Journal of Cleaner Production 311 (2021) 127692

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro

A sustainability maturity model for micro, small and medium-sized enterprises (MSMEs) based on a data analytics evaluation approach

Jenifer Vásquez[a,b,*], Santiago Aguirre[a], Edwin Puertas[a,c], Giulia Bruno[b], Paolo C. Priarone[b], Luca Settineri[b]

[a] Pontificia Universidad Javeriana, Departamento de Ingeniería Industrial, Carrera 7 No. 40-62, 110231, Bogotá, Colombia
[b] Politecnico di Torino, Department of Management and Production Engineering, Corso Duca Degli Abruzzi 24, 10129, Torino, Italy
[c] Universidad Tecnológica de Bolívar, Facultad de Ingeniería, Km 1 Vía Turbaco, 130011, Cartagena, Colombia

ARTICLE INFO

Handling editor: Yutao Wang

Keywords:
Sustainability
Maturity model
Data analytics
Small and medium-sized enterprises

ABSTRACT

A maturity model for micro, small and medium enterprises (MSMEs) is introduced to assess the level of implementation of sustainability strategies and practices in this type of business. According to the literature, only a few sustainability-maturity assessment models intended for MSMEs have integrated the following three factors: environmental knowledge management, environmental strategies and good practices, and environmental management systems. The sustainability maturity model proposed here for MSMEs is capable of supporting the efforts of companies in their attempts to achieve both environmental sustainability and an improvement in their production systems. The model encompasses a four-level qualitative scale and uses supervised classification algorithms to categorize companies through data analysis techniques. After applying the model to a group of MSMEs from different productive sectors in Colombia, the results have shown that 6% of the companies were at an insufficient level, 31% were at an initial level, 45% at a developed sustainability maturity level, and 18% at a consolidated level. This result implies that the decision makers in the latter group have paid greater attention to the strengthening of sustainability progress capabilities and, hence, to the definition of a maturation route.

1. Introduction

As the commitment to contributing to sustainable development in industrial and service sectors is increasing day by day, companies are trying to establish social responsibilities, comply with environmental laws and standards, improve the efficient use of their resources and include environmental practices in their management and operational processes (Heikkurinen et al., 2019). Therefore, the base line and goals that companies must reach to implement environmental sustainability in such processes are issues to which managers, researchers and governments need to respond. The literature on the topic provides different definitions of corporate environmental "sustainability". The quoted word has been linked to "sustainable development" as "progress that integrates immediate and long-term objectives and local and global action, regarding social, economic and environmental issues as inseparable and interdependent components of human progress" (Brundtland, 1987). According to this definition, the corporate environmental sustainability can be considered as a reduction in waste and environmental impacts produced by internal activities, as a result of making efficient use of economic, social and environmental resources. Companies that have adopted environmental sustainability practices achieve tangible benefits, such as budget savings and new financial sources, as well as intangible benefits like brand improvement, the moral satisfaction of workers, and innovative products and processes, among others (Parker et al., 2017). However, the effective implementation of sustainable strategies in management and industrial operational processes is still a subject of debate due to the lack of standards available to approve it (Shi et al., 2019).

Emerging countries have issued government agreements to support business initiatives that contribute to environmental sustainability (Schwan, 2019). These countries are in fact mostly constituted by micro, small and medium enterprises (MSMEs), which represent 90% of all the businesses in these countries. MSMEs promote employment, help to create innovative ideas and are flexible enough to adapt to the niches of very diverse markets and to meet their specific needs (Jové-Llopis and Segarra-Blasco, 2018). However, MSMEs suffer from certain

* Corresponding author. Pontificia Universidad Javeriana, Departamento de Ingeniería Industrial, Carrera 7 No. 40-62, 110231, Bogotá, Colombia.
E-mail addresses: jenifer.vasquez@javeriana.edu.co, jenifer.vasquez@polito.it (J. Vásquez).

https://doi.org/10.1016/j.jclepro.2021.127692
Received 19 August 2020; Received in revised form 7 April 2021; Accepted 25 May 2021
Available online 28 May 2021
0959-6526/

administrative weaknesses, such as lack of strategic planning, limited financial capital and a low adoption of new technologies, all of which hinders their capability to compete on markets with intense technology adoption within the framework of Industry 4.0 (Shi et al., 2019). As a result, MSMEs often appear to struggle in this fourth industrial revolution, because they find it difficult to adapt to the digital era, which implies making use of data analysis, the internet of things, cloud systems, machine learning, cyber security and changes in environmental-practice-related business processes (Froger et al., 2019; Lin et al., 2020). This has prompted entrepreneurs and the scientific community to design and implement maturity models that help to identify installed capacities, strengths and critical points in business processes, in order to establish continuous improvement mechanisms in the areas where they are most required. Since their inception, business maturity evaluation models have been associated with statistical analysis and software tools. They have been applied to multiple business areas as reference frameworks to analyze company progress through different levels or stages of sequential development (Lacerda and von Wangenheim, 2018). In the context of environmental sustainability in management and operational processes, sustainability maturity models have made it possible to measure the progress of economic, social and environmental activities in companies. Moreover, sustainable measurements have been related to Key Performance Indicators (KPIs) (Odważny et al., 2019).

In addition, sustainability maturity models are related to progress in product design and manufacturing (Watz and Hallstedt, 2020), the assessment of specific areas of a company (Shi et al., 2019), or the identification of maturity criteria for those companies that have been certified and accredited (Lopes et al., 2017). However, considering the sustainability maturity models analyzed in the literature, it becomes clear that only a few studies have comprehensively addressed the key elements of environmental sustainability in internal processes. Examples of the latter are provided through the assessment of a series of social, environmental and economic aspects: (*i*) employee-related social attributes, as they affect the capacity of the company to acquire and internally disseminate environmental knowledge; (*ii*) environmental aspects of internal activities related to the degree of implementation of environmental strategies and practices; and (*iii*) economic features related to the development of environmental management systems for decision-making in MSMEs (Plasencia Soler et al., 2018).

By analyzing the current research, a series of important gaps have emerged. First, only a few studies have specifically focused on measuring the maturity of the small manufacturing and service business sector. Second, sustainability factors have rarely been evaluated, because the models are usually addressed on business or project assessment. For this reason, the present study concentrates on the following research question: What factors should an MSME sustainability maturity model include, and how should they be quantitatively evaluated? To answer this question, the aim of this study has been the design of a Sustainability Maturity Model for MSMEs (SMMM) that can measure the degree of implementation of environmental knowledge management and strategies, good environmental practices, and environmental management systems. This objective has been methodologically addressed through DSR (Weber, 2017) and "Design of Computational Intelligence Experiments" (Fernandez-Lozano et al., 2016). These tools were employed to analyze the behavior of 327 MSMEs from different Colombian economic sectors. The main contributions of the current research study are manifold. First, the development of an SMMM that allows the degree of sustainability maturity to be assessed and the management of environmental knowledge, strategies, good environmental practices, and management systems to be integrated in the internal processes of MSMEs. Second, regardless of the economic activity, the model should allow MSMEs to gradually move through sequential levels of implementation of sustainable practices and strategies, that is, from the basic or incipient to the optimal or consolidated. Third, the identification of the factors on which MSMEs should focus their improvement efforts within the framework of environmental sustainability. Such a recognition would likely contribute to the decision-making processes of entrepreneurs, researchers and government entities. Finally, data analysis allows internal management in companies to be quantitatively evaluated, which, in turn, facilitates the diagnosis and prediction of company needs within the framework of environmental sustainability.

This paper is organized as follows: Section 2 presents a literature review to provide theoretical support to the model development. Section 3 presents the research methodology to design the sustainably maturity model for MSMEs (SMMM). The results of the application of the model to a dataset of Colombian MSMEs are presented in Section 4 and discussed in Section 5. Finally, Section 6 presents the main conclusions and future research perspectives.

2. Literature review

Maturity models describe gradual improvement paths toward the development of good practices, to the point of achieving a desirable state in any organization (Lacerda and von Wangenheim, 2018). A company's activity can be assessed in terms of capacity and maturity. The former has to do with the ability to achieve or contribute to the fulfillment of a required objective, whereas maturity addresses the extent to which an organization has consistently implemented processes or practices with a clearly defined scope, thus contributing to the achievement of project objectives (Froger et al., 2019; Von Scheel et al., 2014). Experts have developed various maturity models that can be applied to the different functional areas of a company. They are used as diagnostic and analysis tools to assess the level of development of a product, project or process.

Among the various general maturity simulations that are available, the *Capability Maturity Model* (*CMM*) was conceived to evaluate software development processes (Lacerda and von Wangenheim, 2018). This elaboration model defines a series of sequential steps that determine the state of development of a computer artifact. *CMM* has inspired other business maturity models, such as the *Project Management Maturity Model* (*PM3s*), which is based on the guidelines of the Project Management Book of Knowledge (PMBOK) (Backlund et al., 2014). This is an essential tool for project managers since it allows the progress made in activity fulfillment to be described at each stage of a project. Similarly, the *Business Process Maturity Model (BPMM)* is based on analyses of the current conditions of a company conducted to establish a process management path (Froger et al., 2019). Five evolutionary levels are defined in the *BPMM* by relating capabilities and processes. This allows process maturity to be measured and continuous process improvement to be promoted (Dharmawan et al., 2019).

In the logistics area, the *Maturity Model - Logistics 4.0* (Ellefsen et al., 2019), which is supported by artificial intelligence, establishes maturity levels on the basis of the production and storage activity needs. In this study, businessmen have been interviewed about the term *Logistics 4.0* and/or *smart manufacturing* solutions. The authors were able to classify the studied companies by correlating their competitive position, size, development dynamics and operation internationalization levels. A novel maturity model has recently been used to introduce a measurement framework for Taiwanese SMEs in the context of the evaluation and execution of strategies intended to implement *Industry 4.0* in the processes, technology and organization areas (Lin et al., 2020). Despite the contributions presented in this study, a particular evaluation bias toward the manufacturing sector has been perceived.

