IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received February 14, 2019, accepted March 7, 2019, date of publication March 12, 2019, date of current version March 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2904242

Causal Analysis of the Spanish Industrial Sector Through Smarta

MIGUEL LLORET-CLIMENT, JOSUÉ-ANTONIO NESCOLARDE-SELVA, (Member, IEEE), KRISTIAN ALONSO-STENBERG, AND MARTHA-ARELIS SELVA-BARTHELEMY
Department of Applied Mathematics, University of Alicante, 03690 Alicante, Spain

Corresponding author: Josué-Antonio Nescolarde-Selva (josue.selva@ua.es)

This work was supported in part by the Valencian Government, under Project GV/2018/061.

ABSTRACT When examining a sector of the economy, it can be sometimes difficult to identify the relationships between the underlying variables that compose it. Therefore, we developed a causal analysis technique, capable of converting large amounts of data into a directed graph of cause-effect relationships. The main objective of the technique is to locate the attractors associated with the system, that is, the sets of variables toward which the system tends dynamically. This methodology is based not only on General System Theory, but also on the Graph Theory and a discrete version of Chaos Theory. However, when systems have a large number of variables, applying the technique can be a tedious task. We thus implemented the Smarta application, a causal analysis simulator that allows automating this methodology. The software constitutes a reimplementation and continuation of the application already developed by our research group. We conducted a causal analysis of a system extracted from a database of structural statistics of Spanish industrial sector companies between 2008 to 2015 (the data were obtained from Spain's National Institute of Statistics). We focused on the yearly analysis of companies' structural and economic properties, based on 21 proxy variables. Based on the proposed analysis, we attempted to answer the following questions: how were the survey variables causally related? Were there any groups of independent variables within the system? And what trends did the system follow over the 2008-2015 period? The aim was to propose an alternative to classical statistical methods employed until now.

INDEX TERMS Attractor, causality, industrial sector, smarta.

I. INTRODUCTION

A. BACKGROUND

A great scientific challenge throughout history has been that of creating mathematical models able to accurately describe different complex systems found in nature. During the second half of the twentieth century, various experts felt the need to elaborate a theory that would allow these systems to be modelled, but they did so from a global perspective without taking into account the elements or the relationships that composed them.

Ludwig Von Bertalanffy founded General System Theory (hereon GST) in which he defined the basic principles underlying these types of systems [1], [2]. The author's conception of a system was that of a set of elements that were not only interrelated among themselves, but also with their environment.

Other authors, such as M. D. Mesarovic and Y. Takahara, formulated their own version of GST based on Set Theory, according to which a system is defined as a series of non-empty sets [3]. Later, they published an article on their Abstract Systems Theory [4], in which they introduced two major concepts: the abstract system (which includes the system definition above) and the complex system (a system constituted by several subsystems).

However, the problem is that many systems in real life such as economic and social systems are actually very difficult to model using the complex system model. To address the problem, Y. H. Ma and Y. Lin developed a new system definition based on two components: the object set, which collects all the objects or variables of the system; and the relationship set, which includes all the relations between the objects of the system [5]. This definition is visibly similar to that currently used in Graph Theory [6].

Furthermore, the concept of causality, i.e. the ability to predict future consequences of a given event is fundamental.

The associate editor coordinating the review of this manuscript and approving it for publication was Jianquan Lu.

The idea can be easily extrapolated to the field of GST, where causality can be defined as the ability of one variable to be the cause of another within a complex system. In 1956, N. Wiener postulated that an initial variable can be considered to have a causal relationship with another if the prediction capacity of the second variable is improved when introducing information relating to the first variable [7]. Moreover, it was not until 1969 that C. Granger published a mathematical implementation of this idea applied to the field of linear autoregressive models of stochastic processes [8]. This technique is nowadays highly applicable to a range of fields including neuroscience. S. Bressler and A. Seth claim to have obtained significant results based on Wiener-Granger causality that helped them to understand the neural basis of cognition [9]. For his part, P. M. Senge states that although we see cause-effect relationships as linear (linear causality), reality is in fact constituted by circular relations (circular causality) [10].

B. RELATED WORK

During the decade of the 50s appears the theory of Systems Dynamics (SD), whose father was Forrester [11]–[13]. The SD arises from the problem that the General Electric company proposes to Forrester, since they needed to model and understand the economic and labor dynamics of some of their factories [14]. Thus, this methodology allows modeling and simulating complex systems present in numerous areas (economics, social sciences, biology, engineering, environment, etc.), with the main objective of understanding its operation, analyzing scenarios and alternatives and helping in decision making [15].

In 2000, Lin and Liu [16], [17] proposed analyzing complex systems using data and applied the technique to the prediction of dry and warm winds to tackle the problem posed by winds for crop growth. Three years later, Csató *et al.* [18] advanced a technique to analyze probabilistic data models using a large number of hidden variables. These models aim at explaining the complexity of the observed data by means of hidden or unobservable causes modelled as random variables. In 2011, Li *et al.* [19] published a method to describe, express and distinguish time series through complex network graphs based on these series' spatial distribution properties.

