A new fuzzy linguistic approach to qualitative cross impact analysis

Pablo J. Villacorta\textsuperscript{a}, Antonio D. Masegosa\textsuperscript{a}, Dagoberto Castellanos\textsuperscript{b}, Maria T. Lamata\textsuperscript{a}.

\textsuperscript{a}Models of Decision and Optimization Research Group,
CITIC-UGR, Department of Computer Science and Artificial Intelligence,
University of Granada, 18071 Granada, Spain

\textsuperscript{b}Intelligent Computation Group,
Department of Statistics, OR and Computation,
University of La Laguna, 38271 La Laguna, Spain

Abstract

Scenario Planning helps explore how the possible futures may look like and establishing plans to deal with them, something essential for any company, institution or country that wants to be competitive in this globalized world. In this context, Cross Impact Analysis is one of the most used methods to study the possible futures or scenarios by identifying the system’s variables and the role they play in it. In this paper, we focus on the method called MICMAC (Impact Matrix Cross-Reference Multiplication Applied to a Classification), for which we propose a new version based on Computing with Words techniques and fuzzy sets, namely Fuzzy Linguistic MICMAC (FLMICMAC). The new method allows linguistic assessment of the mutual influence between variables, captures and handles the vagueness of these assessments, expresses the results linguistically, provides information in absolute terms and incorporates two new ways to visualize the results. Our proposal has been applied to a real case study and the results have been compared to the original MICMAC, showing the superiority of FLMICMAC as it gives more robust, accurate, complete and easier to interpret information, which can be very useful for a better understanding of the system.

Keywords: Scenario Planning, Cross Impact Analysis, MICMAC, Soft Computing, Computing with Words, Fuzzy Sets, Linguistic Labels

1. Introduction

How will the future look like? How to be prepared for possible future situations? Finding good answers to these kind of questions is essential for any company, institution or country that wants to be competitive in the current globalized world, as evidenced by the large number of foresight agencies, organizations and departments operating around the world\textsuperscript{1}. Events such as climate change, global warming, the current financial crisis or the emerging economies of countries as China, Brazil or Russia give even more importance to these studies.

\textsuperscript{1}http://www.globalforesight.org/foresight-organizations/orgs—top-foresight

Preprint submitted to Elsevier June 2, 2014
Many methods to address the above questions can be found in the literature. One of the most employed approaches is Scenario Planning \([1, 2, 3, 4, 5]\). It is widely understood as a family of methods to help managers imagine possible futures by stimulating creative thinking in order to consider a wide variety of scenarios in a systematized way \([5]\). These scenarios are coherent descriptions of alternative hypothetical futures that reflect different perspectives on past, present and future developments, which can serve as a basis for action \([6]\).

Broadly speaking, the diversity of techniques employed in Scenario Planning can be classified into qualitative and quantitative, based on the nature of the procedures they employ \([5]\). Among the most popular quantitative methods \([5]\), we find Interactive Cross Impact Simulation \([7]\), Interactive Future Simulation \([8]\), Fuzzy Cognitive Maps (FCM) \([9]\), Trend Impact Analysis (TIA) \([10]\) and Cross Impact Analysis (CIA) \([11]\). Both TIA and CIA use probabilities given by human experts. They represent probabilities of deviation from a model that has been fitted to historical data, in the former, and prior conditional probabilities of the events involved \([12]\), in the latter (in its original conception). Finally, FCMs capture causal relationships in a weighted directed graph that enables the study of loops and indirect relations. The use of FCMs in developing scenarios is very recent \([9]\).

In this work we focus on Cross Impact Analysis (CIA) \([11]\). In the initial proposal by Gordon \([13]\), experts are asked about probabilities (conditional or marginal) of the factors that constitute a scenario, that are later operated to obtain the probability of the scenario. However, a lot of different variants of CIA have been proposed since then in the literature, some of which do not necessarily make use of probabilities. Roughly, they can be classified in four groups \([14]\): deterministic \([15, 16]\), probabilistic \([13, 17]\), equation-based \([18]\), and fuzzy \([14, 19, 20]\). One well-known variant of CIA method proposed by Duperrin and Godet is MICMAC \([4, 15]\). Godet suggests applying MICMAC at the first stage of the Scenario Planning process, when the experts define the main variables of the system and their interactions in order to identify the role they play in the system. MICMAC analyses the importance of a given set of variables through a matrix that contains the influence that each variable has on the others. The influence is not expressed with probabilities but using integer values between 0 and 3. The main characteristic feature of this procedure lies on its ability to uncover both global direct and indirect influence/dependence among variables.