Additionally, several quality models can be found within the certification and business-excellence-achievement families. Among the common prescriptive models, mention can be made of *Standardization Norms,* which are provided by the International Organization for Standardization (ISO), who have published around 23,047 international standards covering such aspects as technology, service and manufacturing (ISO, 2020); and of the excellence model proposed by *European Foundation Quality Management (EFQM)* (EFQM, 2020) which, despite not being normative in nature, is oriented toward total quality standards through a series of self-evaluation principles that are measured, monitored and audited. However, international standards have been the target of criticism by businessmen, who consider that certifications can generate high investment costs and an excess of internal procedures, which in turn can result in internal structures that are too bureaucratic (Carmona-Calvo et al., 2016).

The literature also presents *Sustainability Maturity Models*, which have been developed to analyze the degree of maturity of a specific area, product or process. A sustainable product design maturity model has recently been introduced, pertaining to the analysis of 17 medium to large Swedish companies, which presents an innovative way of associating the design and creation phases of a product (Watz and Hallstedt, 2020). An analysis is conducted through the multi-case methodology and product development funnel method, and this has resulted in a conceptual model for the management of sustainable product requirements. However, some limitations appear when detailing the necessary knowledge or skills for eco-product design experts to manufacture decision-making products.

A maturity model has recently been introduced to sustainably improve the manufacturing of products in the *smart manufacturing* era (Shi et al., 2019). This model, which is known as *Kaizen Level of Smart Manufacturing*, was specifically proposed for process automation systems, for which it defines a series of generic-automation-architecture implementation levels. Its limitations have to do with a lack of model testing in business practice. Therefore, the same authors concluded that it would be necessary to expand their model to all links in the value chain.

Additionally, accreditations and certifications make it possible to identify whether an organization has reached a desirable status. Environmental management system certifications have been studied in the health services area. After classification, through a non-parametric test, thirty-eight certified Brazilian hospitals exhibited management maturity, a parameter for which most of them attained elevated scores when considering the application of sustainable tools (Nascimento et al., 2017). However, some as yet uncertified hospitals may also be implementing good sustainable practices. In addition, the authors did not assess the social and economic dimensions of sustainability.

A maturity model review has recently been conducted to evaluate the sustainability of organizations, as framed at the business level and approached from the ethical perspective of corporate social responsibility theories (Plasencia Soler et al., 2018). These authors addressed the following models: *a*) pressure - state - response and its variations, *b*) the Lowell Center model for sustainable production, *c*) the CUBRIX model and *d*) the sustainable balanced scorecard. In doing so, they established not only the advantages but also the disadvantages of the theories behind these models and their applications, despite their approval and endorsement by governmental entities and international institutions. The downsides of such models have to do with the lack of practical and reliable data analysis tools and the difficulty of assessing the effects of interventions on society and the environment when they are not numerically assessed.

By analyzing the previous developments in maturity models and relating them to MSMEs, it is evident that, due to the intrinsic characteristics of these companies, some studies have emphasized the sequential steps proposed by maturity models as related to the management systems. According to a recent study, ISO 9004 standards are clear quality pattern models for the processes of this type of company, since they allow diagnoses to be made and improvement routes to be established (Páez et al., 2018). However, these models are only linked to the quality and implementation costs of operative processes, and thus gaps remain in meeting the integration of the environmental and strategic aspects of a company.

In this respect, a sustainability maturity model that resorts to the ISO/IEC 15504 Information Technology Guide has been introduced for the remanufacturing of SMEs (Golinska and Kuebler, 2014). The authors adapted some maturity models and related them to the application of sustainable practices. They evaluated four specific factors by surveying five European and five American companies: energy efficiency, compressed air, ergonomics and emissions. They then performed a statistical analysis with the data thus obtained data which allowed them to determine that the studied American companies showed a better performance than the European ones. However, the sample was small, and the conclusions did not cover all the factors that can lead to sustainability.

The current literature review has revealed several important gaps and research problems. In fact, only a few studies have focused on measuring the maturity of the small manufacturing and service business sector. Furthermore, sustainability factors have rarely been evaluated. Environmental knowledge management and environmental management systems have been considered separately when evaluating the implementation of environmental practices or strategies, which has led to the present research in which an integral evaluation of both components has been conducted. Finally, in the current literature review, sustainability maturity has been assessed from “insufficient” (level 0) to “consolidated” (level 5). In this context, the main objective of this research has been to design an SMMM for the comprehensive evaluation of environmental knowledge management, environmental strategies and practices, and environmental management systems, in order to contribute to the decision-making of businessmen, governments and researchers. In addition, a need was observed to identify the strengths and weaknesses of MSMEs in order to establish continuous improvement mechanisms as a contribution to sustainable development. On these grounds, the following section describes the methodology used to design the SMMM proposed in this research.

3. Sustainability maturity model for MSMEs

The current study is a descriptive and exploratory research in which qualitative and quantitative analysis have been integrated. The methodology is based on DSR (Weber, 2017), which involves establishing a general research framework and designing an innovative and useful model for the solution of a particular problem through different phases and sequential activities. Taking into account that the proposed SMMM model should be applied in order to ensure its usefulness for the problem in question, the Design of Computational Intelligence Experiments methodology (Fernandez-Lozano et al., 2016) has been proposed. This methodology establishes the synergy between each of the established phases as well as the impacts of each of them, and proposes new requirements and patterns for data analysis. Fig. 1 shows the research methodologies employed in this study.

The sequence of phases followed in the current methodology is reported hereafter:

- **Phase 1.** Relevance: Developed in the previous Sections 1 and 2, this phase provided a conceptual reference framework. Moreover, the literature review also allowed the research problem, the research gap, and the research objectives to be established.
- **Phase 2.** Design: Developed in the current Section 3, this phase involves identifying the inherent characteristics of knowledge management factors, environmental practices and strategies, and environmental management systems. These characteristics are subsequently classified according to maturity fulfillment levels, which provides the basis for the design of the SMMM model and allows the population that has to be studied to be identified.
- **Phase 3.** Rigor: Developed in Section 4 and discussed in Section 5, this phase, which is based on the results of a survey applied to Colombian MSMEs, corresponds to model validation and data analysis. A series of algorithms are supervised and trained by means of data preprocessing. The best algorithm use in the application of the MSMM is then selected.

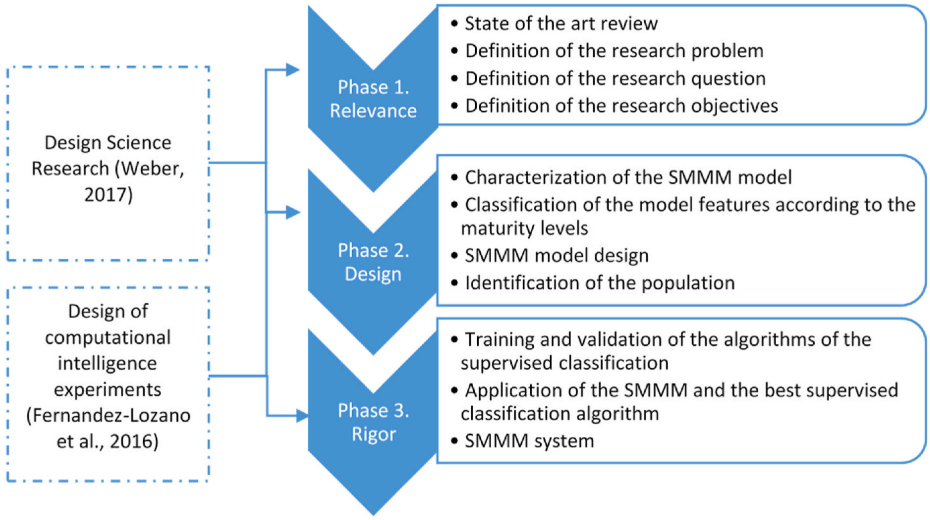

Fig. 1. Research methodology based on Design Science Research and Design of Computational Intelligence Experiments.

Focusing on the development of Phase 2 of the current methodology, the sustainability maturity model for MSMEs (SMMM) is designed, and three factors are integrated: (*i*) environmental knowledge management, (*ii*) environmental strategies and good practices, and (*iii*) environmental management systems. These factors are closely related to the social, environmental and economic components usually applied in environmental sustainability analyses in internal business processes. Moreover, Phase 2 also describes the set of companies analyzed to apply the SMMM and defines the algorithm training and validation processes that facilitated the maturity classification.

3.1. Description of the factors and levels of the SMMM

3.1.1. Factor 1: environmental knowledge management in MSMEs

Environmental knowledge management is used by workers in companies to reduce the environmental impacts of production and consumption processes by acquiring knowledge in the field of environmental protection (Edvardsson et al., 2016). "Knowledge" is expressed as the ability to acquire information individually or in a group, in order to improve effective actions in companies (Biscotti et al., 2018). However, because of certain features of MSMEs, they lack specific capabilities to effectively carry out knowledge acquisition. Therefore, it is necessary to identify the fundamental characteristics of environmental knowledge management when implementing it in MSMEs. The management of environmental knowledge, as applied for the improvement of internal processes in companies, requires technological, human and financial capabilities to introduce amendments and improve their products and business processes. These capabilities constitute the key aspects of the adoption and implementation of environmental management practices (Edvardsson et al., 2016).