In 2014, Lloret and Nescolarde [20] developed a technique to convert large amounts of data into a directed graph of cause-effect relationships which was then applied to a qualitative version of Chaos Theory. The technique aimed at finding the complex system's attractor sets or trends. A major advantage of the technique is that it allows to predict the behavior of the system dynamically, according to time series updates. As in the SD, this methodology also makes use of causal graphs; however, on this occasion the cause-effect relationships are calculated from the data, while in the SD they should be established initially. During that year, they applied this methodology to ecosystems and biological pest control in Mediterranean greenhouses, where they modelled different ecological systems adopting an alternative approach, making use of the concepts of structural function, coverage and invariability [21]. They also obtained important results when applying causality to different variables underlying national tourist flows [22]. In 2015, Alonso-Stenberg *et al.* [23] conducted a causal analysis study using the 73.3 Eurobarometer database of June 2010, based on a survey of 26,602 EU citizens on the potential health effects of electromagnetic fields and other environmental and health factors.

However, the applications of Chaos Theory are numerous and not only limited to the phenomenon of causality; we also find examples in the economy and networks.

In recent works, Akkaya *et al.* analyze the monetary dynamics of Bitcoin for the period 2011-2014, through a technique that combines delay time, embedding dimension and maximal Lyapunov exponents. When they find that this exponent is maximum and positive, then there could be a chaotic behavior in the monetary dynamics of Bitcoin [24].

Rusyn and Savko have managed to control the chaotic behavior of a particular economic model: Cournot's duopoly model (presented by M. Kopel). Specifically, they do this by adding a new time function to the model called the state feedback controller [25].

Haley states that it is possible to predict non-recurring economic cycles if chaotic Sprott systems are applied. Even so, such cycles could disappear if a certain short-term interest value is established [26].

Harikrishnan *et al.* have been able to create a recurrent complex network from chaotic time series. Unlike recurrent networks that are usually obtained in this way (which are undirected and unweighted), the authors get weighted recurrent networks, which facilitates the discrimination of one chaotic time series from another that is noisy [27].

If we compare the technique used by Harikrishnan *et al.* with that proposed in this paper, we can see that one of the similarities is that in both cases we turn the time series into a complex network. However, in the case of Harikrishnan *et al.* we obtain a network that is undirected and weighted, whereas in our methodology the graphs obtained are always directed, with the possibility that the relationships have weights or not. In our technique, the network is obtained through the use of correlation and causality, while in Harikrishnan's approach *et al.* it is obtained through recurrence. Finally, the approach we propose has as its final objective the search for the system's attractor sets, while that of Harikrishnan *et al.* pursues the characterization of the structure of chaotic attractors.

In the disciplines of networks and economics, A. S. García's new approach is interesting. He uses it to carry out an analysis of the Spanish economy for the period 2000-2005, discovering that there have been important structural changes during this period. The proposed method makes use of the concept of structural equivalence present in network theory, but applied to the economic field of input-output [28].

P. Balland *et al.* have used novel network dynamic models to study how the behavior of two types of networks evolves: technical knowledge and business knowledge. By using this methodology in the case of a toy cluster in Spain, they have perceived that the two types of networks have different behavior [29].

In the Spanish industrial context, Köhler tries to restore industrial relations from the end of the dictatorship to the present day. He also analyses the existing dilemmas of social and political actors in the period of the economic crisis [30].

Finally, regarding causality and cybernetics, Chen *et al.* [31] propose a good graphical model of multivariable alarm prediction applied to cyber-physical systems. This model is based on multivariate causal analysis and network parameter learning. The great advantage of this model is that it is capable of: (a) accurately predict possible future alarm events, (b) detect failures in the system and (c) find the origin of these errors.

C. MOTIVATION

Determining cause-effect relationships and locating attractors can be truly arduous when working on systems with a large number of variables. An application capable of automating the entire process is therefore essential: this is how Smarta software was born. Smarta allows entering the variables of the complex system under study; based on the data, it asks the user for different parameters that it will use to calculate cause-effect pairs and represent an interactive directed graph as well as determine the system's trends, in terms of structural functions, coverage, invariability, orbits, attractors, and basins of attraction. The application comes in two versions: (a) the desktop version, developed in the C++ language and (b) the web version in PHP and JavaScript [32]. Both versions are follow-ups of the software developed by M. Lloret, P. Esteve, and E. Almenara in 2009 [33].

In this study, we conducted a causal analysis of a system based on a database of structural statistics of companies part of Spain's industrial sector (obtained from the Spanish National Institute of Statistics, or INE by its Spanish acronym), using the 2008 to 2015 time series. The survey focused on companies' annual structural and economic properties, based on 21 proxy variables grouped into 5 categories: occupied personnel, income, stock variations, costs, and investments. Based on this analysis, we attempted to answer the following questions: how are the variables of the survey causally related? Are there any groups of independent variables within the system? And, what trends does the system follow during the study period? We used the Smarta causal simulator, which allowed us to obtain results automatically and helped us to interpret them. We thus advance an alternative to classical statistical methods.

II. METHODOLOGY

A. THE PROPOSED TECHNIQUE

The technique used in the present study is outlined in Fig. 1. As illustrated, the first step consists in selecting the variables to study and that will make up the complex system. In our example, these variables are: $A, B, \ldots, G$ (Fig. 1, step 1).

