ X_2,\ X_3, X_4) \tag{22}$$

6. Results and discussion

6.1. Comparison of the predictive analysis ANN models of the first and second scenario

The ANN model that we trained for the analysis of the information flow for the first and second scenarios has the following criteria, as shown in Tables 3 and 4. The splitting of datasets 1 and 2 corresponding to scenarios 1 and 2 was as follows: 80% and 20% for the training and validation, respectively.

The ReLU activation function was used for the hidden layers because the output of the hidden layers must deliver a positive output from the negative input. The sigmoid and hyperbolic tangent activation functions could not be used here because of the positive value of the VIF and the

Table 3. Criteria of the ANN model for the first scenario.

Dataset (Number of data)	Number of Inputs	Number of hidden layers	Number of nodes of the hidden layers	The activation function of hidden layers	The activation function of the output	Output (VIF_Pred)
1249038	20	3	5	ReLu	Linear	1

Table 4. Criteria of the ANN model for the second scenario.

Dataset (Number of data)	Number of Inputs	Number of hidden layers	Number of nodes of the hidden layers	The activation function of hidden layers	The activation function of the output	Output (VIF_pred)
1003782	16	3	4	ReLu	Linear	1

Table 5. Performance metrics of the ANN in the first scenario.

Computing Average time (s)	MAE	MSE	RMSE	Coefficient of determination
1203	0.0073	0.000715	0.03265	0.987

Table 6. Performance metrics of the ANN in the second scenario.

Computing Average time (s)	MAE	MSE	RMSE	Coefficient of determination
962	0.00656	0.00043	0.0216	0.9948

Table 7. Simple presentation of the data used for PSO-ANN and GA-ANN.

Optimizer algorithms	Population	Inputs	Output	Generation
PSO	Dataset 1 and Dataset 2	CIF	VIF	None
GA	Dataset 1 and Dataset 2	CIF	VIF	250

vanishing gradient problem. A linear function was used for the output (VIF) to ensure the continuity of the positive output data resulting in the last hidden layer.

The performance metrics for the ANN model of the first and second scenarios are listed in Tables 5 and 6. The accuracy values show that the model has been trained and has the ability to predict the VIF at 99% for both scenarios, but the second one is more accurate, with computing average times of 1200 s and 960 s for the first and second scenarios, respectively, and with mean absolute errors of 0.00731 and 0.00656, respectively.

The coefficients of determination for scenarios 1 and 2 are almost 0.99 and 0.98, respectively. This shows that the VIF, which is a function of all the characteristics of information, has successfully been used for the training of ANN and validation for scenarios 1 and 2.

The training of the ANN model concerning the MSE of the first scenario until 58 epochs was not 100% accurate; meanwhile, for the second model, the training was almost 100%. For the validation of the model, the MSE was almost equal to zero after 58 epochs for the first scenario, whereas for the second, the MSE was almost equal to zero for all epochs. This is the reason why the MSE loss function curve of the second scenario is smoother than that of the first one, as shown in Figures 7 and 8, which occurs because of the presence of the parameters of information flow in the first scenario, namely volatility, complexity, and viscosity.

The visualization of residuals of the first scenario (Figure 9) shows how the error in the prediction is mostly -0.2 and 0.2, and it occurs mostly when VIF_Pred is between 0.5 and 0.7; therefore, the model had to be improved. For the second (Figure 10) scenario, the prediction error is less than that for the first scenario because it ranges from the interval -0.2 to 0.1, although it also occurred when VIF_Pred was between 0.5 and 0.7.

The analysis of the MAE history for the first scenario shows that the value of the MAE tends towards 0 when the number of epochs is equal to or greater than 20, and for the second scenario, the value of MAE tends towards 0 when the number of epochs is equal to or greater than 5, as shown in Figures 11 and 12. In other words, the model of the second scenario makes more accurate predictions than the first because of the absence of disruptive information in information flow sharing. Likewise, as shown in Figures 13 and 14, the previous analysis results of the loss function metrics for training and validation of the ANN show that the ANN of the second scenario, concerning MAPE history, is more acceptable than that for the first one.

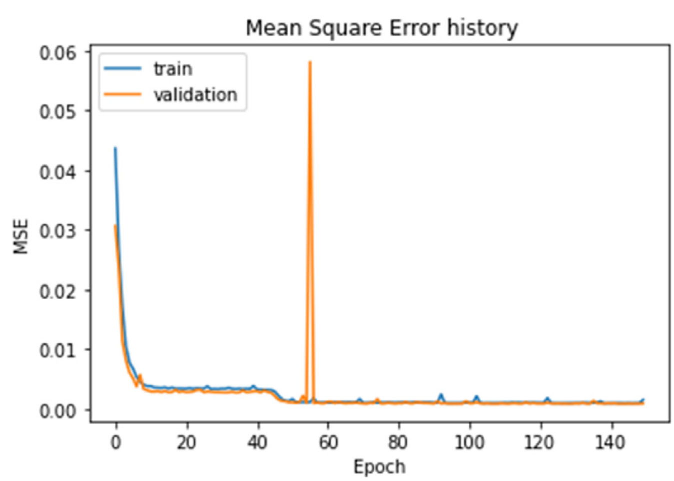


Figure 7. Model loss of the first scenario.

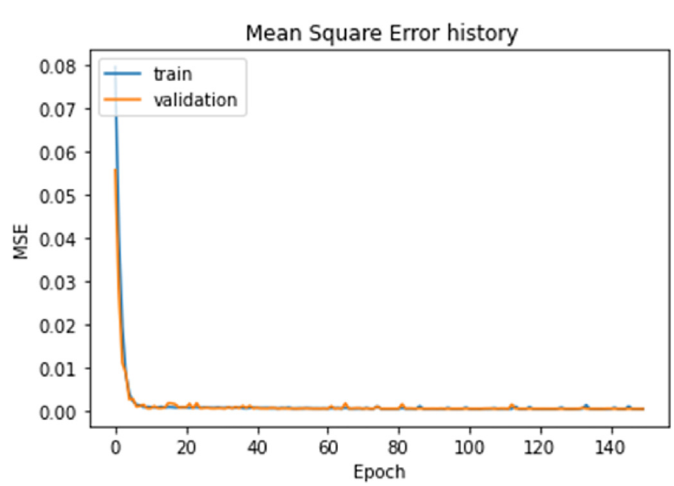


Figure 8. MSE history of the second scenario.