Moreover, it also is essential that the management of environmental knowledge in companies should be carried out by first training the employees and then allowing them to share and socialize their knowledge. This communicative interaction, which requires specifically arranged spaces, allows the productive processes to be improved. Another characteristic of environmental knowledge management is the documentation of strategic and operational processes, which convey a clear notion to the employees of the steps that are necessary to complete a task or process while minimizing the environmental impact (Johnson, 2017). In this way, and framed within corporate environmental sustainability, knowledge management begins to be an important aspect of business administration, all the more so because of the necessary willingness that employees must have to understand and implement environmental practices or prioritize the concomitant needs and requirements.

3.1.2. Factor 2: strategies and good environmental practices in MSMEs

Good environmental practices and strategies are useful measures and actions taken by companies to reduce the negative impacts caused by their productive activities. These are actually variations or transformations of their internal or external processes, which allow them to achieve sustainable development, improve the quality of their products or services and comply with environmental legislation (Čater et al., 2018). The main environmental strategies and practices include cleaner production, eco-efficiency, circular economy and 3R, all of which are featured by low application complexity, improvement of the environmental performance indicators, and the "producing more with less environmental impact" vision (WBCSD, 2009). Performance improvement indicators include the reduction of solid waste, an efficient use of energy, a quality enhancement of air and water, noise moderation, and environmental commitment by the employees.

The main benefits of implementing good strategies and environmental practices in a company include competitive improvement through image enhancement, budget savings, compliance with environmental policies and reliability in the relationship with customers (Arbolino et al., 2018). However, MSMEs struggle to include good practices in their operative processes, due to the barriers or internal factors related to their current business functioning. For this reason, it is important to analyze the specificities of a company to identify any useful and simple practices they can apply to their internal processes (Vásquez et al., 2019). Environmental strategies and practices play important roles within the framework of corporate environmental sustainability, because they produce good and rapid results that actively contribute to sustainable development.

3.1.3. Factor 3: environmental management systems in MSMEs

Environmental management systems are tools that companies can use to organize management, operational and support processes which, in turn, facilitate compliance with responsible environmental policies. These systems are characterized by guidelines and practical approaches for organizations to continuously improve their operations and environmental performance. They contribute to the formulation of environmental policies aimed at minimizing and optimizing environmental resources associated with the compliance with legal requirements (Johnstone, 2020). Another feature of environmental management system is the support they provide to establish organizational leadership and responsibilities. This, in turn, allows significant changes to be made to the internal processes to reduce nonconformities and increase preventive actions accompanied by the formulation and control of environmental indicators. ISO 14001 defines environmental management systems as an international standard that should be adopted to achieve operational, strategic and tactical goals (ISO, 2015). Although MSMEs lack the necessary financial resources to invest in the environmental management systems laid out in ISO standards, some of these systems have been adapted to the needs of these companies and are available in business contexts. These procedures allow a sustainable business culture to be established that promotes the efficient use of natural resources and the design of efficient and effective strategies for the continuous improvement of internal corporate processes.

By taking into account the factors defined above, it is possible to assess the *level* of maturity of a particular SMMM. For such a purpose, each factor is further described through a series of specific characteristics which are, in turn, evaluated in terms of whether they have been fulfilled (value 1) or not (value 0) at a given level of sustainability maturity. The latter is assessed on a four-level scale, which is based on some of the aforementioned generic models (MMC, Kaizen, PM3s, BPMM). In this scale, *level* 1 is considered to be insufficient, while consolidated maturity corresponds to *level* 4 (Santos-Neto and Costa, 2019). Accordingly, the higher the level is, the higher the number of characteristics that have to be fulfilled. Thus, *level* 1 implies the accomplishment of only 5 characteristics, while *level* 2 implies 14 characteristics have been fulfilled. By the same token, *levels* 3 and 4 require the fulfillment of 27 and 37 characteristics, respectively (see Table 1). In this way, a conceptual definition is provided for each of the maturity *levels* of the SMMM.

3.1.4. Definition of SMMM levels

The different levels of the SMMM are defined upon the fulfillment of the characteristics that describe each factor (Fig. 2).

- **Level 1. Companies with insufficient sustainability-maturity**

Table 1
Classification of the factor-describing characteristics according to their fulfillment requirement at each maturity level.

Factor	No.	Characteristics	Level 1	Level 2	Level 3	Level 4
Environmental knowledge management	1	Employees are considered to have a high level of environmental management knowledge.	0	0	0	1
	2	Sufficient human resources are available for environmental knowledge acquisition.	0	1	1	1
	3	Sufficient financial resources are available for environmental knowledge acquisition.	0	0	1	1
	4	Sufficient technological resources are available for environmental knowledge acquisition.	0	0	0	1
	5	Sufficient physical resources are available for environmental knowledge acquisition.	0	1	1	1
	6	Employees are trained on environmental care.	0	1	1	1
	7	Environmental training is conducted on a monthly, bimonthly or trimonthly basis.	0	0	0	1
	8	Environmental training is conducted every six months.	0	0	1	0
	9	Environmental training is conducted on a yearly basis.	1	1	0	0
	10	Adequate channels are offered for the employees to present environmentally innovative ideas.	0	0	1	1
	11	The employees use Environmental Training to improve production processes.	0	0	1	1
	12	The company has documented its operative processes	0	0	1	1
	13	The employees are well-informed about the environmental impacts caused by the organization and their corrective measures.	0	1	1	1
	14	Environmental sustainability is part of the central strategy of the company and is among the first priorities on the agenda.	0	0	0	1
	15	Environmental sustainability is relevant for some parts of the company agenda, but not for all of it.	0	0	1	0
	16	Environmental sustainability is not among the priorities of the company agenda, but has some importance	0	1	0	0
	17	Environmental sustainability is not relevant for any activity on the company agenda.	1	0	0	0
Environmental practices and strategies	18	The employees understand the topics involved in environmental eco-efficiency and cleaner production.	0	0	0	1
	19	The Company has an environmental practice plan or program.	0	0	1	1
	20	The employees understand the topics involved in circular economy.	0	0	0	1
	21	The company has established a plan to report outflows, such as liquid substance spills, gas leaks, gas-liquid mixtures and other non-eco-efficient situations.	0	0	1	1
	22	Environmental practices have been implemented.	1	1	1	1
	23	A water saving and efficient use program is in operation.	0	1	1	1
	24	A solid residue collection and classification program is in operation.	1	1	1	1
	25	The company verifies the final disposal of hazardous residues (corrosive, reactive, explosive, toxic and flammable waste).	0	0	0	1
	26	Employees tend to reuse office stationery materials.	1	1	1	1
	27	Treatments to extend the use of industrial resources, such as oils, lubricants, acids, etc., are in force.	0	0	1	1
	28	The products that the customers no longer use are recovered.	0	0	0	1
	29	An after-sales repair service to extend the useful life of products is in force.	0	0	0	1
	30	Noise level measuring systems are in operation.	0	1	1	1
	31	The company keeps a statistical record of power and water bills.	0	0	1	1
	32	Environmental purchasing criteria are in place in the selection of suppliers.	0	0	0	1
	33	The suppliers comply with the environmental laws in force.	0	1	1	1
Management systems	34	Employees have knowledge of environmental management systems.	0	1	1	1
	35	An environmental management system is in operation.	0	0	0	1
	36	The Company has been environmentally certified.	0	0	0	1
	37	An environmental policy is currently in force.	0	1	1	1
	38	The main environmental policies and legislation established by the government that have to be applied within the company are clear.	0	0	1	1
	39	The company makes use of operating environmental indicators.	0	0	1	1
	40	Manufacturing or service delivery processes and environmental practices are related to decision-making support.	0	0	1	1
	41	An Enterprise Resource Planning (ERP) system is in operation.	0	0	1	1
	42	The company has an HSEQ (Health, Safety and Environmental Quality) module for strategy control in its ERP systems.	0	0	1	1
		Total fulfilled characteristics	5	14	27	37

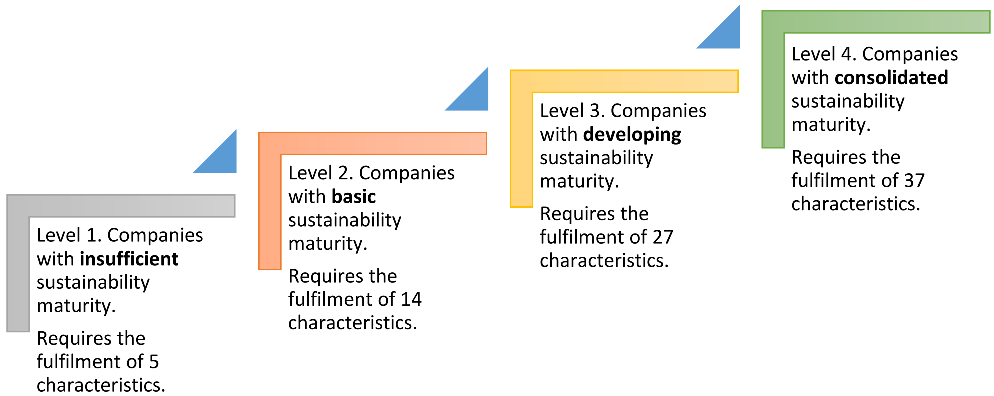

Fig. 2. SMMM levels.

Environmental sustainability is not a relevant issue in any of the internal activities contemplated on the agenda of these companies. Therefore, environmental training is only sporadic or is simply not carried out. However, such companies have implemented some environmental practices, such as noise control in operational processes, solid waste collection and classification programs, as well as activities to encourage the reuse of stationery materials in each department.