MICMAC has been successfully applied in many fields. Some recent examples are listed next. In \([21]\), the authors employ this method to categorize the drivers and barriers of mobile banking (also known as M-Banking) in India. In \([22]\), MICMAC was used to identify factors that represent a major threat for the dynamic loosening, under soft foot conditions, of a bolted joint. Another interesting application is found in \([23]\), where it was used to assess the driving power and dependence of supply chain risks. In \([24]\), MICMAC was employed to analyze the effect and dependence among the overall design components, and to consider the relationship network graph of distribution of components in the system. A recent study by Guo et al. \([25]\) employed a four-stage novel approach for analysing and developing a structured hierarchy framework for students’ usage of computer-mediated communication media in learning contexts. In this work, the authors used MICMAC to analyse the driver and dependence power for each media use reason and identify the hidden and indirect relationships among all reasons. Diabat and Govindan \([26]\) presented a model of green supply chain management. The model developed was validated on a case study by applying MICMAC.

MICMAC and some other CIA methods, despite being very successful tools, still exhibit some drawbacks. First, the information is given by various experts through opinion pools, panels, etc. Such information is vague due to the subjective character of the data, imprecision on
the opinions, not enough consensus among the experts, etc. This vagueness in the information is not properly addressed by these methods since they model and aggregate experts’ opinions using integer numbers, which cannot cope with uncertainty. Another drawback caused by the same fact is the low interpretability of the results, which are numerical values that lack a deep meaningfulness. Furthermore, much of the information provided by MICMAC is relative. For example, consider two systems with the same influence interrelation network among variables (Fig. 1(a)). In one of them, all interrelations are very weak (Fig. 1 (b)) and in the other, they are very strong (Fig. 1 (c)). MICMAC gives the same output in both cases, since it only provides relative information, expressed as rankings. This issue also arises when displaying the results: the same plot is obtained for the two situations. For a good understanding of the system, it is important to provide the user with both relative and absolute measures of the importance of the variables.

The drawbacks mentioned above might lead to a distorted view of the system and therefore, of future scenarios. In order to overcome these problems and obtain a more robust approach, we introduce Computing with Words (CW) techniques taken from the Soft Computing field. To be precise, the use of linguistic labels and fuzzy sets to represent the information allows establishing linguistic assessments (e.g. high, medium, low) of mutual influence relations and, at the same time, captures the uncertainty behind such judgments in a satisfactory way. Moreover, systems designed following the CW paradigm both give and receive linguistic data, which makes the information closer to human language and thus, easier to interpret. This approach has been recently applied with success in other fields as selection of personnel [27, 28], risk assessment [29], product design and development [30, 31, 32], quality service evaluation [33, 34] or renewable energies [35, 36].

Based on this idea, we aim at developing an improved version of MICMAC that overcomes these problems, as well as presenting novel ways to visualize the linguistic results. This new method is called Fuzzy Linguistic MICMAC (FLMICMAC) and has shown good results in preliminary studies with an earlier version [37]. The particular objectives we pursue with this method are the next:

- Allow the user to input data in linguistic format.
- Address the inherent vagueness present in foresight and forecast studies.
- Show linguistic output information to the user at different granularity levels.
- Give the decision-maker information in both relative and absolute terms to check the role of the variables and the real strength of their impact in the system.
• Provide new graphical representations to achieve a better understanding of the system.

The contribution is structured as follows. We start describing the MICMAC methodology in Section 2. Then, our proposal is presented in Section 3. Section 4 deals with the application of our method to a real case-study, and the comparison of the results with those obtained by the original method. Finally, Section 5 is devoted to conclusions and further work.

2. MICMAC

2.1. Description of the methodology

The MICMAC method for structural analysis is aimed at determining the most important variables within a system among a set of variables, initially specified by an expert committee, and establishing their role in the system. This is accomplished by studying the influence relation among the variables. Basically, MICMAC consists in the following three steps:

1. Define relevant variables. The variables of the system are defined basing on the opinion of several experts, brainstorming and literature review. An unsorted list of variables is given as an output in this phase. Let \( n \) be the number of variables identified.

2. Specify the relations between the variables. The experts provide a \( n \times n \) integer matrix that states the influence that each variable has over the rest of variables of the system. This matrix is called the Matrix of Direct Influence, MDI. Every cell \( \text{MDI}_{ij} \) denotes to what extent variable \( i \) influences variable \( j \). This value can be 0 if variable \( i \) exerts no influence on variable \( j \), 1 if there is low influence, 2 if there is medium influence, and 3 if there is a strong influence. The cells \( \text{MDI}_{ii} \) of the diagonal are all set to 0. According to Godet, in real systems only about 30% of the cells of the MDI matrix have values different from 0.