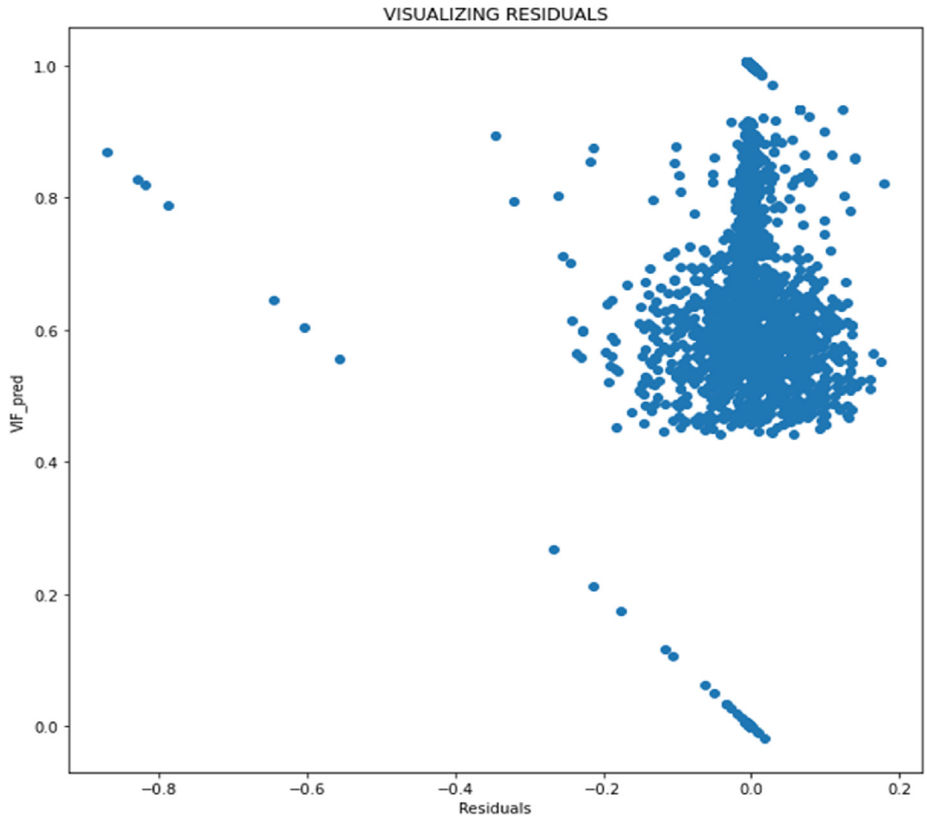


Figure 9. Visualizing residuals of the first scenario.

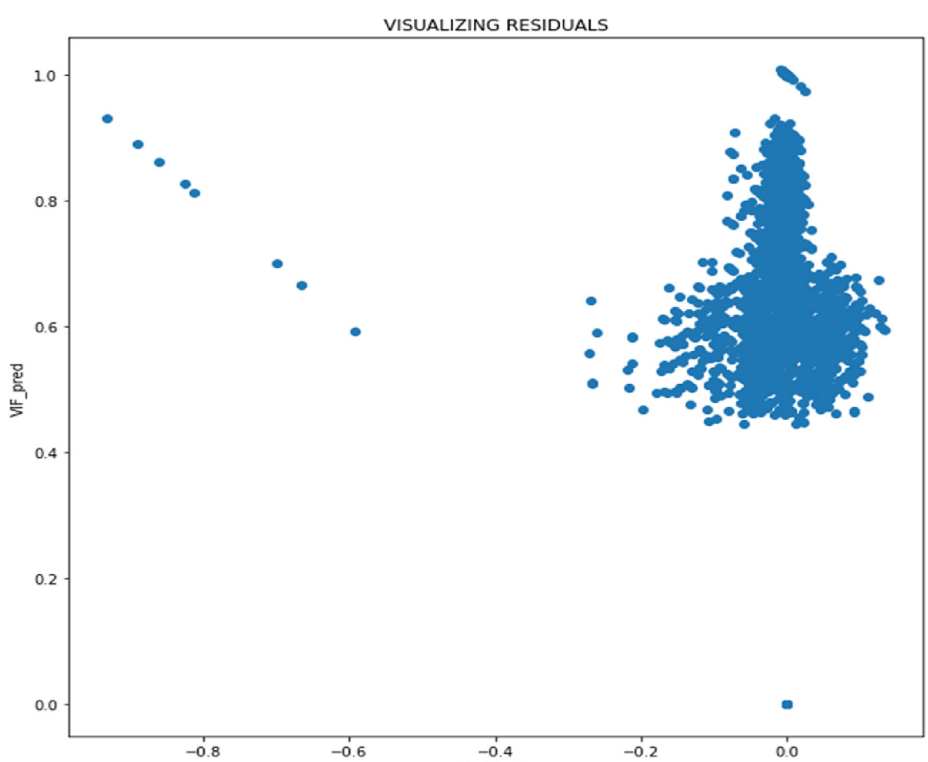


Figure 10. Visualizing residuals of the second scenario.

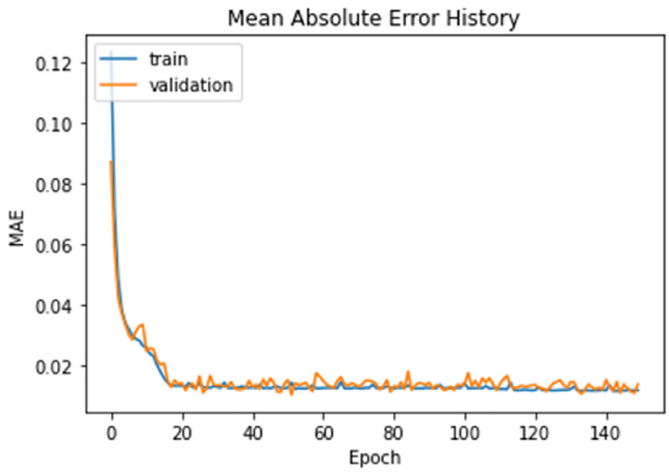


Figure 11. MAE history of the first scenario.

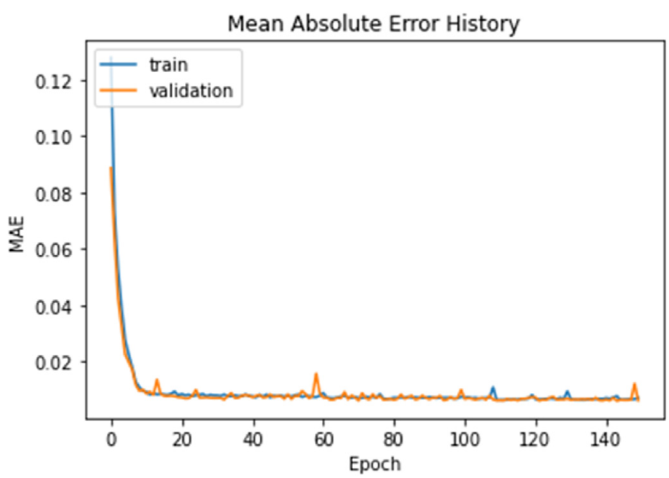


Figure 12. MAE history of the second scenario.

Based on the comparison of the performance metrics of the models of scenarios 1 and 2 and considering the loss function expressed by the different types of curves, a suitable model that should be considered is the model of the second scenario.

The Huber loss of the first and second scenarios Figures (15 and 16) is smaller than the values of MAE and MSE of the first and second scenarios because the Huber loss function is less sensitive to outliers than the MAE. From the two scenarios of the Huber loss function, the Huber loss function of the second scenario has a more optimized (minimized) error than that of the first one. It also justifies that, when the parameters of information flow are not included in the information flow for the production process, the predicted model will have a show better prediction with a tremendous minimization of errors.