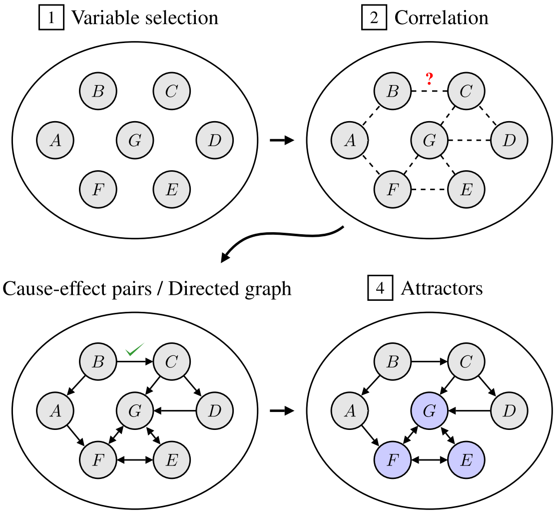

FIGURE 1. Outline of the proposed causal analysis.

Next, the correlation between each pair of system variables is calculated using the Pearson coefficient. Pairs of variables whose correlation exceeds a certain threshold are then validated and the rest are discarded. As can be seen, relationships between correlated pairs are marked with a dashed line (Fig. 1, step 2). This is due to the fact that a relationship may exist between correlated pairs, but the direction of the relationship is unknown.

The next step is to determine whether each pair variable behaves as a cause, as an effect or as both. For example, in the diagram, $A \rightarrow F$ indicates that variable A is the cause of F (unidirectional relationship), while $F \leftrightarrow E$ indicates that F is the cause of E, and in turn, E is the cause of F (bidirectional relationship).

Once all cause-effect pairs are defined, we can represent them by means of a directed graph, where graph nodes represent the variables and the edges of the graph illustrate the relationships (Fig. 1, step 3). The objective of this graph it to visualize the system and thus facilitate its analysis.

Finally, we analyze the graph to locate the system's attractors as well as other properties related to discrete Chaos Theory. In Fig. 1 (step 4), the set $\{E, F, G\}$ is highlighted. Indeed, if we take any variable in the system and follow its chain of relationships, we will always end up in this set, which is why this set constitutes the system's attractor.

B. SYSTEM ATTRACTORS AND BASINS OF ATTRACTION

An *attractor* is a set in a system capable of attracting the rest of the variables that are part of its basin of attraction. Attractors correspond to areas that delimit the variables' apparently disorganized behavior, so they are fundamental to predict a complex system's behavior or trend over time. For its part, the *basin of attraction* is the set formed by all the variables from which the attractor can be reached, including the attractor itself. Therefore, the basin of attraction is an indicator of the attractor's area of influence within a system. If we

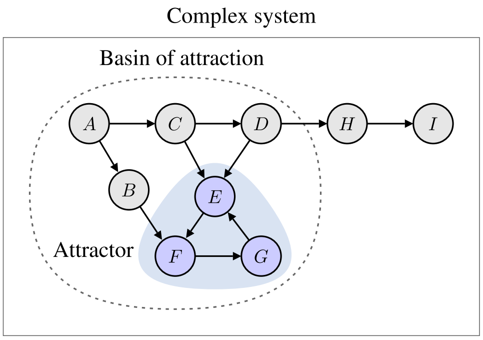

FIGURE 2. **Concept of attractor and basin of attraction. The system attractor $\{E, F, G\}$ can be reached from any variable $\{A, B, C, D, E, F, G\}$ of the basin of attraction.**

consider the directed graph of the complex system in Fig. 2, formed by the variables $\{A, B, C, D, E, F, G, H, I\}$, we can see that most of these variables (all except H and I) have causal links towards the set $\{E, F, G\}$, which corresponds precisely to the system's attractor. As we can see, the attractor forms at least one causal circle or loop that prevents them leaving and going to other parts of the system.

We can also see that the basin of attraction of this latter attractor would be formed by the variables $\{A, B, C, D, E, F, G\}$, because if we follow the chain of relationships of any of these variables, we can always reach the attractor $\{E, F, G\}$. This does not apply to the H and I variables, since they do not lead to the attractor, which is why they do not belong to the basin of attraction. Furthermore, we can see that the basin of attraction will always contain its corresponding attractor, so the size of the basin will always be greater than or equal to that of the attractor, and never below it (*attractor* $\subseteq$ *basin of attraction*).

Significantly, not all complex systems need to have attractors, but when attractors are present, each will have an associated basin of attraction. A complex system usually consists of one or more subsystems that are isolated from one another (also called *independent sets*), so each independent set can have a maximum of one attractor.

To finish, identifying attractors is particularly important when systems contain a large number of variables and data is periodically updated, as attractors enable deducing and identifying changes in trends.

C. THE CASE STUDY

First, the aim was to conduct a causal analysis of a system based on the structural statistics of Spain's industrial sector companies. Specifically, we worked on a time series from 2008 to 2015, that included different bind variables according to sectors of activity. This database was obtained from INE and can be accessed through [34]. The statistical study consisted of an annual survey of manufacturing industries, extractive industries, energy, gas and water companies, and sanitation, waste management and decontamination companies. The main objective of the survey was to identify companies' structural and economic properties, based on a set of 21 proxy analysis variables, grouped into different categories (see Table 1). The units of measurement of the variables were as follow:

- Variables related to occupied personnel: number of people (occupied personnel) and thousands of hours (worked hours).
- Variables related to income, stock variations, costs, and investments: thousands of euros (€).