- **Level 2. Companies with basic sustainability-maturity**

Environmental sustainability is not among the priority issues on the agenda of these companies, although they try to give importance to the topic. Training on environmental issues and practices is carried out annually or every two years or more. Employees are informed about the environmental impacts caused by their organizations and the corresponding corrective measures that should be introduced. In addition, these companies have implemented such environmental practices as stationery reuse, management of the noise level, and solid waste collection and classification programs. Although they do not have environmental management systems in place, they do have an environmental policy.

- **Level 3. Companies with developing sustainability-maturity**

Environmental sustainability does not rank high on the agenda of these companies, although it does have a certain relevance. They consider that their employees have some environmental management knowledge as a result of adequate investments in human, technological and physical resources. Moreover, the employees are trained every six months in the implementation of the environmental strategies that are used to improve the operating processes of these companies, which provide spaces for workers to express innovative environmental ideas. These organizations have also documented their operative procedures. The employees are informed about the environmental impacts caused by their organizations and the corresponding corrective measures. They have implemented an environmental practice agenda that includes plans to prevent leaks and spills of toxic substances; noise management systems; campaigns to reuse industrial resources; and programs for office stationery reuse, collection and classification of solid waste, recovery of products that customers no longer use, and efficient water and energy use and measurement. These businesses have defined environmental management systems that contribute to their decision-making processes, and thus affect the corporate policies, indicator monitoring and control activities, resource planning and HSEQ modules in their business information systems.

- **Level 4. Companies with consolidated sustainability -maturity**

Environmental sustainability is part of the core strategy of these companies and is hence placed at the top of the priority agenda. These companies have sufficient human, financial and technological resources for the acquisition and management of environmental knowledge. For this purpose, they constantly train their employees, who are able to make good and responsible use of the environmental resources of these organizations. These companies have tools or mechanisms in place for their employees to express topical innovative ideas. They document their operational procedures to share knowledge among the employees, who are informed about the environmental impacts and corrective measures associated with the corporate activities. These organizations implement environmental practice programs, such as leak reporting plans to cover liquid spills, gas leaks, gas-liquid mixtures and other non-eco-efficient situations; water and energy measurement, saving and efficient use; solid waste collection and classification; office stationery reutilization; noise management; industrial resource reuse; and verification of the final disposal of hazardous waste. In addition, these organizations have overt corporate strategies associated with their environmental policy, supplier selection and recovery programs for products that their customers no longer use. They also have incorporated control systems to measure environmental, social and economic indicators in all the links of their value chain. These systems have allowed them to obtain quality, environmental and/or occupational health management certification.

3.2. The studied MSME set

In the present study, Colombian MSMEs from different sectors were considered to evaluate and validate the studied SMMM. The trade, service, construction, logistics and manufacturing sectors in Colombia are represented by micro, small and medium enterprises scattered throughout the territory. However, the highest concentrations of MSMEs are in the five main cities in the country: Barranquilla, Bogota, Cartagena, Cali, and Medellin. The contact data of the companies located in these cities were obtained from the Orbis® database, which is characterized by "[…] including information on some 360 million companies worldwide. It is the resource for obtaining company data" (Bureau van Dijk, 2020). By using Orbis®, a list of more than 3500 active MSMEs from the studied cities was prepared. A total of 1350 MSMEs were eliminated from this list because they contained erroneous or misspelled contact data. Therefore, a population size of 2150 MSMEs was calculated, which resulted in a sample size of 327 MSMEs for the five main cities in Colombia. According to the methodology used in this research, it was necessary to apply the SMMM, and this led to the development of a data collection instrument. The main objective of the questionnaire that was administered to these companies was to collect information about the three studied factors in the companies:

1. **Factor 1.** Environmental knowledge management: To what extent do MSMEs have knowledge on environmental issues?

2. **Factor 2.** Environmental practices: What strategies and environmental practices are implemented in the MSMEs?
3. **Factor 3.** Environmental management systems: What tools are used to plan activities, processes and procedures for a continuous improvement of the environmental performance of the company?

The questionnaire was based on characteristics that described each of the studied factors, which had been validated by experts (Golinska and Kuebler, 2014). It contained four specific sections: Part 1. Six questions about general company data, which were intended to characterize and identify the type of evaluated companies. Part 2. Sustainable integration, which separately addressed the three factors being evaluated: (*i*) environmental knowledge management, (*ii*) environmental strategies and practices, and (*iii*) management systems. The questionnaire contained 42 questions aimed at identifying the strengths and weaknesses of the specific sectors under consideration. The questionnaire contained multiple choice, open and yes/no questions. The former allowed the respondents to choose the most satisfactory option from a range of alternatives. Yes/no questions were introduced to obtain immediate information from questions with only two possible options, while open questions were used to collect additional information that could have been of interest for the research. Using the contact data registered in the aforementioned databases (company name, telephone number, e-mail, address, city and production activity), interviews were conducted with the managers or personnel responsible for the production, quality or environment in the companies in 2019. The decision was made to send the survey via e-mail, together with the research objectives and an invitation to participate in this study. On the basis of data from Part 1 of the questionnaire, Fig. 3 shows the number of MSMEs that participated in the study, as classified by city. It can be observed that the cities of Bogota (26%) and Cali (23%) contributed the most. The cities of Barranquilla (19%) and Cartagena (18%) scored similar intermediate participation levels, whereas Medellin (14%) had a low participation level.

Fig. 4 shows the distribution of the sample across the economic sectors. According to the *International Standard Industrial Classification of All Economic Activities* (ISIC. Rev. 4), the companies were classified as coming from three main sectors: manufacturing, services and trade, and construction and civil works. The manufacturing sector had the largest representation, with 39% of the sampled businesses, followed by the service and trade sector, with 38% and the construction and civil works sector, with 23%. The studied businesses were also classified by size, according to the number-of-workers' criterion established by Law 590 of 2000 by the Colombian Ministry of Commerce, Industry and Tourism (Bancóldex, 2017) (see Table 2).

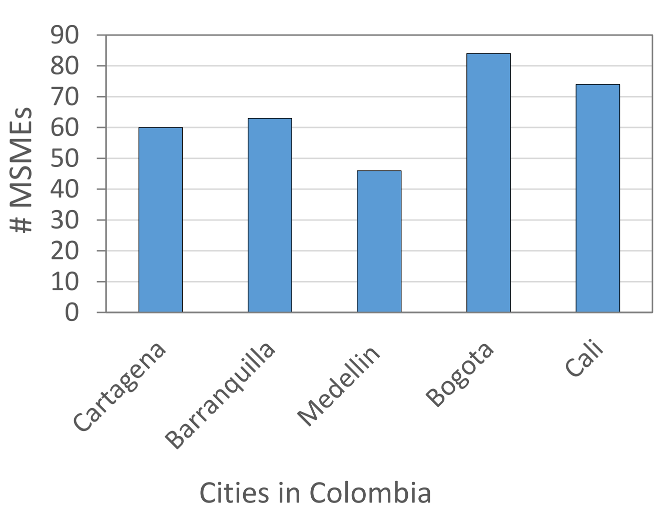

Fig. 3. MSMEs evaluated for the city of Colombia.

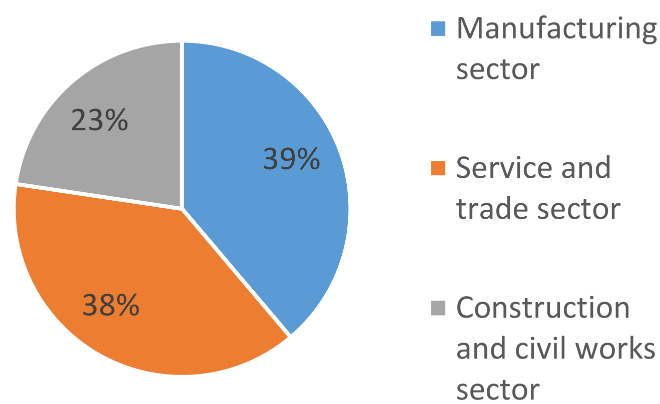

Fig. 4. Economic activities of the evaluated MSMEs.

Table 2
Company size classification according to the number of workers.

Business type	No. of workers	Surveyed businesses
Microenterprises	No more than 10	48%
Small businesses	Between 11 and 50	37%
Medium businesses	Between 51 and 200	15%

Other general information about the evaluated companies is that pertaining to the guild affiliation, to which only 12.8% of the sample belonged. The remaining 87.2% of the companies were found to not belong to any type of business guild. The companies were affiliated to the most representative guilds in Colombia, namely the Colombian Association of Micro, Small and Medium Enterprises (*Asociación Colombiana de las Micro, Pequeñas y Medianas Empresas - ACOPI*), the Colombian Chamber of Construction (*Cámara Colombiana de la Construcción - CAMACOL*), the Colombian Hotel Association (*Asociación Hotelera Colombiana - ASOTELCA*), and the Hotel and Tourism Association of Colombia (*Asociación Hotelera y Turística de Colombia - COTELCO*). Details of the questions used for the evaluation of the SMMM are presented in Table 3. After 327 MSMEs had filled in the questionnaire, the supervised classification algorithms were used to validate the SMMM through data analytics.

4. Results

Focusing on Phase 3 of the proposed methodology, this section reports the results of the survey applied to Colombian MSMEs, together with the model validation and data analysis.

4.1. Training and validation process

Data analytics uses different analysis methods to interpret data. Among the most commonly used methods is machine learning (ML), which, more than a simple database that processes information, is a programming problem solver. It is a system that, on the basis of a dataset provided by the user, has the "ability to learn" by representing data structures and generalizing the corresponding behaviors (Alpaydin, 2016; Fernandez-Lozano et al., 2016). In phase 3 ("Rigor") of the current research methodology, a sequence of activities is performed to train and validate the supervised classification algorithms used by the SMMM, as shown in Fig. 5.