The adopted ANN model for the next analysis of the VIF is the model of the second scenario; therefore, it can be used for the analysis of VIF in the SF that does not have information disruptions. Even the predictive ANN model of the first scenario can be used on the SF where there are information disruptions in the production process.

6.2. Comparison of the optimized ANN with PSO and GA for scenarios 1 and 2

Based on the ANN model presented in Tables 3 and 4 and the scores in Tables 5 and 6, the PSO and GA were used to optimize the loss function of the ANN model between the predicted and actual values of information flow, as well as the score of each scenario model.

To analyze the performance of PSO-ANN and GA-ANN, the criteria listed in Table 7 were considered in the computing program.

For PSO-ANN, the work was based on the MSE represented by the cost function when the hidden nodes in ANN were activated with the Tanh (PSO-ANN-1) function and with the ReLu (PSO-ANN-2) function. For GA-ANN, the adopted metrics are the MSE and coefficient of determination. The performance metrics obtained for scenarios 1 and 2 for the ANN trained by PSO and GA are presented in Figure 17 and Table 8.

Compared with the ANN metrics presented previously, the PSO-ANN-1 (using Tanh) and PSO-ANN-2 (using ReLu) accurately analyzed the VIF, and their obtained MSE values were 0.0005 and 0.00012, respectively; these values were less than 0.00071 for the first scenario. Therefore, for the first scenario, PSO-ANN, with excellent performance, is the one in

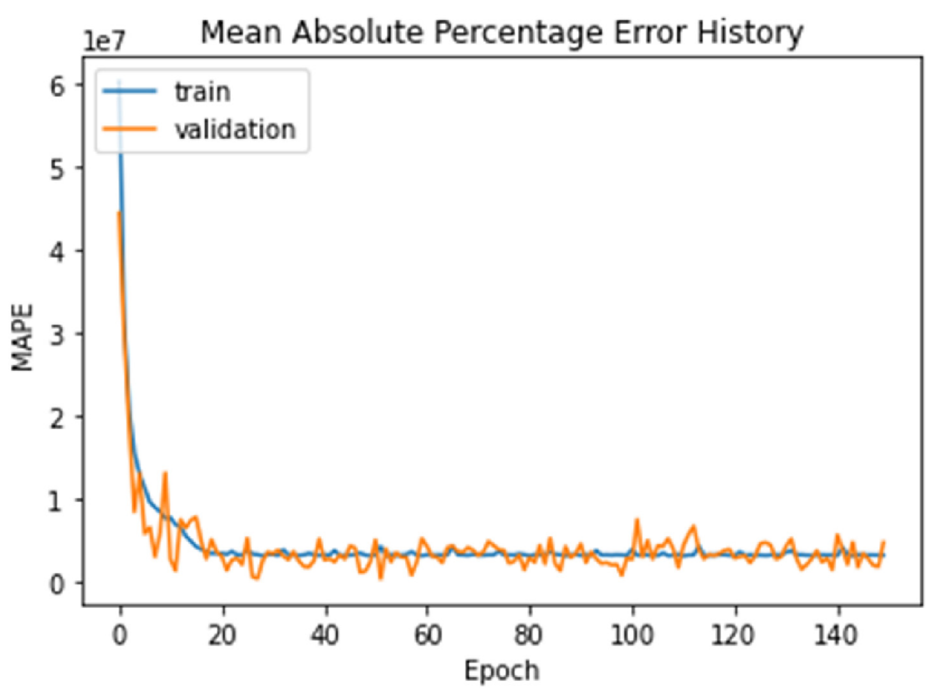


Figure 13. MAPE history of the first scenario.

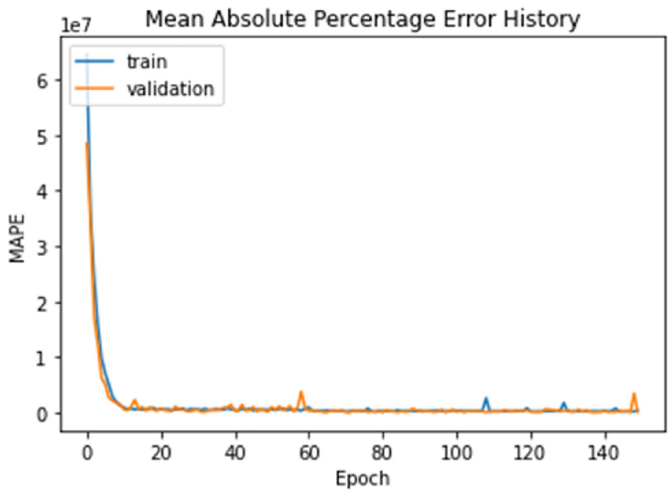


Figure 14. MAPE history of the second scenario.

Figure 15. Huber loss history of the first scenario.

which the hidden layers are activated with the ReLu function. For the second scenario, the obtained MSE values of PSO-ANN-1 and PSO-ANN-2 are 0.00055 and 0.00011, respectively, which are much higher than the obtained value of 0.00043 for ANN. The PSO-ANN-2 for the second scenario had the highest MSE of the PSO-ANN. The history of the cost

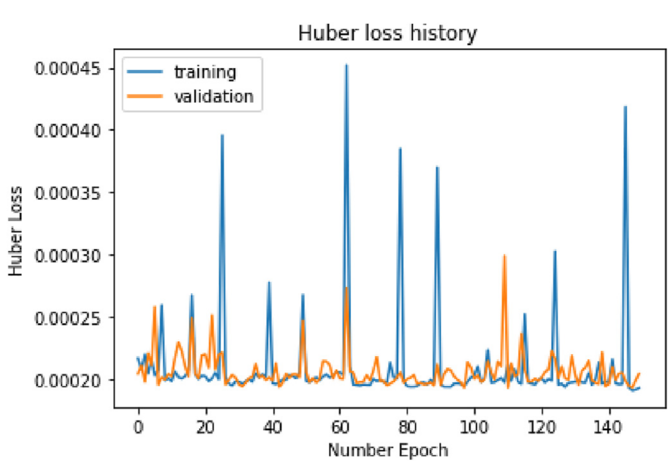


Figure 16. Huber loss history of the second scenario.

Table 8. Comparison of GA-ANN performance for scenarios 1 and 2.

	MAE	MSE	Coefficient of determination
Scenario 1	0.355	0.2923	0.6018
Scenario 2	0.375	0.2560	0.4031

function of the first and second scenarios using the Tanh and ReLu functions is presented in Figures 18 and 19.

Figure 18 shows that the cost function of the PSO-ANN for the Tanh and ReLu curves of the first scenario is the same from the 60th iteration. Likewise, Figure 19 shows that the PSO-ANN model with the Tanh and ReLu curves of the second scenario is almost the same from the 14th iteration. PSO-ANN is also a well-optimized model to be used to analyze the VIF when compared with ANN.

Using GA to train the ANN, we obtained the performance metrics presented in Table 8.