TABLE 1. **Category, name and code of the variables included in companies' structural statistics (variables are coded by their acronyms).**

Variable category	*Variable name and code*
Occupied personnel variables	Occupied Personnel (OP) and Worked Hours (WH).
Income variables	Sales Turnover (ST), Product Sales (PS), Goods Sales (GS), Service Rendering (SR), Work Performed by Company (WPC), Operating Grants (OG), Other Operating Income (OOI), and Total Operating Income (TOI).
Stock variation variables	Changes in Inventories in Finished and current products (CIF) and Changes in Inventories of Raw materials, other provisions and merchandise (CIR).
Cost variables	Goods Purchase (GP), Work Performed by Other companies (WPO), Purchases and Work performed by Other companies (PWO), Staff Costs (SC), External Service Costs (ESC), Other Operating Costs (OOC), and Total Operating Costs (TOC).
Investment variables	Investments in Tangible Assets (ITA) and Investments in Intangible Assets (IIA).

The aim of the study was to uncover the causal relationships between the structural and economic variables of the national industrial sector and determine the system's behavior over the 2008-2015 period based on its attractors, basins of attraction, and independent sets.

To carry out the causal analysis, we used Smarta, a simulator allowing to automate the technique described above. The simulator was elaborated by the *Systemics, Cybernetics, and Optimization* research group. The desktop version of this simulator was implemented using C++ language because it is highly efficient. We also created a web version using the PHP and JavaScript languages. Although we made use of the desktop version in the present article, the methodology described here can also be applied to the web version.

We thus began by introducing the 21 study variables into Smarta, using the table located in the *Variables* panel (Fig. 3, step 1). Next, we proceeded with the correlation analysis for each pair of variables and established the relevant threshold. Pairs with a Pearson coefficient (in absolute value) higher than or equal to the threshold are shown in the *Correlation* panel (Fig. 3, step 2). Next, we performed the calculation

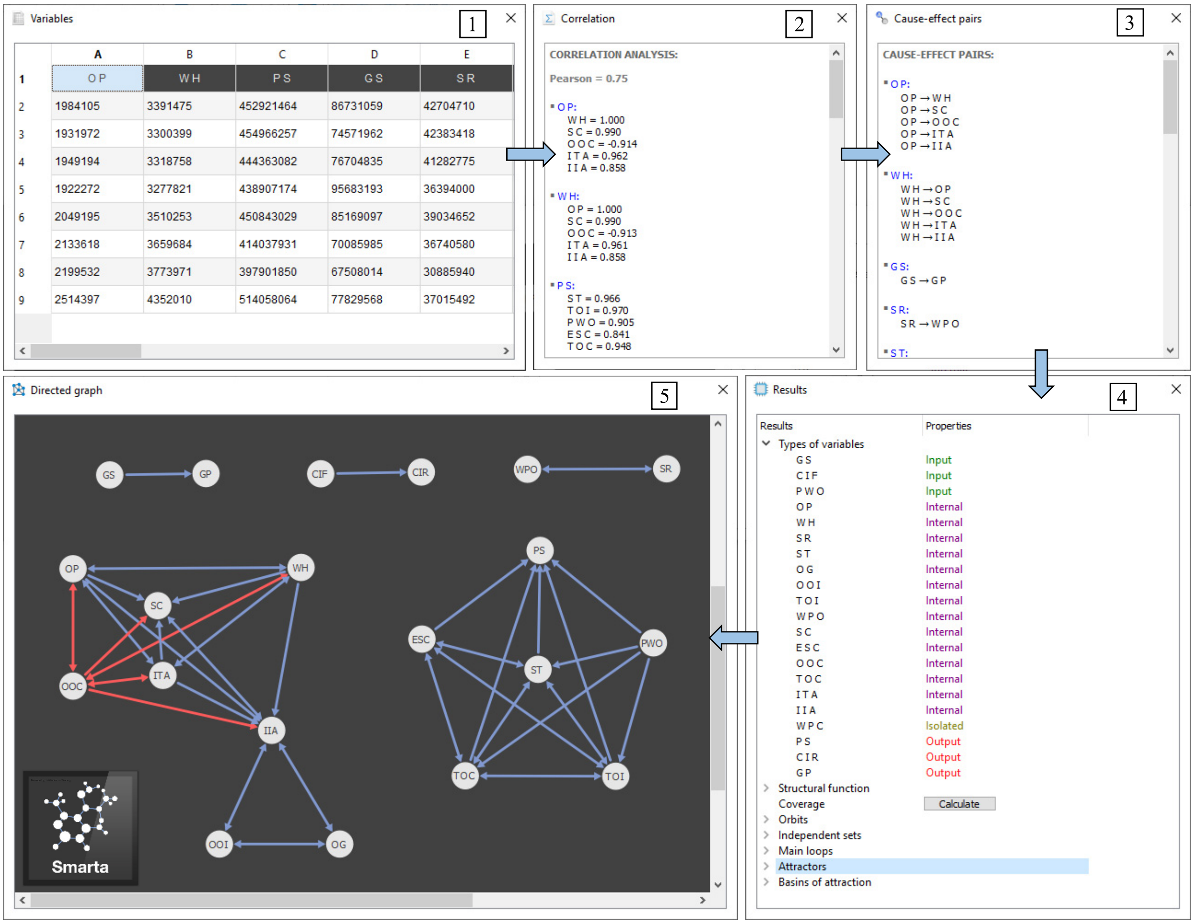

FIGURE 3. Panels and work sequence in Smarta.

of the cause-effect pairs and printed the direct influences of each variable in the *Cause-effect pairs* panel (Fig. 3, step 3). We then obtained the results of the analysis regarding the attractor sets, basins of attraction, and independent sets, among others. These results are shown in the *Results* panel, in the form of a drop-down list (Fig. 3, step 4). Finally, it was also possible to generate the interactive directed graph associated with the system, visualizing it in the *Directed graph* panel (Fig. 3, step 5), where each variable is represented by a circle and each relationship by a single or double arrow, depending on whether the relationship is unidirectional or bidirectional. A ratio marked in purple indicates a positive Pearson coefficient in the cause-effect pair, while a red color indicates a negative Pearson coefficient between the pair.