4.1.1. Dataset

The dataset is the tabulation of the different responses (*i.e.*, numerical values) provided by the companies who responded to the questionnaire sent to the 327 Colombian MSMEs under study.

Table 3
General questions by factors.

Factors under evaluation	Questions
Environmental knowledge management	Q1. Does the company consider that its employees have elevated environmental management knowledge? Q2. Does the company have sufficient human resources for environmental knowledge acquisition? Q3. Does the company have sufficient financial resources for environmental knowledge acquisition? Q4. Does the company have sufficient technological resources for environmental knowledge acquisition? Q5. Does the company have sufficient physical resources for environmental knowledge acquisition? Q6. Have the employees received environmental care training? Q7. Are environmental trainings conducted on a monthly, bi-monthly or tri-monthly basis? Q8. Are environmental trainings conducted every six months? Q9. Are environmental trainings conducted every year or more? Q10. Does the company provide the means for the workers to contribute innovative environmental ideas? Q11. Do the employees use environmental trainings to improve productive processes? Q12. Has the company documented its operative processes? Q13. Are the employees informed about the environmental impacts of the organization and the corresponding corrective measures? Q14. Select the option that best reflects the relevance of environmental sustainability on your company's agenda: it is part of the core strategy and is at the top of the priority agenda. Q15. Select the option that best reflects the relevance of environmental sustainability on your company's agenda: it is relevant for some activities, but not for all the company. Q16. Select the option that best reflects the relevance of environmental sustainability on your company's agenda: It is not among the priority items on the agenda, but it is given some importance. Q17. Select the option that best reflects the relevance of environmental sustainability on your company's agenda: It is not relevant for any of the company's activities.
Environmental practices and strategies	Q18. Do you understand the notions of ecoefficiency and cleaner production? Q19. Does your company have an environmental practice program? Q20. Do you know what circular economy is? Q21. Has the company set up a plan to report liquid spills, gas leaks, gas-liquid mixtures or any other non-ecoefficient situations? Q22. Has your company implemented environmental practices? Q23 Does your company have a water saving and efficient use program? Q24. Does your company have a solid waste collection and classification program? Q25. Does your company check the final disposal of dangerous waste (corrosive, reactive, explosive, toxic and flammable materials)? Q26. Do the employees tend to reuse office stationery materials? Q27. Does the company use treatments to extend the use of industrial resources such as oils, lubricants, acids, etc.? Q28. Does the company recover the products that their customers no longer use? Q29. Does the company offer after sales repair service to extend the duration of a product? Q30. Does your company have any noise level measurement system? Q31. Does the company have a statistical record of the power and water bills? Q32. Does the company apply purchase environmental criteria when selecting its suppliers? Q33. Do you believe that your suppliers comply with the environmental legislation in force?
Management systems	Q34. Do you know what an environmental management system is? Q35. Does your company have an environmental management system? Q36. Has the company received any environmental certification? Q37. Does the company have any environmental policy? Q38. Are the main environmental policies and legislation established by the government that have to be applied in your company clear enough for implementation? Q39. Does your company have any environmental indicators? Q40. Are environmental practices and the manufacturing or service delivery processes related to decision-making support? Q41. Does the company have an Enterprise Resource Planning system (ERP)? Q42. Does the company have a strategy-control HSEQ (Health, Security, Environment and Quality) module in the Enterprise Resource Planning (ERP)?

4.1.2. Data preprocessing

The answers to the questions were standardized in a numerical data format, so they could be analyzed by the classification algorithm. In such a process, yes/no questions are represented as logical values, *i.e.*, "No" is given a value of 0 and "Yes" is given a value of 1. Since the questionnaire included multiple choice questions, a conversion to a binary range was carried out, to adjust all the answers to logical values. Therefore, the "totally disagree", "disagree" and "undecided" answers are given a value of 0. The "totally agree" and "agree" answers are given instead a value of 1. In the present research study, each company was manually labeled according to its sustainability maturity level, and this allowed the sampled organizations to be grouped. Fig. 6 shows the results of this manual labeling: 26 companies were classified as Level 1, 110 companies as Level 2, 136 as Level 3, and 55 corresponded to Level 4. Most of the studied MSMEs were ranked at sustainability maturity Levels 2 and 3, *i.e.*, at the initial and developing levels of environmental knowledge management, practices and management systems.

It was observed, before the supervised-learning algorithms were trained, that the companies were not evenly distributed across the sustainability maturity levels. For this reason, the "oversampling" data analysis technique was applied to those levels that were below the one containing the most companies, thus allowing a homogeneous data distribution to be achieved. Oversampling consists in repeatedly multiplying the sampling frequency to correct for the bias of the original dataset, thus adjusting the distribution (Mukherjee et al., 2019; Patil and Sonavane, 2020). In this way, the sustainability-maturity-level classification algorithm achieves a greater accuracy when applied to a sample of companies. Thus, the level 1, 2 and 4 samples were replicated according to the number of level 3 replicates, that is, 136 replicates.

4.1.3. Supervised learning algorithms

Supervised learning works as the result of the use of classification algorithms that allow the category (e.g., a sustainability maturity level) to which a new instance (e.g., a company) belongs to be identified, on the basis of previous observations. These procedures include "Multi-

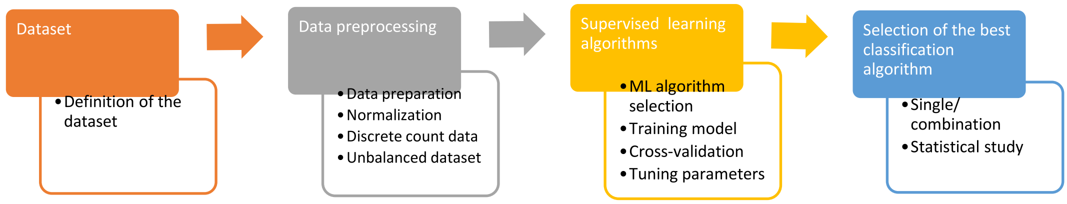

Fig. 5. Training and validation activities of the supervised classification algorithms.

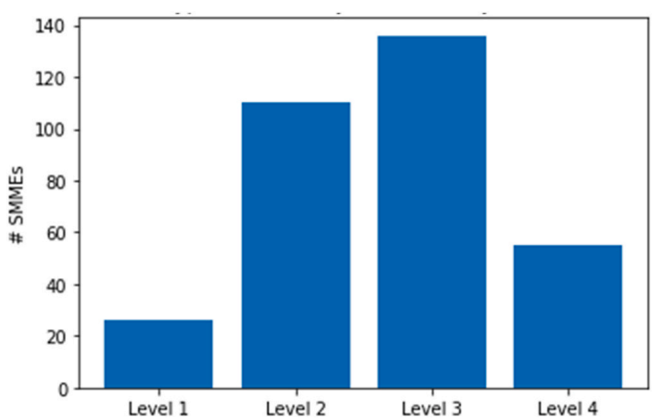

Fig. 6. Classification of the MSMEs according to the SMMM

class" algorithms, which handle several classification categories. In the present study, this ranking corresponded to sustainability maturity levels 1 to 4. It should be pointed out that the success of classification algorithms depends on the particular handling of the dataset in which they are trained. This handling consists in preprocessing and standardizing the data.

The *k*-fold Cross Validation technique is employed to avoid over-adjustment of the samples. In this technique, the set of training data is divided into *k* subsets and, at the time of training, all the subsets are sequentially used to test the model, while the remaining data are used as the actual training set. This process is repeated *i* times, and a different test set is selected in each iteration, while the remaining data are used as the training set. Once the iterations have been completed, the accuracy and error values of each model are calculated. Additional accuracy and final error values are obtained by calculating the corresponding averages of the *k* trained models (Fushiki, 2011; Klyueva, 2019; Moreno-Sandoval et al., 2019).

The Support Vector Machine (SVM), Random Forest (RF) and Naïve Bayes (NV) classification algorithms were employed to perform the training and validation of the algorithms used in the SMMM. These procedures were chosen due to their capability to simultaneously solve classification and regression problems by estimating the internal variability of the samples. SVM is used to predict the maturity level of a new sample, on the basis of a set of previously labeled ones. Similarly, RF classifies new samples on the basis of the attributes of prior ones, where the classifying prediction is the major choice of a decision tree. NV assumes that the predictive variables used to classify a new instance are independent of each other, *i.e.*, the presence of a certain characteristic in the dataset is not related to the presence of any other characteristic in it.