3. Identify the key variables. This is the main step of the method. Some important measures that give us a clue of the degree of importance of the variables can be computed from the MDI after simple operations. There are two procedures to accomplish this:

   a) Direct method. It ranks the variables according to their direct influence/dependence on/of the rest of variables. To this end, MICMAC adds the elements of the \( k \)-th row and the \( k \)-th column of the MDI matrix to obtain the global direct influence, \( I_k = \sum_{j=1}^{n} \text{MDI}(k, j) \), and dependence, \( D_k = \sum_{i=1}^{n} \text{MDI}(i, k) \), of the \( k \)-th variable, respectively.

   With this information, an influence ranking \( \sigma_I \) and a dependence ranking \( \sigma_D \) are built by sorting the variables decreasingly according to their influence and dependence, respectively. Both rankings serve as a first indicator of the importance of each variable in the system.

   b) Indirect method. It ranks the variables according to their indirect influence and dependence, which are caused by the propagation of the influences/dependences through some other intermediate variables. This may bring out some variables that, despite not having very high direct influence or dependence values, are relevant when considering the system as a whole. Such measures are calculated as follows:

   1. Initialization. Let \( \sigma_I^0 \) and \( \sigma_D^0 \) be the influence and dependence rankings obtained with the original MDI in the direct method.

   2. \( i \leftarrow 2 \)

   3. Iteration:
2.2. Establishing the role of the variables in the system

The direct and indirect influence/dependence of the variables can be used to study their role in the system. To this end, MICMAC (and CIA methods in general) establishes a bi-dimensional categorization of the variables by means of the Influence-Dependence plane [15], whose representation and categories can be seen in Figure 2. The chart is a two-dimensional map where the horizontal axis represents the degree of dependence and the vertical axis the degree of influence. The axes dividing the four quadrants are located at mean global influence and dependence, respectively. The zone where a variable is located gives information to the decision-maker about the role the variable plays within the system. Refer to [15] for details.

Finally, recall that the output of MICMAC consists of four rankings (direct and indirect influence/dependence rankings) and two influence/dependence charts (i.e., direct and direct). The rankings only show relative information, that is, they inform us about how influential/dependent a variable is with respect to the others, but there is no measure that establishes if the influence/dependence of a variable is strong or weak in absolute terms, as we showed in the introduction. The charts pose the same problem. The position of a variable in the chart depends on its distance to the average global influence and dependence, which are represented in the horizontal and vertical lines in the middle of the chart of Fig. 2. Coming back to the example of Fig. 1, MICMAC will display the same influence/dependence charts for both cases (b) and (c). Furthermore, although the input data have a very basic linguistic structure, the global direct/indirect influence and dependence are numbers whose interpretability is low or none. Both aspects are addressed by our proposal that will be described in detail in the next section.
3. Fuzzy Linguistic MICMAC

The main idea behind FLMICMAC is to enable the user to give qualitative values and to obtain linguistic outputs whose information comes in absolute terms. With this goal in mind, we develop our novel approach using CW techniques and fuzzy sets.

The use of fuzzy sets and linguistic variables is not new in CIA and MICMAC. For example, in [19], the authors presented a fuzzy cross-impact simulation for non-cyclic technology impact assessment where the interrelations among variables were modeled as linguistic values and the time as a fuzzy number. Triangular fuzzy numbers and linguistic modifiers are employed in [20] to model the experts opinion in the Kane’s simulation through impact matrix, a type of qualitative cross-impact analysis. A proposal closer to our method is found in [14]. In this work, the authors show two extensions to MICMAC, one where the strength of the influence among variables are given by linguistic labels and another one by fuzzy numbers. In the three works we have just mentioned, the use of linguistic labels and fuzzy numbers is oriented to model the uncertainty of the expert’s opinion or to prevent the experts giving accurate values of the impact relationships. However, none of them provides linguistic data to decision-makers or information in absolute terms to check the real strength of the variables’ impact. These are the two major contributions of our proposal.

Definitions and notation. Before defining our method, we will give some basic definitions and the notation used. By a linguistic variable [38] we mean a variable whose values are words or sentences in a natural language. A strict ordering must exist over the possible values of X so that all the values are comparable. As mentioned above, it is also necessary to have a mathematical structure behind such linguistic labels to enable calculations. Every linguistic term has an underlying fuzzy set [39] associated to it. Here, we will focus on triangular fuzzy numbers. A triangular fuzzy number (TFN) is a fuzzy number \( \tilde{A} \) whose membership function is defined by three real numbers \( a, b, c \), where \( a < b < c \). Thus a TFN can be represented as a triplet \([a; b; c]\)

\[
\mu_{\tilde{A}}(x) = \begin{cases} 
  \frac{(x - a)}{(b - a)} & a \leq x \leq b \\
  \frac{(c - x)}{(c - b)} & b \leq x \leq c \\
  0 & \text{otherwise}
\end{cases}
\] (1)

In our computational model, the experts use linguistic labels to evaluate the influence between the variables, and all the computations of the method are done with the membership functions of the underlying TFNs [39].