III. RESULTS

To analyze the database we used a correlation threshold of 0.75, thus obtaining pairs whose absolute value of Pearson correlation coefficient, $|r_{xy}|$, was greater than or equal to this value. The correlated pairs (a total of 36), as well as their correlation coefficient were grouped in Table 2. After conducting the causality analysis for each correlated pair, we obtained the cause-effect pairs shown in Table 3 (a total of 35). The correlation threshold value was selected in such a way as to reach an optimal compromise between the number of cause-effect pairs as strongly related as possible and the number of attractors with the greatest number of variables.

Next, we generated the directed graph associated with the system under study (Fig. 4). A total of 5 independent sets appeared, from S_1 to S_5, whose values are shown below:

$$S_1 = \{\mathrm{GS},\ \mathrm{GP}\} \tag{1}$$

$$S_2 = \{\mathrm{CIF},\ \mathrm{CIR}\} \tag{2}$$

$$S_3 = \{\mathrm{WPO},\ \mathrm{SR}\} \tag{3}$$

$$S_4 = \{\mathrm{OP},\ \mathrm{WH},\ \mathrm{OG},\ \mathrm{OOI},\ \mathrm{SC},\ \mathrm{OOC},\ \mathrm{ITA},\ \mathrm{IIA}\} \tag{4}$$

$$S_5 = \{\mathrm{PS},\ \mathrm{ST},\ \mathrm{TOI},\ \mathrm{PWO},\ \mathrm{ESC},\ \mathrm{TOC}\} \tag{5}$$

Of the 5 independent sets shown, S_1 and S_2 did not generate any attractor (since they did not contain any loops), while S_3, S_4, and S_5 generated attractors $A(S_3)$, $A(S_4)$, and $A(S_5)$:

$$A(S_3) = S_3 \tag{6}$$

$$A(S_4) = S_4 \tag{7}$$

$$A(S_5) = \{\mathrm{PS},\ \mathrm{ST},\ \mathrm{TOI},\ \mathrm{ESC},\ \mathrm{TOC}\} \tag{8}$$

TABLE 2. Pairs of variables correlated with $|r_{xy}| \geqslant 0.75$.

OP – WH (1.000)	PS – ST (0.966)	ST – ESC (0.811)	CIF – TOC (0.754)
OP – SC (0.990)	PS – TOI (0.970)	ST – TOC (0.995)	PWO – TOC (0.987)
OP – ITA (0.962)	PS – PWO (0.905)	OG – OOI (0.800)	SC – ITA (0.960)
OP – IIA (0.858)	PS – ESC (0.841)	OG – IIA (0.797)	SC – IIA (0.916)
OP – OOC (-0.914)	PS – TOC (0.948)	OOI – IIA (0.789)	SC – OOC (-0.900)
WH – SC (0.990)	GS – GP (0.984)	TOI – PWO (0.974)	ESC – TOC (0.800)
WH – ITA (0.961)	SR – WPO (0.859)	TOI – ESC (0.826)	OOC – ITA (-0.944)
WH – IIA (0.858)	ST – TOI (0.999)	TOI – TOC (0.995)	OOC – IIA (-0.796)
WH – OOC (-0.913)	ST – PWO (0.978)	CIR – CIF (0.806)	ITA – IIA (0.888)

TABLE 3. Cause-effect pairs obtained with $|r_{xy}| \geqslant 0.75$.

OP ↔ WH	WH ↔ ITA	ST ↔ TOC	CIF → CIR	ESC ↔ TOC
OP → SC	WH → IIA	OG ↔ OOI	PWO → PS	OOC → SC
OP ↔ ITA	GS → GP	OG ↔ IIA	PWO → ST	OOC ↔ ITA
OP → IIA	SR ↔ WPO	OOI ↔ IIA	PWO → TOI	OOC → IIA
OP ↔ OOC	ST → PS	TOI → PS	PWO → TOC	TOC → PS
WH → SC	ST ↔ TOI	TOI ↔ ESC	SC ↔ IIA	ITA → SC
WH ↔ OOC	ST ↔ ESC	TOI ↔ TOC	ESC → PS	ITA → IIA

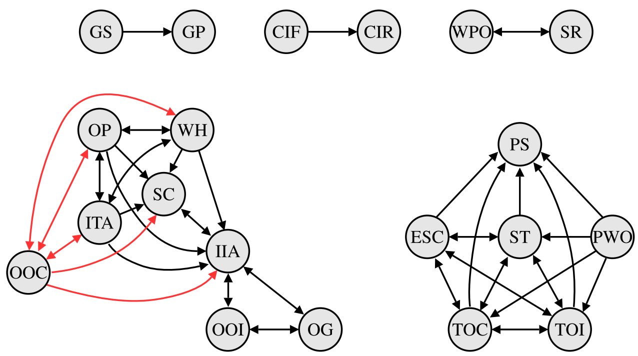

FIGURE 4. Directed graph associated with the system under study.