In addition, and according to the Design of Computational Intelligence Experiments methodology (Fernandez-Lozano et al., 2016; Tsiliki et al., 2015), learning models are used to evaluate the effectiveness of algorithms through different statistical techniques. In the present case, the dimension reduction technique, which involves shortening the number of random variables under study, was applied. This procedure is commonly assessed through the following metrics: F1-Score, Accuracy, Recall, and Precision. Fig. 7 shows the training and validation process used for the classification of the SMMM algorithms. A 10-fold cross validation procedure was employed, for which the data were divided into two sets: 70% for training and 30% for validation. During the execution of this process, the parameters used for the SVM, RF and NV algorithms were established by default, that is, as described below:

- SVM. C = 1.0, kernel = 'rbf', degree = 3, gamma = 'scale', coef0 = 0.0, shrinking = True, probability = False, tol = 0.001, cache_size =

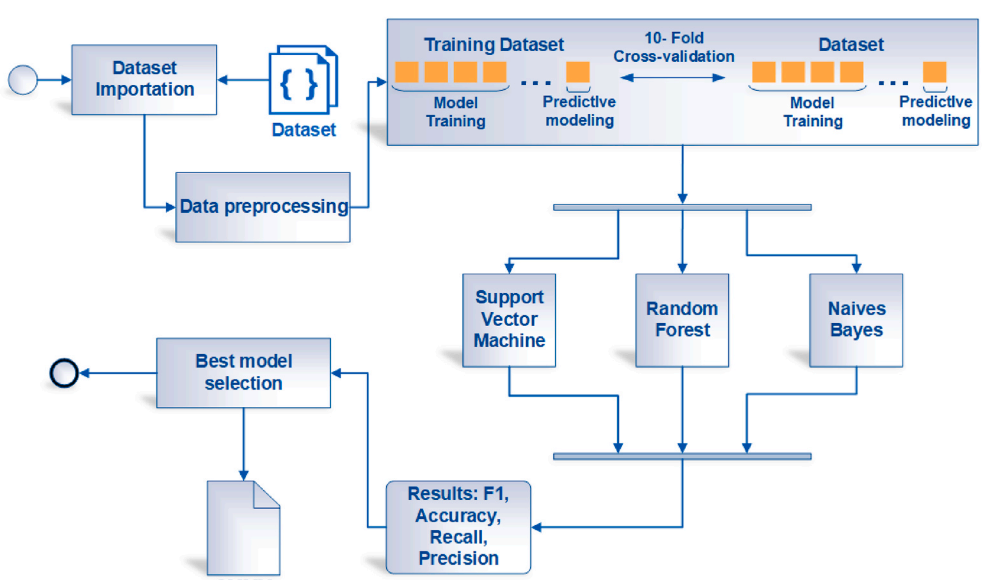

Fig. 7. SMMM training and validation process.

200, class_weight = None, verbose = False, max_iter = − 1, decision_function_shape = 'ovr', break_ties = False, random_state = None.

- Random Forest. n_estimators = 100, criterion = 'gini', max_depth = None, min_samples_split = 2, min_samples_leaf = 1, min_weight_fraction_leaf = 0.0, max_features = 'auto', max_leaf_nodes = None, min_impurity_decrease = 0.0, min_impurity_split = None, bootstrap = True, oob_score = False, n_jobs = None, random_state = None, verbose = 0, warm_start = False, class_weight = None, ccp_alpha = 0.0, max_samples = None.
- Naive Baye. priors = None, var_smoothing = 1e-09

4.1.4. Selection of the best classification algorithm

The Design of Computational Intelligence Experiments methodology (Fernandez-Lozano et al., 2016; Tsiliki et al., 2015) comprises several procedures that are used to validate the relevance and performance of the algorithms in question. On these grounds, it is possible to select the best classification algorithm to improve or complement information heterogeneity through data analytics predictors. Relevance is based on repeatedly executing the classification algorithm under the same conditions and with the same set of data and error rates, in order to ensure that the results are not biased. This is done by evaluating the behavior of the mean and standard deviation of all the executions. In the present case, a 10-fold calculation of the mean and standard deviation values was carried out for each of the metrics employed in the experiment.

With regards to performance, the null and true hypothesis test is used to validate which algorithm is statistically better than the others. For this purpose, the assumptions of independence and normality are verified through parametric and non-parametric tests. In the present study, the Student's t-test (parametric) is used to validate that the means of the training and test samples were independent and significantly different from each other, with all the algorithms reaching higher p values than 0.05. The null hypothesis (non-parametric test) is adopted to establish the precision of the models, using the F1-score value to determine a single weighted value of precision and completeness. In addition, these tests are complemented with the confusion matrix, which is used to determine the performance of the algorithms adopted in the training phase.

4.2. Algorithm training and validation results

Using the data analytics approach, this study has proposed an SMMM model which uses supervised classification algorithms to classify MSMEs according to their compliance with different sustainability features. This section presents the results obtained after applying the SMMM model to the sample of studied companies.

In order to determine the sustainability maturity level of the sampled businesses by productive sector, it was necessary to carry out the training, validation and testing of the supervised classification algorithms. Out of the 327 MSMEs, 70% were selected for training and validation. This implied randomly choosing 231 companies that were, in turn, divided into the three studied sectors: manufactures, services and trade, and construction and civil works. The remaining 30% of the sample was used to test the model. For this purpose, 96 randomly selected companies were divided into the three studied business sectors, in order to prevent over-adjustment. Table 4 shows the numbers of enterprises that were randomly selected, as grouped by size and productive sector.

The training data were used to ensure that the supervised classification algorithms which recognized patterns and characteristics in the data, thus ensuring a more accurate and efficient performance. Thus, the 10-fold cross validation technique explained in the previous section was used to train the classification algorithms through SVM, RF and NV. Out of the 231 companies that made up 70% of the sample (which were assigned to training and validation), 70% were ascribed to training and 30% to testing. In other words, out of the 231 companies, 162 companies are used to train the classification algorithms and 69 companies to validate the efficiency of the algorithms.

The results of the training and validation process of the SMMM model (see Fig. 7) are shown in Table 5. When the SVM, NV and RF algorithms were compared, the one which showed the highest score and best performance was Random Forests (RF), with 93.10% F1-Score, 93.20% Accuracy, 93.67% Recall, and 93.36% Precision.

On the other hand, data analytics uses the "Confusion Matrix" procedure to visualize and determine the performance of the classification algorithms that had previously been developed through supervised learning. Fig. 8 shows the confusion matrix of each of the evaluated algorithms. The main diagonals show the accuracy of the algorithms in predicting each one of the evaluated levels, and RF is able to correctly predict the sustainability maturity level of a new sample of businesses with an accuracy of 93.36%.

Finally, it was found, after taking into account the parametric and non-parametric tests, that the best classification model was RF, as can be observed in Table 6. An alpha level or significance level of $p < 0.5$ was considered for the parametric Student's t-test. Also the non-parametric test of RF shows a higher value of the F1-score, in accordance with the false positives of the confusion matrix.

4.3. Results on the test dataset

The dataset used to apply the test was taken from the aforementioned randomly selected group of companies that made up 30% of the sample. Additionally, the best classification algorithm (RF) was used to group the companies according to their level of sustainability maturity, as indicated by the SMMM. The maturity level classification is illustrated, by productive sector, in Fig. 9.

5. Discussion

As a result of the application of the SMMM model, it can be observed that there is a small number of companies in level 1 (*i.e.*, *insufficient sustainability maturity*). Nine percent of these companies belong to the

Table 5
Algorithm classification results.

Classifier	F1-Score (%)	Accuracy (%)	Recall (%)	Precision (%)
Support vector machine (SVM)	90.13	90.34	90.99	90.92
Naïve Bayes (NV)	92.73	92.84	93.31	92.99
Random Forest (RF)	93.10	93.20	93.67	93.36

Table 4
The numbers of MSMEs assigned by productive sector.

Enterprise size	Manufacturing sector		Service and trade sector		Construction and civil works sector	
	No. of training and validation companies	No. of test companies	No. of training and validation companies	No. of test companies	No. of training and validation companies	No. of test companies
Medium	16	11	9	8	3	3
Small	37	11	36	12	13	13
Micro	42	10	49	12	26	16

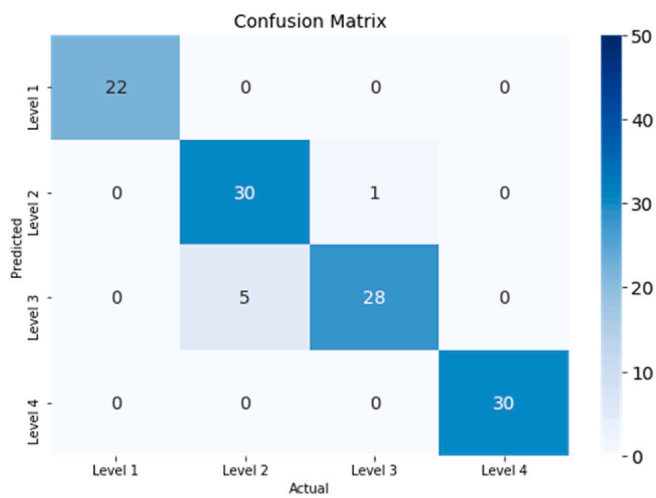

Fig. 8. Confusion matrix of the Random Forest (RF) algorithm.

Table 6
Selection of the best model.

Classification algorithms	F1-Score (%)	Student's t-test
Support Vector Machine (SVM)	90.13	0.84
Naïve Bayes (NV)	92.73	0.89
Random Forest (RF)	93.10	0.86

service and trade sector, 6% of them are in the construction and civil works sector, and 3% in the manufacturing sector. According to the model, these companies have an insufficient level of knowledge management, environmental practices and management systems. In other words, they lack environmental strategies and motivations that allow them to mitigate the effects of their operations and products on the environment.

Level 2 pertains to companies with basic sustainability maturity. Most of them came from the manufacturing sector (38%), followed by the service and trade and the construction and civil works sectors, each of which harbored 28% of organizations at this level. This means that the companies in level 2 do not have a structured environmental management system, but they do have basic environmental strategy knowledge. This allows them to design fundamental environmental care practices and policies to minimize likely negative impacts within their internal processes.

The highest percentages were observed at level 3, which refers to companies with developing sustainability maturity. At this level, the construction and civil works sector accounted for 56% of the businesses, the manufacturing sector for 44%, and the service and trade sector for 34%. Organizations at this level are characterized by the progress and effort they have made, not only in the development of internal environmental strategies, but also in investing their financial capital in human, technological and physical resources to acquire environmental knowledge and improve operational processes.