With regard to the fuzzy numbers, we will show only the mathematical operations that will be used throughout the development of the methodology. Let \( T_1 \) and \( T_2 \) be two TFNs defined by the triplets \([a_1, b_1, c_1]\) and \([a_2, b_2, c_2]\), respectively. Then, we can define mathematical operations between them such as:

Addition: \( T_1 \oplus T_2 = [a_1 + a_2, b_1 + b_2, c_1 + c_2] \) \hspace{1cm} (2)

Multiplication: \( T_1 \odot T_2 = [a_1 \times a_2, b_1 \times b_2, c_1 \times c_2] \) \hspace{1cm} (3)

Power: \( T_1^n = [a_1^n, b_1^n, c_1^n] \) \hspace{1cm} (4)

Distance between TFN’s [41]: \( d(T_1, T_2) = |a_1 - a_2| + 4|b_1 - b_2| + |c_1 - c_2|)/6 \) \hspace{1cm} (5)

Defuzzification method [41]: \( c(T_1) = (a_1 + 4b_1 + c_1)/6 \) \hspace{1cm} (6)
3.1. Description of the methodology

Our method can be divided in four main steps that are described below. Each step is applied in a different manner depending on whether we are running the direct or the indirect method; that is why we make a distinction when necessary. To make the understanding of the methodology easier, every step is illustrated in the Supplementary Material using the toy example of Table 1.

3.1.1. Definition of a set of linguistic labels for the input

We will use the set $L = \{\text{Very weak, Moderate, Very strong}\}$. We will abide to this division during the remainder of this paper but any other set of labels can be valid as well. This division matches the original number of possible (crisp) values proposed by Godet. We will refer to them as $L = \{l_1, \ldots , l_N\}$, assuming that the label None is not included in this notation because it represents no influence at all, in a crisp sense. Therefore, there are $N+1$ terms actually. $N$ must be an odd natural number (usually 3 or 5) so there always exists a central label dividing the set in two subsets of labels with the same cardinality. This is the common assumption when choosing a set of labels in CW. This way the original MDI becomes a linguistic MDI that we call LMDI, i.e. a matrix in which every cell is a linguistic label with a TFN associated to it, as shown in Table 1. The cells that are set to None are ignored (discarded) during all the computations because None is equivalent to an empty cell.

It is important to note that the parameters of the underlying TFN’s may be either predefined or customized by the user. The width of the TFN corresponding to a certain label models the amount of uncertainty assumed by the experts when they use that label to evaluate a mutual influence relation in the LMDI. Thus, the user can adjust the values of a, b and c of the TFN to her/his convenience, in order to properly capture the vagueness he/she considers behind each linguistic judgment. This issue is extremely important as it affects the rest of the calculation process, but usually there is no standard way to accomplish this task as it heavily depends on the circumstances of the problem and the final decision-maker.

3.1.2. Computation of fuzzy direct and indirect global dependence and influence

Both the direct and the indirect method remain unchanged from a high-level perspective. We must only redefine the original crisp sums and products as sums and products of TFN’s, as shown in Eq. (2) and (3). The product of fuzzy matrices can now be defined in terms of sums and products of the TFN’s of the cells. Thus the direct/indirect influence and dependence of a variable now turn into TFN’s. The formal description of this step is given below:

**a) Fuzzy direct method:**

1. For each variable $V_k$, $k = 1, \ldots , n$ compute the following TFNs as indicated in Eq. (2):

$$I_k \leftarrow \oplus_{j=1}^{n} LMDI(k, j); \quad D_k \leftarrow \oplus_{j=1}^{n} LMDI(j, k)$$
2. Build the influence and dependence rankings, \( \sigma_I \) and \( \sigma_D \), by sorting the variables separately according to their defuzzified influence \( c(I_k) \) and dependence \( c(D_k) \) that are calculated using Eq.(6).

b) Fuzzy indirect method:

1. \( \mathrm{LMII} \leftarrow \mathrm{LMDI}^\beta \), where \( \beta \leftarrow \min \{ \alpha \in \mathbb{N} : \text{rankings over LMDI}\alpha \text{ match those over LMDI}\alpha+1 \} \)

2. Compute the global fuzzy influence and dependence over \( \mathrm{LMII} \) of every variable and normalize each component of the resulting TFN’s by the next formula:

\[
\text{Norm}(T = [a, b, c]) = \left[ \sqrt[n]{\frac{a}{n}}, \sqrt[n]{\frac{b}{n}}, \sqrt[n]{\frac{c}{n}} \right]
\] (7)

3. Build indirect influence and dependence rankings, \( \sigma'_I \) and \( \sigma'_D \), by defuzzifying the TFN’s using Eq.(6).