The loss metrics of GA-ANN show that the MSEs for the first and second scenarios are 0.2923 and 0.2560, respectively. These results show that the prediction is not accurate when compared with the obtained values of ANN and PSO-ANN. Figures 20 and 21 show that the fitness of the GA-ANN until the 250 generations is not above 3, which is why the MAE is very high. This result shows that the GA-ANN is not sufficient to analyze the VIF with accuracy when the CIF representing the total population take their values (float) in the unit interval.

From the above comparison, PSO-ANN-2 is the best model that can be used to analyze the VIF when trying to optimize the ANN for scenarios 1 and 2. Each company where the information flow is based on scenarios 1 and 2 can use the presented ANN or PSO-ANN to analyze to predict the VIF. The performance metrics of PSO-ANN-2 for the second scenario seem to be the one of a deep learning model used for regression based on other related deep learning performance (MSE) models used for regression in the literature as compared to the one Hossain et al. (2018), Kim et al. (2020), Shetty et al. (2021) with MSE values of 0.002, 0.00393, and 0.034 respectively.

6.3. Comparison of the analyzed VIF of the first and second scenario using PSO-ANN

The predicted VIF of the ANN is also equal to the predicted VIF of the PSO-ANN because of the light difference between their loss functions. We extracted a sample of some information flow that can be found on the SF, and these are presented in Tables 9 and 10. The analysis of the sub-characteristics depending on the VIF helps to accurately understand the novelty of this study, proving the reality of the SF in developing countries.

In Tables 9 and 10, X^0 shows that, if on a SF where a raw material (t_{11}, which is of good quality and that we have all its properties) is to be processed by an automated machine (x_{55}) monitored by another company or technician which is not from the company (y_{44}) and if the procedures of the product processing are accurately transparent, accessible, and timely without being totally detailed or granular at 30%, the value of information for the production process to be performed on the SF is 0.84773 and 0.72075 for the first and second scenarios, respectively, according to our PSO-ANN.

X^1 can be referred to as a raw material already in the process on an automated machine, which has the same conditions as mentioned above. X^1 shows that we are on a SF where a raw material (t_{11}, which is of good quality and has all its properties) is in a process being operated by a human operator (x_{22}) on a non-automated machine. A human operator receives a signal (audio) or wordily instruction from a supervisor that he does not underestimate (y_{22}) and who is in another (or before) job post; from a machine announcing a breakdown failure or from the approval of good manufacturing operation due to chip (t_{12}). The audio information is transferred in a noisy environment, assuming that the audio signal or the

	PSO-ANN-1 Scenario 1	PSO-ANN-2 Scenario 1	PSO-ANN-1 Scenario 2	PSO-ANN-2 Scenario 2
MSE	0.0005	0.00012	0.00055	0.00011

Figure 17. PSO-ANN performance for scenarios 1 and 2.

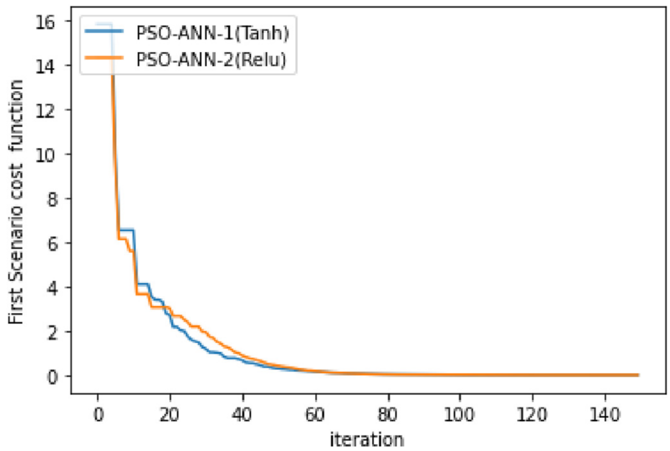


Figure 18. Cost history of PSO-ANN-1 and PSO-ANN-2 of scenario 1.

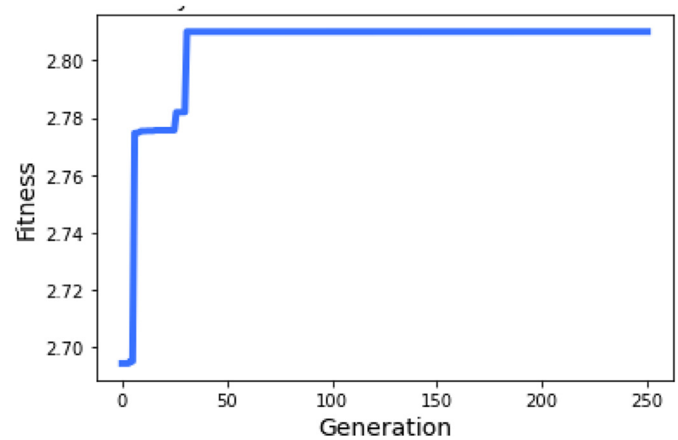


Figure 20. Fitness history of GA-ANN of scenario 1.

wordy instruction to the human operator is accurately transparent, accessible, and timely without being detailed or granular at 30%. The value of information for this described production process on the SF is 0.80106 and 0.80236 for the first and second scenarios, respectively, according to our PSO-ANN.

X^2 is a case where raw material (t_{11}, which is of good quality and that we have all its properties and the processing techniques) is to be processed by the Internet of Things associated with the machine (x_{66}, Industry 4.0), which performs self-monitoring because the program is from an external company (y_{44}), and if the procedures of the product processing are fairly transparent, accessible, and timely without being detailed or granular at 10% because of its digitalization, the value of information for the production process to be performed on the SF is 0.9340 and 1.0131 for the first and second scenarios, respectively, according to our PSO-ANN.

X^3 illustrates a situation where raw material is to be processed by a semi-automated machine (x_{44}) monitored by a technician who is not from the company. If the machines are monitored by an external automated program (y_{44}) and if the procedures of the product processing are accurately transparent, accessible, and timely without being detailed or granular, the value of information for the production process to be performed on the SF is 0.0017 and 0.7054 for the first and second scenarios, respectively, according to our PSO-ANN. Notably, we have a low value for the first scenario because, in this case, the volatility of information can be due to the obsolescence of the automated machines or flexible cable.

X^4 describes a job in which raw material is processed by a human operator (reading technical drawing information about or paper procedures) (x_{11}) on a machine. If the human operator receives maintenance information concerning the lack of availability of the downward machine to perform the job of the downward post (y_{11}) and if the procedures of the product processing are accurately transparent, accessible, and timely without any details, the value of information for the production process to be performed on the SF is 0.7748 (no volatility, no complexity, viscosity reduced to null) and 0.6475 for the first and second scenarios, respectively, according to our PSO-ANN.

X^9 illustrates a production sequence in which raw material is processed by a human operator (x_{33}) viewing information on a board or Kanban board to run the production. If the information that he is viewing is at the same job post (y_{33}) and if the procedures of the product processing are 80% transparent, accessible, and timely without totally detailed or granularity is at 30%, then the value of information for the production process to be performed on the SF is 0.5208 (complex but not volatile though 30% of viscosity) and 0.688 for the first and second scenarios, respectively, according to our PSO-ANN.