Level 4 accounts for companies with consolidated sustainability maturity, which characteristically exhibit optimal management and sustainability strategies. The largest fraction at this level corresponds to the service and commerce sector, which accounts for 28% of these companies. The manufacturing sector instead has 16%, and is followed by the construction and civil works sector, with 9%. Companies at this level possess quality management or environmental management certification, in addition to environmental practice programs and structured management indicators that allow them to provide and control information on the environmental performance of their operations. Therefore, the results of the classification of the companies in different sustainability maturity levels allowed us to establish that they are aware of the importance of environmental strategies and practices when implemented in internal processes. However, they exhibit weaknesses regarding the mechanisms used to strengthen environmental plans or programs.

Considering the above, and following the results also highlighted in the previous works of (Hsu et al., 2017; Aboelmaged, 2018; Caldera et al., 2019; Matinaro et al., 2019), in order for MSMEs to gradually move through each of the proposed levels of sustainability maturity, it is recommended that they strengthen their business associations and cooperative activities. This can be done by designing technological applications that allow them to identify and characterize all the elements they use in their internal operations, so that they can be continuously reused in different stages of their operations.

Another result of this research is the design of the SMMM system, which provides support for both business sustainability maturity assessments and decision-making in companies and governmental entities. The proposed system consists of a series of activities, as shown in Fig. 10. First, a company's user conducts a self-diagnosing survey through a web application. The system then captures the information and processes the data and features of the company, in terms of environmental knowledge management, environmental practices and environmental management systems. Subsequently, the system takes the extracted features and sends them automatically to the RF classification algorithm, which matches them with those that define the SMMM levels, and then estimates the maturity of the company. Finally, the system provides a series of suggestions and guidelines for the company to follow in order to increase its maturity level.

6. Conclusions and future research perspectives

Micro, small and medium enterprises (MSMEs) constitute the backbone of the business in several countries, and particularly in emerging realities. However, in addition to some common structural and organizational weak points, the lack of guidelines and/or standards available to support environmental sustainability implementation in the internal processes of businesses appears to be a major concern in the MSME sector. In this context, only a few studies and models available in the literature have holistically accounted for the key elements of sustainable development. Therefore, this research has introduced a new sustainability maturity model for MSMEs (SMMM), which helps companies to

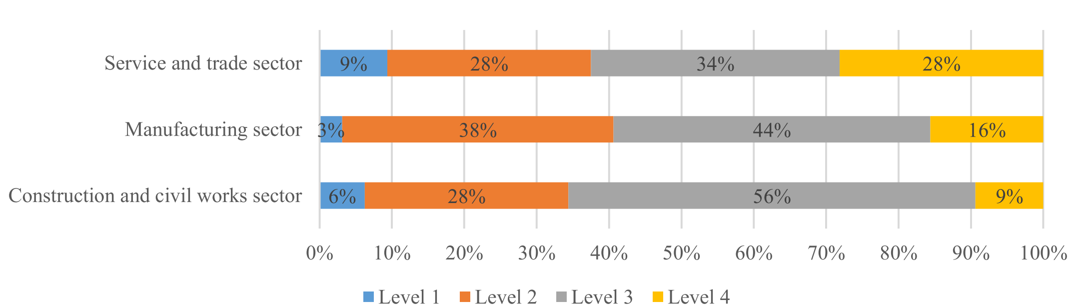

Fig. 9. Classification of the sample according to the business sectors in the SMMM.

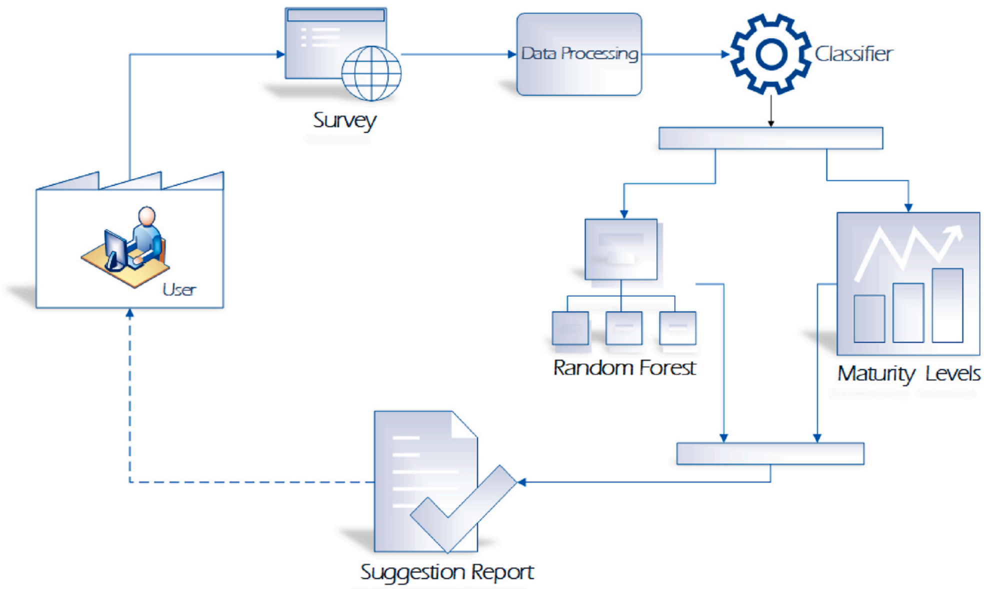

Fig. 10. The SMMM system.

detect their environmental sustainability strengths and weaknesses. The model is divided into four levels that integrate environmental knowledge management compliance, environmental practices, and management systems. The SMMM has been applied to a group of companies in Colombia, using the DSR and Design of Computational Intelligence Experiments techniques. Making use of data analytics and supervised classification algorithms, the model found that the studied group of Colombian MSMEs was composed of 6% of companies that were at an insufficient sustainability maturity level, 31% at a basic level, and 45% and 18% were at the developing and consolidated levels, respectively. The case study highlighted, for several companies, the need to progress through sequential levels of implementation of sustainable strategies. Since the model can be used to assesses environmental knowledge management, environmental practices and management systems, it can be adopted by MSMEs to continue strengthening such capabilities. In this way, companies will have the possibility of advancing their internal processes in order to implement good environmental practices and strategies that have balanced economic, social and environmental aspects. It is worth remarking that, regarding the potential replicability of this research, the approach here proposed can be adapted for empirical investigations and extended to other relevant industry domains/sectors. Moreover, future research should contemplate not only the evaluation of external aspects concerning corporate environmental issues, but also extend the sampling to other cities, in order to assess sustainable business maturity at the national level. It is suggested that the SMMM system should be used by MSMEs through government entities or business associations. In such a context, the model constitutes an agile tool to allow strength-and-weakness-based business decision-making and evaluation to be made, in order to establish continuous (and sustainable) improvement mechanisms.

CRediT authorship contribution statement

Jenifer Vásquez: Writing – original draft, Conceptualization, Methodology, Data curation, Visualization, Investigation. **Santiago Aguirre:** Writing – original draft, Supervision. **Edwin Puertas:** Data curation, Validation, Software, Writing – original draft, Investigation. **Giulia Bruno:** Writing – original draft, Writing – review & editing. **Paolo C. Priarone:** Writing – original draft, Supervision, Writing – review & editing. **Luca Settineri:** Writing – original draft, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are grateful to all the MSMEs involved in this research. Without their cooperation, this study would not have been possible. The authors are thankful for the support they received to carry out this study from the CEIBA Colombia foundation and Bolivar governmental entities that promote science and technology development in Colombia.

References

Aboelmaged, M., 2018. The drivers of sustainable manufacturing practices in Egyptian SMEs and their impact on competitive capabilities: a PLS-SEM model. J. Clean. Prod. 175, 207–221. https://doi.org/10.1016/j.jclepro.2017.12.053.

Alpaydin, E., 2016. In: Machine Learning : the New AI. MIT Press, Cambridge, MA.

Arbolino, R., De Simone, L., Carlucci, F., Yigitcanlar, T., Ioppolo, G., 2018. Towards a sustainable industrial ecology: implementation of a novel approach in the performance evaluation of Italian regions. J. Clean. Prod. 178, 220–236. https://doi.org/10.1016/J.JCLEPRO.2017.12.183.

Backlund, F., Chronéer, D., Sundqvist, E., 2014. Project management maturity models – a critical review. Procedia - Soc. Behav. Sci. 119, 837–846. https://doi.org/10.1016/j.sbspro.2014.03.094.

Bancóldex, 2017. Clasificación de empresas en Colombia.

Biscotti, A.M., D'Amico, E., Monge, F., 2018. Do environmental management systems affect the knowledge management process? The impact on the learning evolution and the relevance of organisational context. J. Knowl. Manag. 22, 603–620. https://doi.org/10.1108/JKM-08-2017-0344.

Brundtland, G.-H., 1987. Our Common Future: the World Commission on Environment and Development. Oxford University Press, New York.

Bureau van Dijk, 2020. Bureau van Dijk | private company information – Orbis [WWW Document]. URL. https://www.bvdinfo.com/en-gb. accessed 2.27.20.

Caldera, H.T.S., Desha, C., Dawes, L., 2019. Evaluating the enablers and barriers for successful implementation of sustainable business practice in 'lean' SMEs. J. Clean. Prod. 218, 575–590. https://doi.org/10.1016/j.jclepro.2019.01.239.

Carmona-Calvo, M.A., Suárez, E.M., Calvo-Mora, A., Periáñez-Cristóbal, R., 2016. Sistemas de gestión de la calidad: un estudio en empresas del sur de España y norte de Marruecos. Eur. Res. Manag. Bus. Econ. 22, 8–16. https://doi.org/10.1016/j.iedee.2015.10.001.