Moreover, each of these attractors' areas of influence would be determined by their basins of attraction, which in this case, were represented as $C(S_3)$, $C(S_4)$, and $C(S_5)$, corresponding respectively to their own sets S_3, S_4, and S_5. As we can see, these basins are characterised by the fact that when selecting a variable in any of them and following a sequence of causal iterations, we are led to its associated attractor.

Regarding the results, it is possible to observe that attractor $A(S_3)$ causally relates the variable "Service Rendering" (SR) to "Work Performed by Other companies" (WPO), indicating that the income received by the company for services rendered to third parties can favour the acquisition of external services performed by other companies, and vice versa.

On the other hand, we can see that attractor $A(S_4)$ links variables related to the categories of occupied personnel, investments, costs, and income. Specifically, occupied personnel's only two variables would be related in a bidirectional way: "Occupied Personnel" (OP) and "Worked Hours" (WH). The reason for this could be that an increase in number of hours worked would cause an increase in amount of employed personnel, and vice versa.

We also found that the variable "Investment in Tangible Assets" (ITA) directly influences the variable "Investment in Intangible Assets" (IIA), but not the other way around, so bidirectionality does not apply here. This phenomenon reflects the fact that tangible assets (stock, furniture, machinery, money, etc.) are usually more economically significant than intangible assets (know-how, copyright, brands, etc.) in companies.

We also found that both the variables relating to occupied personnel and to investment causally affected the variable "Staff Costs" (SC). It may be deduced that bigger investments in assets causes the company to expand, so a greater amount of personnel must be hired.

If we focus on income variables, as in the case of "Operating Grants" (OG) and "Other Operating Income" (OOI), we can observe a bidirectional relationship, which makes sense because they belong not only to the same category, but also to the same domain (Operation). This cause-effect pair is related to the rest of the attractor variables via the variable "Investment in Intangible Assets" (IIA).

Finally, it is necessary to note the negative influences within the attractor, which would be originated by the variable "Other Operating Costs" (OOC) (relationships coloured in red in Fig. 4). The negative relationships between OOC and "Occupied Personnel" (OP), "Worked Hours" (WH), "Staff Costs" (SC), "Investment in Tangible Assets" (ITA), and "Investment in Intangible Assets" (IIA), would thus indicate that increases in operating costs (such as equipment repairs, salaries, travel costs, advertising, etc.) would have a negative impact on other investments, such as the hiring of personnel, material goods (raw materials, stock, furniture, etc.) and intangible assets (goodwill, customer loyalty, brand influence, etc.). Therefore, it seems logical that if a company has a certain amount of resources and allocates them to certain sections, other economic aspects would be more neglected.

Next, attractor $A(S_5)$ links the variables of main income and costs. First of all, it is worth noting that a loop is formed by the variables "Total Operating Costs" (TOC), "Total Operating Income" (TOI), "External Service Costs" (ESC), and "Sales Turnover" (ST): this loop tells us that major dependency and feedback relationships exist between them. For example, an increase in turnover would cause an increase in total income based on product operations. Similarly, if a cost increase is related to external services, it would also affect the company's total operating costs. In addition, the loop would represent the feedback process, in this case, between the company's main inputs and outputs.

Secondly, we can see that although the variable "Purchases and Work performed by Other companies" (PWO) does belong to the basin of attraction $C(S_5)$, it is not included in attractor $A(S_5)$. Indeed, the PWO variable influences the other variables but it is not itself influenced by other variables, because it constitutes a starting variable and is independent from the others.

To finish, it is worth mentioning that the variable "Product Sales" (PS) acts as a drain, since not only all the variables of the basin $C(S_5)$ end in PS, but PS does not influence any variable. This reveals the importance of a company's income and costs regarding the sale of products or services to customers. Although the independent sets S_1 and S_2 do not generate any attractor, they are relevant because of the relationships between the variables. In S_1, the variable "Goods Sales" (GS) would influence "Goods Purchases" (GP), without any bidirectionality. This result may reveal that it is possible that an increase merchandise purchase does not lead to an increase in sales.

In S_2, the variable "Changes in Inventories in Finished and current products" (CIF) would be causally directed towards "Changes in Inventories of Raw materials, other provisions and merchandise" (CIR). We can thus interpret this by the fact that a reduction in stocks of products created by a company (due to sales) would also lead to a reduction in stocks of raw materials (since the products are made from them).

Finally, it is worth noting that the variable "Work Performed by Company" (WPC) is isolated: although it is considered in the initial analysis, it does not present a sufficiently robust relationship with the rest of the variables to go beyond the established correlation threshold (0.75). Therefore, it would not be part of any cause-effect pair, so we can deduce that the contribution of WPC to system trends is almost negligible.

On the other hand, results of the survey provided by the INE offer general conclusions such as the following: by 2014, the industrial sector's turnover grew by 1.7%; machinery and equipment repair and installation reached its highest turnover increase (1.5%); and 29.7% of industrial sector sales took place in foreign markets [35].

However, none of the provided results show how the variables are causally related among themselves, which independent sets of variables can be found, or the system's trends over a period of time. To address these questions, we analysed the system using the Smarta causal simulator.

IV. LIMITATIONS AND FUTURE WORK

Before concluding our paper, we would like to mention the limitations of this study, as well as the possible lines of future work.

Regarding the limitations of the study, we can point out that although Smarta is capable of working with a large number of variables (hundreds or a few thousand variables), it would be necessary to improve the application so that it could work on a larger scale (hundreds of thousands or millions of variables), in order to adapt it to the Big Data era. Likewise, this also generates the disadvantage that the larger the system under study and the more relationships are found, the more complicated can be the interpretation of the results.