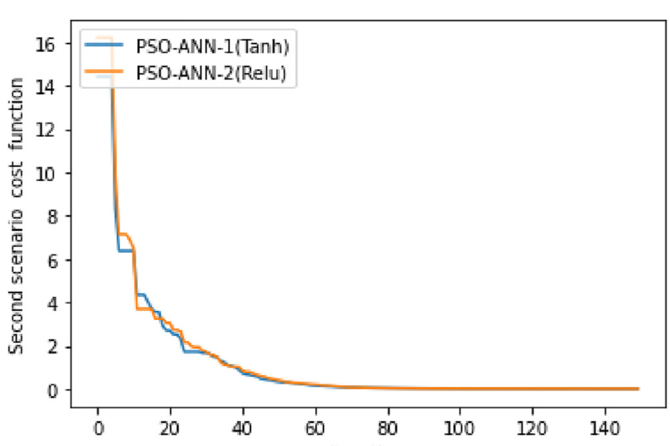


Figure 19. Cost history of PSO-ANN-1 and PSO-ANN-2 of scenario 2.

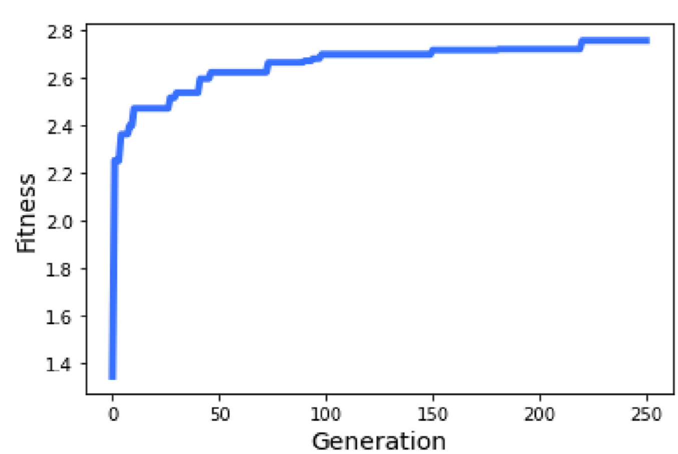


Figure 21. Fitness history of GA-ANN of scenario 2.

Table 9. Extracted values of the analysis of the VIF of the first scenario.

CIF of the first scenario																					Output
X	t_{11}	,t_{12}	x_{11}	x_{22}	x_{33}	x_{44}	x_{55}	x_{66}	y_{11}	y_{22}	y_{33}	y_{44}	p_{11}	p_{12}	p_{13}	p_{14}	Q_{11}	Q_{12}	Q_{13}	Q_{14}	VIF_Pred
X^0	1	0	0	0	0	0	0.8	0	0	0	0	0.9	1	0	0	0	1	1	1	0.7	0.84773
X^1	1	1	0	0.4	0	0	0	0	0	0.6	0	0	0.8	1	0	0	1	1	1	0	0.80106
X^2	1	0	0	0	0	0	0	1	0	0	0	0.3	1	0.1	0	0	0.7	1	1	0.1	0.9340
X^3	1	0	0	0	0	0.6	0	0	0	0	0	1	1	0	0.8	1	1	1	1	0	0.0017
X^4	1	0	0.2	0	0	0	0	0	0.9	0	0	0	1	0.1	0	0	1	1	1	0	0.7748
X^5	1	0	0.2	0	0	0	0	0	0	0	0	1	1	0.3	0.2	0.5	1	1	1	0	0.5618
X^6	1	0	0	0	0	0	0	1	0	0	1	0	1	0	0	0	1	1	1	0	0.93402
X^7	1	0	0	0	0	0.6	0	0	0	0	0.8	0	1	1	0	0	1	1	1	0.9	0.6823
X^8	1	0	0	0	0.4	0	0	0	1	0	0	0	1	1	0	0	1	0.4	1	0.8	0.0008
X^9	1	0	0	0	0.4	0	0	0	0	0	1	0	1	0.3	0	0	0.8	1	1	0.7	0.5208
X^{10}	1	1	0	0	0.4	0	0	0	0	0	0	0.7	1	0	0	0	1	1	1	0	0.5208
X^{11}	1	1	0	0.4	0	0	0	0	0	0.3	0	0	1	1	1	0	1	1	1	0	0.5208
X^{12}	1	0	0	0	0	0	0.8	0	0.8	0	0	0	1	1	1	0	1	1	0.2	0.9	0.5208
X^{13}	1	0	0	0	0	0	0	1	0	1	0	0	1	1	1	0.8	1	0.1	1	0	0.5208
X^{14}	0	1	0	0	0	0	0	1	0	0	0	1	1	0.2	1	0.2	1	1	1	0.9	0.5208
X^{15}	1	0	0	0	0	0.6	0	0	0	0.7	0	0	0.6	1	1	0.4	1	1	1	0.8	0.5208
X^{16}	0	0	0	0	0	0	0	1	0	1	0	0	1	1	1	1	1	1	1	1	0.5208
X^{17}	0	1	0.2	0	0	0	0	0	0	0	0.8	0	1	1	1	1	0.7	1	1	0	0.5208
X^{18}	1	0	0	0	0.4	0	0	0	0	1	0	0	1	0.5	1	1	1	0.9	1	0.5	0.5208
X^{19}	1	0	0	0	0	0	0.8	0	0.7	0	0	0	1	1	1	1	1	1	1	0.6	0.5208
X^{20}	1	1	0.2	0	0	0	0	0	1	0	0	0	1	1	1	0.6	1	1	1	0.6	0.8635

Table 10. Extracted values of the analysis of the VIF of the second scenario.

CIF of the second scenario																	Output
X	t_{11}	t_{12}	x_{11}	x_{22}	x_{33}	x_{44}	x_{55}	x_{66}	y_{11}	y_{22}	y_{33}	y_{44}	Q_{11}	Q_{12}	Q_{13}	Q_{14}	VIF_pred
X^0	1	0	0	0	0	0	0.8	0	0	0	0	0.9	1	1	1	0.7	0.72095
X^1	1	1	0	0.4	0	0	0	0	0	0.6	0	0	1	1	1	0	0.80236
X^2	1	0	0	0	0	0	0	1	0	0	0	0.3	0.7	1	1	0.1	1.0131
X^3	1	0	0	0	0	0.6	0	0	0	0	0	1	1	1	1	0	0.7054
X^4	1	0	0.2	0	0	0	0	0	0.9	0	0	0	1	1	1	0	0.6475
X^5	1	0	0.2	0	0	0	0	0	0	0	0	1	1	1	1	0	0.6590
X^6	1	0	0	0	0	0	0	1	0	0	1	0	1	1	1	0	1.0070
X^7	1	0	0	0	0	0.6	0	0	0	0	0.8	0	1	1	1	0.9	0.6780
X^8	1	0	0	0	0.4	0	0	0	1	0	0	0	1	0.4	1	0.8	0.0007
X^9	1	0	0	0	0.4	0	0	0	0	0	1	0	0.8	1	1	0.7	0.6880
X^{10}	1	1	0	0	0.4	0	0	0	0	0	0	0.7	1	1	1	0	0.6880
X^{11}	1	1	0	0.4	0	0	0	0	0	0.3	0	0	1	1	1	0	0.6880
X^{12}	1	0	0	0	0	0	0.8	0	0.8	0	0	0	1	1	0.2	0.9	0.6880
X^{13}	1	0	0	0	0	0	0	1	0	1	0	0	1	0.1	1	0	0.6880
X^{14}	0	1	0	0	0	0	0	1	0	0	0	1	1	1	1	0.9	0.6880
X^{15}	1	0	0	0	0	0.6	0	0	0	0.7	0	0	1	1	1	0.8	0.6880
X^{16}	0	0	0	0	0	0	0	1	0	1	0	0	1	1	1	1	0.6880
X^{17}	0	1	0.2	0	0	0	0	0	0	0	0.8	0	0.7	1	1	0	0.6880
X^{18}	1	0	0	0	0.4	0	0	0	0	1	0	0	1	0.9	1	0.5	0.6880
X^{19}	1	0	0	0	0	0	0.8	0	0.7	0	0	0	1	1	1	0.6	0.6880
X^{20}	1	1	0.2	0	0	0	0	0	1	0	0	0	1	1	1	0.6	0.7205