The interested reader can refer to Table 5 of the Supplementary Material that summarizes the results of this step using the toy example. With the purpose of illustrating one important aspect of our method, a part of such results is shown in Table 2. Focusing on variables \( V_1 \), \( V_3 \) and \( V_4 \), the original MICMAC with four different influence degrees (0, 1, 2, 3), as in our example, would yield the same dependence value for these three variables, as it occurs with the central values of the TFN’s of our method. However, the uncertainty behind the global dependence of the three variables is different. Looking at the \( D_k \) values, \( V_4 \) has the lowest uncertainty (smallest difference between \( a \) and \( c \) values) followed by \( V_3 \) and \( V_1 \). The reason for the lower vagueness of the dependence of \( V_4 \) is that the extreme labels VS and VW involve less uncertainty than W, M or S. However, the different fuzziness of \( V_3 \) and \( V_1 \) is due to the fact that the accumulation of imprecision when we aggregate three subjective opinions (\( V_1 \)) is greater than when we aggregate only two (\( V_3 \)). In this way, unlike the original MICMAC, our method is able to show the uncertainty behind the global influence/dependence of a variable to the decision-makers.

3.1.3. Computation of the output linguistic term sets for the direct/indirect influence/dependence

Since the input is given in a linguistic manner, the output direct/indirect influence/dependence, which are TFN’s, should also be given in natural language. It would be desirable to know if the resulting direct/indirect influence/dependence of a variable is Strong, Moderate or Weak because it is more informative than just having the resulting TFN \([a, b, c]\) with a ranking that only provides

<table>
<thead>
<tr>
<th>( V_1 )</th>
<th>( V_2 )</th>
<th>( V_3 )</th>
<th>( V_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D_k )</td>
<td>[2, 5, 8]</td>
<td>[0, 0, 0.5]</td>
<td>[3, 5, 7]</td>
</tr>
<tr>
<td>( c(D_k) )</td>
<td>5</td>
<td>0.083</td>
<td>5</td>
</tr>
<tr>
<td>Dep. rank</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Dep. label</td>
<td>M</td>
<td>VW</td>
<td>M</td>
</tr>
</tbody>
</table>

Table 2: Fuzzy direct dependence TFN’s and labels assigned to them on the toy example.
relative information. The overall aim, therefore, is to assign a linguistic label to each variable’s fuzzy influence/dependence calculated in the precedent step\(^2\).

In order to provide information in absolute terms using the output labels, the new output term sets should be defined from the ideal case in which all the non-null relations have the same value. That is, the output label Very strong is defined from the ideal case in which all the non-null relations of our system are Very strong, and analogously for Moderate and Very weak. This idea has a resemblance with the concept of ideal solution of some linguistic multicriteria decision making models, which is also obtained after substitution of linguistic labels [43].

The output granularity \(N\) (i.e. the number of different output linguistic terms) can be adjusted by the user. This value can be different from the number \(N\) of input labels, but it must be odd as well. Possible values for \(N\) are 3, 5 or 9. Four different output term sets are calculated, namely two for the direct method (one for direct influence and another for direct dependence) and two for the indirect method. The requested granularity is created recursively by appropriately dividing the output universe of discourse (Fig. 3). Note that a fine-grain division is more accurate in describing the influence/dependence, but having too many labels may be problematic as it is difficult to understand the differences between them, even in natural language. The user should establish a trade-off to his convenience when choosing the value of \(N\).

The detailed calculation of the output labels for both the direct and indirect method is given next. To better explain this step, it has been structured in three stages. Again, the interested reader may refer to Fig. 7 of the Supplementary Material for details of the results of the fuzzy direct method on the toy example.

\[\text{Stage 1. Definition of the ideal LMDI's}\]

1. \(L_s, l_m\) and \(l_g\) be the smallest, the middle and the greatest labels from the set of input labels \(L = \{l_s, ..., l_m, ..., l_g\}\) \(^3\).
2. For each \(l_i\) in \(\{l_s, l_m, l_g\}\)
   - \(\text{a) Let } M_i \leftarrow \text{Substitute the LMDI cells different from None by } l_i,\)
   - \(\text{b) For each variable } V_k, k = 1, ..., n \text{ compute the following TFN's:}\)
     \[I^M_{k} \leftarrow @_{j=1}^{n} M_j(k, j) \quad D^M_{k} \leftarrow @_{j=1}^{n} M_j(j, k)\]

\[\text{Stage 2. Compute the left (L), central (C) and right (R) values of the influence and dependence output term set universes.}\]