Čater, B., Čater, T., Prašnikar, J., Ivašković, I., 2018. Environmental strategy and its implementation: what's in it for companies and does it pay off in a posttransition context? J. East Eur. Manag. Stud. 23, 55–88. https://doi.org/10.5771/0949-6181-2018-1-55.

Dharmawan, Y.S., Divinagracia, G.G., Woods, E., Kwong, B., 2019. Inter-dependencies on BPM maturity model capability factors in deriving BPM roadmap. Procedia Comput. Sci. 161, 1089–1097. https://doi.org/10.1016/j.procs.2019.11.220.
Edvardsson, I.R., Yigitcanlar, T., Pancholi, S., 2016. Knowledge city research and practice under the microscope: a review of empirical findings. Knowl. Manag. Res. Pract. 14, 537–564. https://doi.org/10.1057/s41275-016-0003-0.
EFQM, 2020. EFQM model [WWW Document]. URL. https://www.efqm.org/index.php/efqm-model/. accessed 6.17.20.
Ellefsen, A.P.T., Oleśków-Szłapka, J., Pawłowski, G., Toboła, A., 2019. Striving for excellence in ai implementation: ai maturity model framework and preliminary research results. Logforum 15, 363–376. https://doi.org/10.17270/J.LOG.2019.354.
Fernandez-Lozano, C., Gestal, M., Munteanu, C.R., Dorado, J., Pazos, A., 2016. A methodology for the design of experiments in computational intelligence with multiple regression models. PeerJ 2016 1–22. https://doi.org/10.7717/peerj.2721.
Froger, M., Bénaben, F., Truptil, S., Boissel-Dallier, N., 2019. A non-linear business process management maturity framework to apprehend future challenges. Int. J. Inf. Manag. 49, 290–300. https://doi.org/10.1016/j.ijinfomgt.2019.05.013.
Fushiki, T., 2011. Estimation of prediction error by using K-fold cross-validation. Stat. Comput. 21, 137–146. https://doi.org/10.1007/s11222-009-9153-8.
Golinska, P., Kuebler, F., 2014. The method for assessment of the sustainability maturity in remanufacturing companies. In: Procedia CIRP. Elsevier, pp. 201–206. https://doi.org/10.1016/j.procir.2014.06.018.
Heikkurinen, P., Young, C.W., Morgan, E., 2019. Business for sustainable change: extending eco-efficiency and eco-sufficiency strategies to consumers. J. Clean. Prod. 218, 656–664. https://doi.org/10.1016/J.JCLEPRO.2019.02.053.
Hsu, C.H., Chang, A.Y., Luo, W., 2017. Identifying key performance factors for sustainability development of SMEs – integrating QFD and fuzzy MADM methods. J. Clean. Prod. 161, 629–645. https://doi.org/10.1016/j.jclepro.2017.05.063.
ISO, 2020. ISO - certification [WWW document]. ISO. https://www.iso.org/certification.html accessed 6.17.20.
ISO, 2015. ISO 14001:2015 Environmental management systems — requirements with guidance for use. ISO/TC 207/SC 1 Environ. Manag. Syst [WWW Document]. https://www.iso.org/standard/60857.html. accessed 2.27.20.
Johnson, M.P., 2017. Knowledge acquisition and development in sustainability-oriented small and medium-sized enterprises: exploring the practices, capabilities and cooperation. J. Clean. Prod. 142, 3769–3781.
Johnstone, L., 2020. A systematic analysis of environmental management systems in SMEs: possible research directions from a management accounting and control stance. J. Clean. Prod. 244 https://doi.org/10.1016/j.jclepro.2019.118802.
Jové-Llopis, E., Segarra-Blasco, A., 2018. Eco-efficiency actions and firm growth in European SMEs. Sustain. Times 10. https://doi.org/10.3390/su10010281.
Klyueva, I., 2019. Improving quality of the multiclass SVM classification based on the feature engineering. In: Proceedings - 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency, SUMMA 2019. Institute of Electrical and Electronics Engineers Inc., pp. 491–494. https://doi.org/10.1109/SUMMA48161.2019.8947599
Lacerda, T.C., von Wangenheim, C.G., 2018. Systematic literature review of usability capability/maturity models. Comput. Stand. Interfac. https://doi.org/10.1016/j.csi.2017.06.001.
Lin, T.C., Wang, K.J., Sheng, M.L., 2020. To assess smart manufacturing readiness by maturity model: a case study on Taiwan enterprises. Int. J. Comput. Integrated Manuf. 33, 102–115. https://doi.org/10.1080/0951192X.2019.1699255.
Lopes, C.M., Scavarda, A., Hofmeister, L.F., Thomé, A.M.T., Vaccaro, G.L.R., 2017. an analysis of the interplay between organizational sustainability, knowledge management, and open innovation. J. Clean. Prod. 142, 476–488. https://doi.org/10.1016/j.jclepro.2016.10.083.
Matinaro, V., Liu, Y., Lee, T.R., Poesche, J., 2019. Extracting key factors for sustainable development of enterprises: case study of SMEs in Taiwan. J. Clean. Prod. 209, 1152–1169. https://doi.org/10.1016/j.jclepro.2018.10.280.
Moreno-Sandoval, L.G., Puertas, E., Plaza-Del-Arco, F.M., Pomares-Quimbaya, A., Alvarado-Valencia, J.A., Alfonso Ureña-López, L., 2019. Celebrity profiling on twitter using sociolinguistic features notebook for PAN at CLEF 2019. CEUR Workshop Proc.
Mukherjee, A., Mukhopadhyay, S., Panigrahi, P.K., Goswami, S., 2019. Utilization of oversampling for multiclass sentiment analysis on amazon review dataset. In: 2019 IEEE 10th Int. Conf. Aware. Sci. Technol. iCAST 2019 - Proc, pp. 1–6. https://doi.org/10.1109/ICAwST.2019.8923260.
Nascimento, G., Araujo, C.A.S., Alves, L.A., 2017. Corporate sustainability practices in accredited Brazilian hospitals: a degree-of-maturity assessment of the environmental dimension. Rev. Adm. 52, 26–35. https://doi.org/10.1016/j.rausp.2016.10.001.
Odważny, F., Wojtkowiak, D., Cyplik, P., Adamczak, M., 2019. Concept for measuring organizational maturity supporting sustainable development goals. Logforum 15, 237–247. https://doi.org/10.17270/J.LOG.2019.321.
Páez, G., Rohvein, C., Paravie, D., Jaureguiberry, M., 2018. Revisión de modelos de madurez en la gestión de los procesos de negocios. Ingeniare Rev. Chil. Ing. 26, 685–698. https://doi.org/10.4067/s0718-33052018000400685.
Parker, D.W., Loh, A., Chevers, D., Minto-Coy, I., Zeppetella, L., 2017. Operations sustainability maturity model: preliminary findings of financial services in developing and developed countries. Meas. Bus. Excell. 21, 309–336. https://doi.org/10.1108/MBE-08-2016-0044.
Patil, S., Sonavane, S., 2020. Minority–majority mix mean oversampling technique: an efficient technique to improve classification of imbalanced data sets. In: Advances in Intelligent Systems and Computing. Springer, pp. 501–509. https://doi.org/10.1007/978-981-32-9515-5_48.
Plasencia Soler, J.A., Marrero Delgado, F., Bajo Sanjuán, A.M., Nicado García, M., 2018. Modelos para evaluar la sostenibilidad de las organizaciones. Estud. Gerenciales 34, 63–73. https://doi.org/10.18046/j.estger.2018.146.2662.
Santos-Neto, J.B.S. dos, Costa, A.P.C.S., 2019. Enterprise maturity models: a systematic literature review. Enterprise Inf. Syst. 13, 719–769. https://doi.org/10.1080/17517575.2019.1575986.
Schwan, G., 2019. Sustainable development goals. GAIA. https://doi.org/10.14512/gaia.28.2.1.
Shi, X., Baba, T., Osagawa, D., Fujishima, M., Ito, T., 2019. A maturity model for sustainable system implementation in the era of smart manufacturing. IEEE Int. Conf. Emerg. Technol. Fact. Autom. ETFA 2019-Septe 1649–1652. https://doi.org/10.1109/ETFA.2019.8869446.
Tsiliki, G., Munteanu, C.R., Seoane, J.A., Fernandez-Lozano, C., Sarimveis, H., Willighagen, E.L., 2015. RRegrs: an R package for computer-aided model selection with multiple regression models. J. Cheminf. 7, 46. https://doi.org/10.1186/s13321-015-0094-2.
Vásquez, J., Aguirre, S., Fuquene-Retamoso, C.E., Bruno, G., Priarone, P.C., Settineri, L., 2019. A conceptual framework for the eco-efficiency assessment of small- and medium-sized enterprises. J. Clean. Prod. 237 https://doi.org/10.1016/j.jclepro.2019.117660.
Von Scheel, H., Von Rosing, G., Skurzak, K., Hove, M., 2014. BPM and maturity models. In: The Complete Business Process Handbook: Body of Knowledge from Process Modeling to BPM. Elsevier Inc., pp. 395–426. https://doi.org/10.1016/B978-0-12-799959-3.00019-7
Watz, M., Hallstedt, S.I., 2020. Profile model for management of sustainability integration in engineering design requirements. J. Clean. Prod. 247, 119155. https://doi.org/10.1016/j.jclepro.2019.119155.
WBCSD, 2009. Learning module [WWW Document]. URL. http://www.wbcsd.org/Projects/Education/Resources/Eco-efficiency-Learning-Module. accessed 2.5.18.
Weber, R., 2017. Design-science research. In: Research Methods: Information, Systems, and Contexts, second ed., pp. 267–288. https://doi.org/10.1016/B978-0-08-102220-7.00011-X.