This analysis is then completed by comparing the VIF predicted for the first and second scenarios, as shown in Figure 22, where it can be seen that the maximum value of the predicted VIF is 1 and the minimum is 0 for both scenarios; however, for the first information flow X^0, the predicted VIF of the second scenario is greater than that of the first one. This confirms the technical aspect of information flow sharing, which aims to eliminate the occurrence of volatility and viscosity of information and sometimes the complexity of information which depends on the information receiver.

6.4. The predictive analysis system of the VIF for managers of SMEs in developing countries

From the above analysis, we demonstrated how the parameters of information flow negatively influenced the performance of the ANN and consequently the VIF. The MSE values of the ANN of the second scenario are not different from those of PSO-ANN-2 for the second scenario, so the predicted VIF will be the same for both models because it is a deep learning approach. Good characteristics that enable appropriate

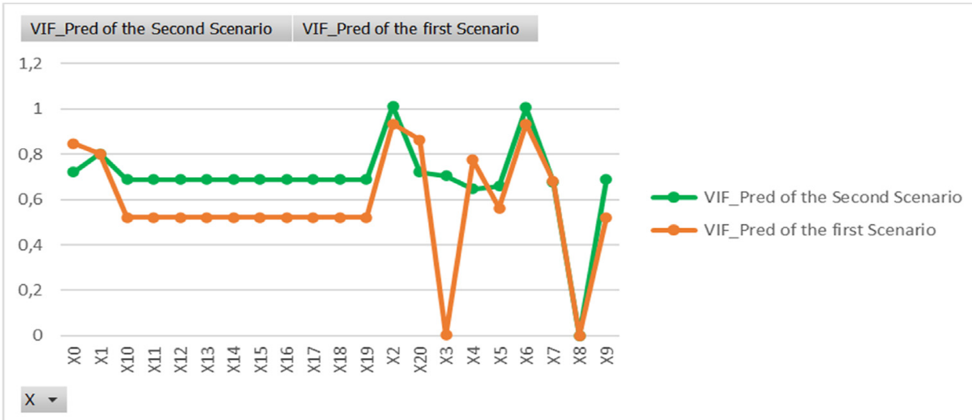


Figure 22. Comparison of the predicted VIF of the first and second scenario.

information sharing are those of the second scenarios, which are the type, dimension, direction, and quality of information flow. In a SF operation where production order is released, a machine can have a breakdown or failure and a human operator could be tired or wounded; furthermore, there can be power cuts and missing tools and quality control on the chip (waste of production). Alternatively, in a SF where we have any order information, the VIF will influence the production time and the quality product based on the decision-making of the operator (human, machine), when the information to be received is analyzed by the ANN. This should be how the SF information management system corresponding to developing countries ought to look like.

When the information flow arrives on the SF for production processes of raw material or if during the operations processes, and information

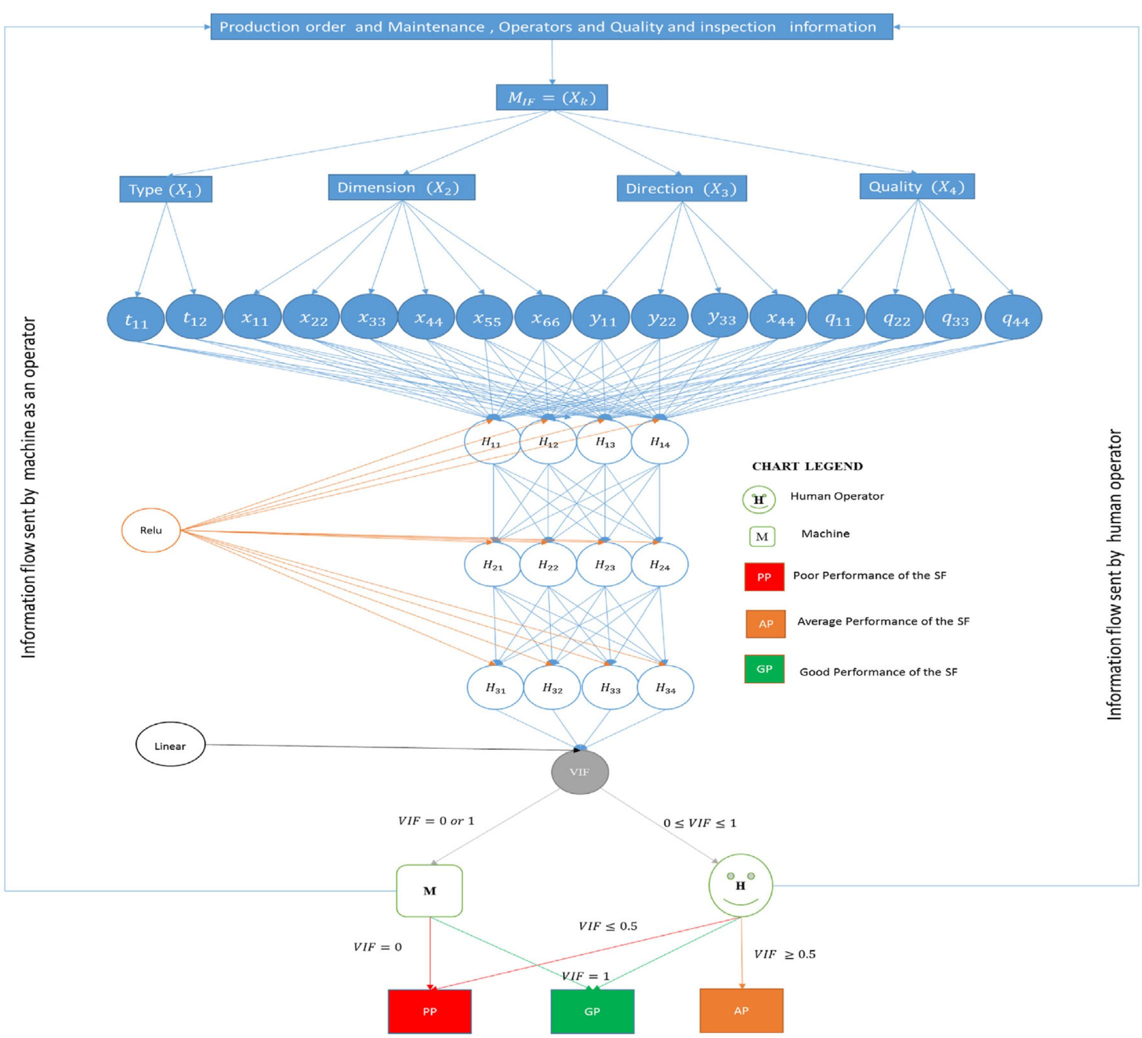


Figure 23. Predictive system analysis of the VIF with a proposed ANN Trained with PSO.