<table>
<thead>
<tr>
<th>Influence term set universe</th>
<th>Dependence term set universe</th>
</tr>
</thead>
<tbody>
<tr>
<td>(L_{inf} \leftarrow \min_k [a_k : I^M_{k} = [a_k, b_k, c_k]])</td>
<td>(L_{dep} \leftarrow \min_k [a_k : D^M_{k} = [a_k, b_k, c_k]])</td>
</tr>
<tr>
<td>(C_{inf} \leftarrow \text{median}<em>{k} [b_k : I^M</em>{k} = [a_k, b_k, c_k]])</td>
<td>(C_{dep} \leftarrow \text{median}<em>{k} [b_k : D^M</em>{k} = [a_k, b_k, c_k]])</td>
</tr>
<tr>
<td>(R_{inf} \leftarrow \max_k [c_k : I^M_{k} = [a_k, b_k, c_k]])</td>
<td>(R_{dep} \leftarrow \max_k [c_k : D^M_{k} = [a_k, b_k, c_k]])</td>
</tr>
</tbody>
</table>

---

\(^2\text{Selecting the label that best fits a fuzzy number obtained after computations is known as the retranslation problem [42]}\)

\(^3\text{If } N = 3 \text{ then } l_s = l_t, l_m = l_2 \text{ and } l_g = l_3. \text{ If } N = 5 \text{ then } l_s = l_1, l_m = l_3 \text{ and } l_g = l_5\)
Figure 3: Different granularity for the direct influence output labels.
Stage 3. Definition of the output labels according to granularity.

- If \( N = 3 \), then the sets of output linguistic labels are defined as \( \Delta_I = \{ \delta^I_1, \delta^I_2, \delta^I_3 \} \) and \( \Delta_D = \{ \delta^D_1, \delta^D_2, \delta^D_3 \} \) (Fig. 3(a)) where:
  1. Direct influence labels:
     \[ \delta^I_1 \equiv \text{Very weak} = [L_{\inf}, L_{\inf}, C_{\inf}] \]
     \[ \delta^I_2 \equiv \text{Moderate} = [L_{\inf}, C_{\inf}, R_{\inf}] \]
     \[ \delta^I_3 \equiv \text{Very strong} = [C_{\inf}, R_{\inf}, R_{\inf}] \]
  2. Direct dependence labels: analogous, using \( L_{\dep}, C_{\dep}, R_{\dep} \).

- If \( N = 5 \) then \( \Delta_I = \{ \delta^I_1, \ldots, \delta^I_5 \} \) and \( \Delta_D = \{ \delta^D_1, \ldots, \delta^D_5 \} \) (Fig. 3(b)) where:
  1. \( C^I_{\inf} \leftarrow (L_{\inf} + C_{\inf})/2 \)
  2. Direct influence labels:
     \[ \delta^I_1 \equiv \text{Very weak} = [L_{\inf}, L_{\inf}, C^I_{\inf}] \]
     \[ \delta^I_2 \equiv \text{Weak} = [L_{\inf}, C^I_{\inf}, C^I_{\inf}] \]
     \[ \delta^I_3 \equiv \text{Moderate} = [C^I_{\inf}, C^I_{\inf}, C^I_{\inf}] \]
     \[ \delta^I_4 \equiv \text{Strong} = [C^I_{\inf}, C^I_{\inf}, R_{\inf}] \]
     \[ \delta^I_5 \equiv \text{Very strong} = [C^I_{\inf}, R_{\inf}, R_{\inf}] \]
  3. Repeat steps (a) and (b) with the dependence points \( L_{\dep}, C_{\dep}, R_{\dep} \) to obtain the five direct dependence labels.

- If \( N = 9 \) then it is still necessary to compute 4 intermediate points (Fig. 3(c)).

b) Fuzzy indirect method:

Stage 1. Definition of the ideal LMDI’s

1. \( \beta \leftarrow \min \{ \alpha \in \mathbb{N} : \text{rankings over LMDI}^{\alpha} \text{match those over LMDI}^{\alpha+1} \} \).
2. For each \( l_i \) in \{ls, lm, lg\}
   (a) Let \( M_i \leftarrow \text{Substitute the LMDI cells different from None by } l_i \).
   (b) For each variable \( V_k, k = 1, \ldots, n \) compute the following TFN’s:
      \[ I^{'M}_k \leftarrow \bigoplus_{j=1}^n M^j_i(k, j) \]
      \[ D^{'M}_k \leftarrow \bigoplus_{j=1}^n M^j_i(j, k) \]

Stage 2. Compute the left (L’), central (C’) and right (R’) values of the output term sets universe.

All \( I^{'M}_k \) and \( D^{'M}_k \) are first normalized using equation (7), and then \( L'_{\inf}, L'_{\dep}, C'_{\inf}, C'_{\dep}, R'_{\inf} \) and \( R'_{\dep} \) are calculated analogously to the direct method.