Another limitation we can find is that Smarta works with quantitative variables, which is why the software uses Pearson's correlation coefficient. Therefore, the application could not currently work with categorical variables.

In addition, while the proposed technique helps us locate causal relationships within a system, there may also be external factors that are unknown but affect system variables. A final limitation is that there could be causal phenomena within companies that are not reflected in the results obtained by the proposed technique, due to the nature of the data.

With regard to possible future lines of work, we could highlight the following:

- Causal analysis of the industrial survey for the years 2016, 2017 and 2018.
- Comparison of the results when calculating cause-effect pairs using another technique, such as Granger's causality.
- Application of the technique to the tourism sector, related to the stock market.
- Enhanced application features, so that Smarta can work with a greater number and type of variables.

V. CONCLUSIONS

GST is a highly valuable tool that can be applied to the general modelling of systems in different areas of study, including economic sectors. Thanks to the technique presented in the present article, which combines GST, Graph Theory, and Discrete Chaos Theory, it is possible to discover a complex system's trends as long as attractor sets associated with the system can be determined. Attractors are groups of variables that attract the rest of the variables in their basin of attraction.

A major advantage of the proposed technique is that it allows to predict a system's behavior in an alternative and dynamic way. One setback, however, is that the technique's calculation procedure can prove to be a truly tedious, especially when large numbers of variables are present in the system. We therefore implemented the Smarta application, which helped us simplify and automate the process.

By way of synthesis, we can deduce that the trends of the system at a general level during the period 2008-2015 have been given by the set of attractors found: $A(S_3)$, $A(S_4)$ and $A(S_5)$. In this sense, we have discovered that the trends are divided into three groups that link different categories, namely:

1) Income and acquisition of services.
2) Occupied personnel, investments, costs and income.
3) Main income and principal costs.

Based on these results, it could be interesting for a company to analyze its trend network, in order to find out which variables should increase or decrease in order to optimize other variables of interest. For example, for attractor $A(S_5)$, a company could study which variable pertaining to main costs would be less expensive to promote the sale of products and services (pertaining to main income).

If we observe the period of study, we must highlight that the year 2008 coincides with the beginning of the world economic crisis (Lehman Brothers bank's collapse), while the year 2015 marks the end of this crisis in Spain. Although at present the Spanish economy has not recovered the levels prior to the crisis, we can affirm that the companies, as a whole, have contributed positively to neutralize the negative effects caused during this period.

Finally, we can emphasize that this study offers a complementary analysis to the statistical techniques used in the INE survey; and in turn, the use of Smarta allows opening new lines of work in the field of GST and the business sector.

REFERENCES

[1] L. Von Bertalanffy, *General Systems Theory: Foundations, Development, Applications*. New York, NY, USA: GeorgeBraziller, 1968, pp. 30–86.

[2] L. Von Bertalanffy, "The history and status of general systems theory," in *Trends in General Systems Theory*, G. J. Klir, Ed. New York, NY, USA: Wiley, 1972, pp. 21–41.

[3] M. D. Mesarovic and Y. Takahara, *General Systems Theory: Mathematical Foundations*. New York, NY, USA: Academic, 1975.

[4] M. D. Mesarovic and Y. Takahara, *Abstract Systems Theory*. New York, NY, USA: Springer-Verlag, 1989.

[5] Y.-H. Ma and Y. Lin, "Some properties of linked time systems," *Int. J. Gen. Syst.*, vol. 13, no. 2, pp. 125–134, 1987.

[6] K. M. Koh, F. M. Dong, and E. G. Tay, *Introduction to Graph Theory*. Singapore: World Scientific, 2007, pp. 6–7.

[7] N. Wiener, "The theory of prediction," in *Modern Mathematics for Engineers*, E. Beckenbach Ed. New York, NY, USA: McGraw-Hill, 1956, pp. 165–190.

[8] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, Aug. 1969.

[9] S. L. Bressler and A. K. Seth, "Wiener–Granger causality: A well established methodology," *NeuroImage*, vol. 58, no. 2, pp. 323–329, 2011.

[10] P. M. Senge, *The Fifth Discipline: The Art & Practice of the Learning Organization*. New York, NY, USA: Knopf Doubleday, 1990, pp. 58–60.

[11] J. Forrester, *Industrial Dynamics*. Cambridge, U.K.: MIT Press, 1961.

[12] J. Forrester, *Industrial Dynamics*. Cambridge, MA, USA: MIT Press, 1969.

[13] J. Forrester, *World Dynamics*. Cambridge, MA, USA: Wright-Allen Press, 1973.

[14] A. Ghosh, *Dynamic Systems for Everyone*. Cham, Switzerland: Springer, 2016, pp. 45–48.

[15] J. Sterman, *Business Dynamics*. Boston, MA, USA: Irwin/McGraw-Hill, 2000.

[16] Y. Lin and S. Liu, "A systemic analysis with data (I)," *Int. J. Gen. Syst.*, vol. 29, no. 6, pp. 989–999, 2000.

[17] Y. Lin and S. Liu, "A systemic analysis with data (II)," *Int. J. Gen. Syst.*, vol. 29, no. 6, pp. 1001–1013, 2000.

[18] L. Csato, M. Opper, and O. Winther, "Tractable inference for probabilistic data models," *Complexity*, vol. 8, no. 4, pp. 64–68, 2003.