flow is released to the SF management information system, the PSO-ANN analyzes the information flow and gives its value. Alternatively, the PSO-ANN model may predict the VIF for a particular job in the production process. The VIF metrics can then be considered as such; when the VIF is less than 0.5, the system has to note the poor performance of the SF. If the VIF is greater than or equal to 0.5, the system will mention the good performance of the SF while indicating which characteristic of information has to be removed or improved to have an excellent VIF (VIF = 1), as presented in Figure 23. Tomanek et al. (2020) provided a value of 5 on a scale from 0 to 5 to the VIF when the SF was digitalized (Industry 4.0), but in this work, we considered the VIF analyzed to be from 0 to 1; this is just a matter of scale.

7. Conclusion and future research work

The MIF has always been the focus in the amelioration of production performance on the SF for customer satisfaction. SMEs in developing countries such as Cameroon are still lacking in performance improvement, due to poor information management on the SF. Previous studies on management of information flow have used methods of integrating process information, VAHM, quantification of information, and digitalization degree. The CIF that was considered with these methods was the dimension and quality of the information. This study aimed to develop a predictive analysis model of the VIF of the SF of developing countries using ANN, PSO-ANN, and GA-ANN by considering two scenarios. The first scenario included all CIF, and the second one did not include the parameters of information flow in the analysis. The results of the developed ANN model were as follows: the ANN is a suitable tool to analyze the VIF with a score of 0.99 for the loss function includes, 0.00656 for the mean absolute error (MAE), and 0.00043 for the MSE. To improve the performance, we used the PSO-ANN model. The results of PSO-ANN (MSE) for the first scenario were 0.0005 and 0.00012 when using Tanh and ReLu as activation functions. For the second scenario, the MSEs were 0.00055 and 0.00011 when using Tanh and ReLu as activation functions, respectively. The results obtained for GA-ANN for both scenarios were not satisfactory for this information flow analysis based on the following hypothesis: The best model for the analysis of the information was PSO-ANN 2 for the second scenario. The PSO-ANN 2 model was used to present a predictive analysis system of the VIF that will be of use to managers of developing countries. Knowing that the dataset obtained from the audit could not be the same in all countries, the PSO-ANN 2 can cover the small changes that may occur during the prediction, but not large ones. A job or a task that can be performed by an operator on the SF will henceforth have a VIF. How can the production supervisor schedule the different types of jobs or tasks based on the VIF and the due time? This question will be answered in future work.

Declarations

Author contribution statement

André Marie Mbakop: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Florent Biyeme: Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Joseph Voufo: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Jean Raymond Lucien Meva'a: Conceived and designed the experiments.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

We would like to thank the esteemed reviewers for their valuable and unique contributions by providing critical feedback and comments to improve the manuscript. We would like to thank the editors and editor-in-chief for their encouragement and framework for keeping the paper in shape.

References

Akpan, I.J., Udoh, E.A.P., Adebisi, B., 2020. Small business awareness and adoption of state-of-the-art Technologies in emerging and developing markets, and lessons from the COVID-19 pandemic. J. Small Bus. Enterpren. 1–18.

Aldahoul, N., Essam, Y., Kumar, P., et al., 2021. Suspended sediment load prediction using long short-term Memory neural network. Sci. Rep. 11, 7826.

Bardeji, S.F., Saghih, A.M.F., Pooya, A., Mirghaderi, S.H., 2020. Perishable inventory management using GA-ANN and ICA-ANN. Int. J. Procure. Manag. 13 (3), 347.

Benyahya, P., Macurová, L., 2021. Utilization of Shop floor management as a tool for communication and knowledge sharing in the framework of lean logistics-Case study. Serbian J. Manag. 16 (1), 267–276.

Busert, T., Fay, A., 2020. Information quality focused value stream mapping for the coordination and Control of production processes. Int. J. Prod. Res. 1–20.

Çınar, Z.M., Abdussalam, N.A., Zeeshan, Q., Korhan, O., Asmael, M., Safaei, B., 2020. Machine learning in predictive maintenance towards sustainable smart manufacturing in industry 4.0. Sustainability 12, 8211.

Deshwal, S., Kumar, A., Chhabra, D., 2020. Exercising hybrid statistical tools GA-RSM, GA-ANN and GA-ANFIS to optimize FDM process parameters for tensile strength improvement. CIRP J. Manufact. Sci.

Drigo, E.D.S., Rodriguez, J.L.M., Embiruçu, M., Filho, S.A., 2020. Development of a neuro-fuzzy system for assessing information management on the shop floor. In: IEEE Access, 8, pp. 207063–207075.

Durugbo, C., Erkoyuncu, J.A., 2014. Managing integrated information flow for industrial service partnerships: a case study of aerospace firms, product services systems and value creation. Procedia CIRP 16, 338–343.

Durugbo, C., Hutabarat, W., Tiwari, A., Alcock, J.R., 2011. Modelling collaboration using complex networks. Inf. Sci. 181 (15), 3143–3161.

Eberhart, K.,J., 1995. A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science. MHS'95), Nagoya, pp. 39–43.

Ferrero Bermejo, J., Gómez Fernández, J.F., Olivencia Polo, F., Crespo Márquez, A., 2019. A review of the use of artificial neural network models for energy and reliability prediction. A study of the solar PV, hydraulic and wind energy sources. Appl. Sci. 9 (9), 1844.

Göppert, A., Robert, L.M., Schmitt, H., 2021. Predicting Performance Indicators with ANNs for AI-Based Online Scheduling in Dynamically Interconnected Assembly Systems.

Habtamu, T., Megersa, O.D., 2021. Application of ANN and HEC-RAS model for flood inundation mapping in lower Baro Akobo River Basin, Ethiopia. J. Hydrol.: Reg. Stud. 36, 100855.

Haidar, A., Field, M., Sykes, J., Carolan, M., Holloway, L., 2021. PSPSO: a package for parameters Selection using particle swarm optimization. SoftwareX 15 (1-5).

Hossain, M.A., Karim, R., Thulasiram, R., Bruce, N.D.B., Wang, Y., 2018. Hybrid deep learning model for stock price prediction. In: 2018 IEEE Symposium Series on Computational Intelligence (SSCI).