Stage 3. Definition of the output labels according to granularity.

The definition of output labels is also analogous to the direct method. In this case the resulting labels are \( \Delta'_I = \{ \delta'^I_1, \ldots, \delta'^I_K \} \) and \( \Delta'_D = \{ \delta'^D_1, \ldots, \delta'^D_K \} \).
3.1.4. Assignment of the closest label of the output term set to the fuzzy global influence/dependence

The last step is to assign to every influence/dependence TFN the closest label of the corresponding output term set, as follows:

a) Fuzzy direct method:

For each variable $V_k$, $k = 1, ..., n$ assign to the fuzzy influence/dependence TFN’s the closest label of the corresponding output term set:

$$
\delta_k^I \leftarrow \arg \min_{\delta_j \in \Delta_I} (d(I_k, \delta_j)) \text{ and assign that label to } I_k. \text{ The TFN of } I_k \text{ is not modified.}
$$

$$
\delta_k^D \leftarrow \arg \min_{\delta_j \in \Delta_D} (d(D_k, \delta_j)) \text{ and assign that label to } D_k. \text{ The TFN of } D_k \text{ is not modified.}
$$

b) Fuzzy indirect method:

The process is analogous to the former. For each variable $k$, $I_k^{IM}$ and $D_k^{IM}$ are assigned the closest output label in $\Delta'_I$ and $\Delta'_D$, respectively.

As before, the interested reader may refer to Table 5 of the Supplementary Material for the computation of the closest labels for both the fuzzy direct and indirect method (Tables 5(a) and 5(b), respectively) on the toy example.

3.2. Visualization of the linguistic results

As we explained above, the original version of MICMAC also presents drawbacks when displaying the results in the influence/dependence plane due to the relative character of this visualization mode. For this reason, along with the FLMICMAC algorithm, we also introduce two new plots to visualize the linguistic results obtained and show both relative and absolute information. The description of the plots is given next:

Heat map of linguistic results. First, in order to have a global view of the system and the strength of mutual influence/dependence interactions happening in it, we represent the number of variables to which each of the output influence/dependence labels have been assigned at the end of the FLMICMAC process (Fig. 4(a)). Each square of the map corresponds to a possible combination of linguistic influence and dependence, and the colour intensity indicates the number of variables having such combination of labels at the output. This plot can be done for both the direct and the indirect method, and displays absolute information in a way that is easily interpretable and gives an overall view of how influential/dependent the variables in our system are. In this way, the global behaviour of the whole system is summarized in a plot facilitating a better understanding. The detail level of this representation depends on the granularity $\bar{N}$ set by the user. For this sample figure, we have set $\bar{N} = 5$.

Fuzzy influence-dependence plane with absolute information. In this second graphic (Fig. 4), we represent the influence and dependence of every variable in the direct/indirect method in a bi-dimensional chart that is similar to Fig. 2 but allows visualizing information in absolute terms. Every point in the space represents a variable, defined by a pair of crisp values which stand for the defuzzified global influence and dependence of the variable. Note we have added a linguistic scale to the axes. The TFNs represented on them are the output term sets, computed as explained in Section 3.1.3. Therefore, it is possible to know the output label assigned to the fuzzy
influence and dependence of a variable, by looking in the axes of the chart for the fuzzy sets to which the point belongs most. In this way, we can categorize the role of a variable in the system by comparing its global influence/dependence respect to the other variables, as in the original plot, and at the same time, we can check the strength of the global dependence/influence of this variable in absolute terms, which gives a more accurate idea of its behaviour in the system.

4. Application and comparison in a real case study

To assess our method, we have used a real case study about the determinant factors of the rural spaces within time horizon 2010 in France [44]. This example can be downloaded from the LIPSOR web site. The results obtained by our method will be compared against those from Godet’s original MICMAC method.

We must point out that the focus of this section is not on the economic or social interpretation of the results concerning the problem itself, but on the comparison of FLMICMAC and Godet’s MICMAC, as well as the advantages of our method regarding absolute linguistic information and novel plots to visualize the linguistic output. For that reason, no comments or explanations about the specific results of the case study will be given, as it is out of the scope of this work.

In this study the group of experts initially defined 50 variables to be considered in the analysis, which can be found in Table 6 of the Supplementary Material. In the linguistic version, \( N = 3 \) labels (with their underlying TFNs) were defined to be employed in the linguistic fuzzy MDI: \( L = \{ \text{Weak} = [1, 1, 2], \text{Moderate} = [1, 2, 3], \text{Strong} = [2, 3, 3] \} \), which is the same granularity employed in the original MICMAC method using three integer values \([1, 2, 3]\). Although it seems we are doing here a conversion between original crisp values of the example and linguistic...
labels, it must be noticed that it has been done with the only purpose of a fairer and easier comparison of the results with another existing method, namely crisp MICMAC. When FLMICMAC is to be applied to a new problem, no conversion has to be done, since the first step of the method is the definition of a set of linguistic labels that will be used to assess mutual influence of the variables and construct the LMDI. No previous numerical (crisp) evaluation of such influences has to be conducted, but directly a linguistic one.