[19] Y. Li, H. Caö, and Y. Tan, "Novel method of identifying time series based on network graphs," *Complexity*, vol. 17, no. 1, pp. 13–34, 2011.

[20] M. Lloret-Climent and J. Nescolarde-Selva, "Data analysis using circular causality in networks," *Complexity*, vol. 19, no. 4, pp. 15–19, 2014.

[21] M. Lloret-Climent, R. Amorós-Jiménez, L. González-Franco, and J.-A. Nescolarde-Selva, "Coverage and invariance for the biological control of pests in mediterranean greenhouses," *Ecol. Model.*, vol. 292, pp. 37–44, Nov. 2014.

[22] M. Lloret-Climent and J. Nescolarde-Selva, "A sociocybernetics data analysis using causality in tourism networks," in *Proc. Int. Conf. Complex Syst. Int. J. Social, Manage., Econ. Bus. Eng.*, At. Amsterdarm, Holland, 2014, vol. 8, no. 8, pp. 2285–2291.

[23] K. Alonso-Stenberg, M. Lloret-Climent, and J. A. Nescolarde-Selva, "Causal analysis of databases concerning electromagnetism and health," *Systems*, vol. 4, no. 4, p. 39, 2016.

[24] E. Akkaya, A. Yildirim, and A. Hacinliyan, "Chaos in digital currency markets," in *Proc. Chaotic Modeling Simulation (CMSIM)*, vol. 1, 2015, pp. 65–71.

[25] V. Rusyn and O. Savko, "Modeling of chaotic behavior in the economic model," in *Proc. Chaotic Model. Simul. (CMSIM)*, vol. 3, 2016, pp. 291–298.

[26] J. M. Haley, "Forecasting chaotic business cycles perturbed by noise," in *Proc. 10th Chaotic Model. Simul. Int. Conf.*, Barcelona, Spain, May 30,/Jun. 2, 2017.

[27] K. P. Harikrishnan, R. Jacob, R. Misra, and G. Ambika, "Weighted recurrence networks from chaotic time series," *Chaotic Model. Simul.*, vol. 4, pp. 433–440, 2017.

[28] A. S. G. Muñiz, "Input–output research in structural equivalence: Extracting paths and similarities," *Econ. Model.*, vol. 31, pp. 796–803, Mar. 2013.

[29] P.-A. Balland, J. A. Belso-Martínez, and A. Morrison, "The dynamics of technical and business knowledge networks in industrial clusters: Embeddedness, status, or proximity?" *Econ. Geograp.*, vol. 92, no. 1, pp. 35–60, 2015.

[30] H.-D. Köhler, "Industrial relations in Spain—Strong conflicts, weak actors and fragmented institutions," *Employee Relations*, vol. 40, no. 4, pp. 725–743, 2018.

[31] X. Chen, J. Wang, and J. Zhou, "Process monitoring based on multivariate causality analysis and probability inference," *IEEE Access*, vol. 6, pp. 6360–6369, 2018.

[32] M. Lloret-Climent, J. Nescolarde-Selva, M. Selva-Barthelemy, and K. Alonso-Stenberg. (2018). *Smarta*. Universidad de Alicante. [Online]. Available: http://hdl.handle.net/10045/82028

[33] M. Lloret-Climent, P. E. Esteve-Calvo, and E. Almenara-Carmona, "Algorithm for qualitative analysis of a system's dependencies," *Kybernetes*, vol. 38, no. 9, pp. 1556–1565, 2009.

[34] (2018). *INEbase / Industria, Energía y Construcción /Industria y Energía /Estadística Estructural de Empresas: Sector Industrial / Resultados, Ine.es*. Accessed: Jul. 20, 2018. [Online]. Available: http://www.ine.es/dyngs/INEbase/es/operacion. htm?c=Estadistica_C&cid=1254736143952&menu=resultados&idp=1254735576715

[35] Instituto Nacional de Estadística. (2015). *Encuesta Industrial de Empresas: Año*." Accessed: Jul. 20, 2018. [Online]. Available: http://www.ine.es/prensa/np952.pdf

MIGUEL LLORET-CLIMENT received the Ph.D. degree in mathematics from the University of Alicante, Spain, where he is currently a Professor of applied mathematics. He has authored/co-authored several papers in journals and proceedings. He is devoted to general research on the theory of general systems and biocybernetics.

JOSUÉ-ANTONIO NESCOLARDE-SELVA received the degree in mathematics from the University of Havana, Cuba, in 1999, and the Ph.D. degree in mathematics from the University of Alicante, Spain, in 2010, where he has been with the Department of Applied Mathematics, since 2002. He is the author and co-author of several papers in journals and books. He is devoted to research on the theory of systems and belief systems. He won the award of Gold Title at the University of Havana, in 1999.

KRISTIAN ALONSO-STENBERG received the B.Sc. degree in telecommunications engineering (sound and image) and the M.Sc. degree in mathematics teaching (Hons.) from the University of Alicante, Spain, in 2014 and 2015, respectively, where he is currently pursuing the Ph.D. degree. He is a member of the research group Systemic, Cybernetics and Optimization, University of Alicante. He is devoted to research in general systems theory and systems modelling.

MARTHA-ARELIS SELVA-BARTHELEMY received the degree in systems engineering, and the master's degree in engineering sciences from the Moscow Energy Institute, Russia, in 1976. She has extensive experience in the development of computer applications, with various applications in the field of engineering. She is devoted to research in general systems theory and software engineering.