Hosseinzadeh, A., Baziar, M., Alidadi, H., Zhou, J.L., Altaee, A., Najafpoor, A.A., Jafarpour, S., 2020. Application of artificial neural network and multiple linear regression in modeling nutrient recovery in vermicompost under different condition. Bioresour. Technol. 303, 122926.

Jabur, Z.T., Dawood, L.M., 2015. Analysis of information flow for job-shop production system. Eng. Technol. J. 33 (1), 223–236.

Jalaee, S.A., Lashkary, M., GhasemiNejad, A., 2019. The Phillips curve in Iran: econometric versus artificial neural networks. Heliyon 5 (8), e02344.

Kankar, P.K., Satish, C., Sharma, S.P., 2011. Harsha Fault diagnosis of ball bearings using machine learning methods. Expert Syst. Appl. 38, 1876–1886.

Kim, M., Lee, S., Kim, J., 2020. A Wide & Deep Learning Sharing Input Data for Regression Analysis. In: 2020 IEEE International Conference on Big Data and Smart Computing (BigComp).

Le, L.T., Nguyen, H., Dou, J., Zhou, J., 2019. A comparative study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for smart city planning. Appl. Sci. 9 (13), 2630.

Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., et al., 2018. Lead time prediction using machine learning algorithms: a case study by a semiconductor manufacturer. Procedia CIRP 72, 1051–1056.

Lu, Y., Xu, X., Wang, L., 2020. Smart manufacturing process and system automation – a critical review of the standards and envisioned scenarios. J. Manuf. Syst. 56, 312–325.

Mathiasen, B.J., Haas, D.H., 2021. Digitalization of shop floor management: in blissful ignorance of superfluous work. J. Indust. Integr. Manag. 333–352. World Scientific Publishing Company.

Mbakop, A.M., Voufo, J., Biyeme, F., Ngozag, L.A., Meva'a, L., 2021. Analysis of information flow characteristics in shop floor: state-of-the-art and future research directions for developing countries. Global J. Flex. Syst. Manag. 22, 43–53.

Mourtzis, D., Fotia, S., Boli, N., Ekaterini, V., 2019. Modelling and quantification of industry 4.0 manufacturing complexity based on information theory: a robotics case study. Int. J. Prod. Res.

Mourtzis, D., Fotia, S., Vlachou, E., 2017. Quantification of industry 4.0 manufacturing complexity based on information theory. In: Proceedings of 7th IESM Conference 2017, Saarbrücken, Germany. October 11–13.

Ngoungo, D., 2012. The Practice of Corporate Governance in companies receiving foreign direct investment in Cameroon. Res. World Econom. Res. World Econom. 3 (1), 2–19. Sciedu Press.

Pant, P., Chatterjee, D., 2020. Prediction of clad characteristics using ANN and combined PSO-ANN algorithms in laser metal deposition process. Surf. Interfaces 21, 100699.

Osisanwo, F.Y., Akinsola, J.E.T., Awodele, O., Hinmikaiye, J.O., Olakanmi, O., Akinjobi, J., 2017. Supervised machine learning algorithms: classification and comparison. Int. J. Comput. Trends Technol. (IJCTT) 48 (3), 128–138.

Pandey, K., Kumar, S., Malik, A., Kuroki, A., 2020. Artificial neural network optimized with a genetic algorithm for seasonal groundwater table depth prediction in Uttar Pradesh, India. Sustainability 12 (21).

Paul, J., 2020. SCOPE framework for SMEs: a new theoretical lens for success and internationalization. Eur. Manag. J.

Qiokata, V., Khan, M.G.M., 2014. Artificial neural network and regression models for predicting Fiji population. In: Asia-Pacific World Congress on Computer Science and Engineering.

Richter, A., Vodanovich, S., Steinhuser, M., Hannola, L., 2017. IT on the shop floor—challenges of the digitalization of manufacturing companies. In: Bled eConference, Bled, Slovenia, 18 June 2017–21 June 2017.

Sachdev, H.S., 2020. How to Choosing Number of Hidden Layers and Number of Hidden Neurons in Neural Networks. (12) Choosing Number of Hidden Layers and Number of Hidden Neurons in Neural Networks. LinkedIn. (Accessed 27 March 2021).

Shariati, M., Mafipour, M.S., Mehrabi, P., Ahmadi, M., Wakil, K., Trung, N.T., Toghroli, A., 2020. Prediction of concrete strength in presence of furnace slag and fly ash using Hybrid ANN-GA (Artificial Neural Network-Genetic Algorithm). Smart Struct. Syst. 25 (2), 183–195.

Sheela, G.K., Deepa, S.N., 2014. Review on Methods to Fix Number of Hidden Neurons in Neural Networks. Hindawi Publishing Corporation Mathematical Problems in Engineering, pp. 1–11.

Shetty, S.A., Padmashree, T., Sagar, B.M., Cauvery, N.K., 2021. Performance analysis on machine learning algorithms with deep learning model for crop yield prediction. In: Jeena Jacob, I., Kolandapalayam Shanmugam, S., Piramuthu, S., Falkowski-Gilski, P. (Eds.), Data Intelligence and Cognitive Informatics. Algorithms for Intelligent Systems. Springer, Singapore.

Sütöova', A., Seginakova', S., 2018. Improving information flow for decision making on product quality in the automotive industry. Qual. Innovat. Prosper. 22 (1), 71–78.

Thomas, A.J., Petridis, M., Walters, S.D., Gheytassi, S.M., Morgan, R.E., 2017. Two hidden Layers are Usually Better than One. In: Communications in Computer and Information Science, pp. 279–290.

Tian, H., Shu, J., Liu, H., 2018. The effect of ICA and PSO on ANN results in approximating elasticity modulus of rock material. Eng. Comput.

Tomanek, D.P., Schröder, J., Wirz, M., 2016. Value added heat map—a new method for the optimization of production space. In: 13th International Conference on Industrial Logistics (ICIL 2016), SawiczT. Alnus Sp. z o.o, Krakow, pp. 315–323.

Tomanek, D.P., Schroder, J., 2017. Analyzing the value of information flow by using the value added heat map. In: Dujak, D. (Hrsg) (Ed.), Proceedings of the 17th International Scientific Conference Business Logistics in Modern Management. Faculty of Economics in Osijek, p. 86.

Tomanek, D.P., Hufnagl, C., Schröder, J., 2020. Determining the digitalization degree of information flow in the context of industry 4.0 using the value added heat map. In: Kolinski, A., Dujak, D., Golinska-Dawson, P. (Eds.), Integration of Information Flow for Greening Supply Chain Management. EcoProduction (Environmental Issues in Logistics and Manufacturing). Springer, Cham.

UNIDO, 2021. Challenges to Industry 4.0 | UNIDO Knowledge Hub. AT 12: 08 am. (Accessed 17 March 2021).

Wagner, B., Strobel, M., et al., 2003. Flow management for manufacturing Companies. In: Sustainable Re-organisation of Material and Information Flows, Redline Wirtschaft at Überreuter, Frankfur.