Regarding the granularity of the outputs, we have tested the values $\bar{N} = 3$, $\bar{N} = 5$ and $\bar{N} = 9$ for the number of output labels in order to evaluate the influence of this parameter over the actual distinctness of the linguistic results. The output term sets in each case together with their abbreviations are summarized in Fig. 3, which is a generic figure. The concrete values calculated in this problem (dependence only) can be found in Fig. 8 of the Supplementary Material.

The complete rankings obtained with MICMAC and FLMICMAC are presented in the Supplementary Material (Tables 7, 8, 9 and 10). In order to facilitate the analysis of the results, we only provide a brief summary here (Table 3) that contains the variables on the first and last five positions of the Godet’s global direct and indirect influence rankings. Such information is enough for explaining the benefits of our method.

The variables have been sorted according to Godet’s ranks in all the tables. The meaning of the columns is the following:

- $\text{Var}$ is the numerical identifier of the variable.
- $R_G$ and $R_F$ are the ranks assigned to the variable by Godet’s original MICMAC method and by our FLMICMAC method, respectively.
- $\text{Inf}_G$ and $\text{Dep}_G$ are the integer (crisp) values of the influence and dependence of the variable computed according to Godet’s method, that is, the sum of its row and column, respectively. Godet’s rankings $R_G$ are obtained by sorting the variables according to these values.
- $\text{Inf}_F$ and $\text{Dep}_F$ are the real values obtained after the defuzzification of the fuzzy influence and dependence of the variable, computed according to the FLMICMAC method. The fuzzy rankings $^5 R_F$ are obtained by sorting the variables according to these values.
- $I_k$, $I'_k$, $D_k$ and $D'_k$ are the TFNs of the global direct and indirect influence and dependence of the variables, respectively.
- $L_3$, $L_5$ and $L_9$ are the linguistic labels assigned to the absolute direct/indirect influence/dependence (depending on the meaning of each table) when 3, 5 or 9 different output labels are considered, respectively, as indicated in the FLMICMAC method.

4.1. Comparison of FLMICMAC against MICMAC

A number of remarkable results have been obtained after the application of FLMICMAC to this case study, as can be seen in the tables:

---

$^5$Here we use \textit{fuzzy} to refer to the FLMICMAC method employed to obtain the ranking, not because the ranking itself is fuzzy, which is not the case.
<table>
<thead>
<tr>
<th>Var</th>
<th>$\text{Inf}_G$</th>
<th>$R_G$</th>
<th>$I_F$</th>
<th>$\text{Inf}_F$</th>
<th>$R_F$</th>
<th>$L_3$</th>
<th>$L_5$</th>
<th>$L_9$</th>
<th>Linguistic output</th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>59</td>
<td>1</td>
<td>[37, 59, 88]</td>
<td>60.17</td>
<td>1</td>
<td>M</td>
<td>S</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>49</td>
<td>2</td>
<td>[34, 49, 67]</td>
<td>49.50</td>
<td>3</td>
<td>M</td>
<td>M</td>
<td>MS</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>46</td>
<td>3</td>
<td>[32, 46, 70]</td>
<td>47.67</td>
<td>4</td>
<td>M</td>
<td>M</td>
<td>MS</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>46</td>
<td>4</td>
<td>[34, 46, 80]</td>
<td>49.67</td>
<td>2</td>
<td>M</td>
<td>M</td>
<td>MS</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>43</td>
<td>5</td>
<td>[27, 43, 58]</td>
<td>42.83</td>
<td>6</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>11</td>
<td>46</td>
<td>[8, 11, 19]</td>
<td>11.83</td>
<td>46</td>
<td>W</td>
<td>V</td>
<td>QW</td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>10</td>
<td>47</td>
<td>[8, 10, 18]</td>
<td>11.00</td>
<td>47</td>
<td>W</td>
<td>V</td>
<td>QW</td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>10</td>
<td>48</td>
<td>[6, 10, 16]</td>
<td>10.33</td>
<td>48</td>
<td>W</td>
<td>V</td>
<td>QW</td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>9</td>
<td>49</td>
<td>[7, 9, 16]</td>
<td>9.83</td>
<td>49</td>
<td>W</td>
<td>V</td>
<td>QW</td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>7</td>
<td>50</td>
<td>[5, 7, 12]</td>
<td>7.50</td>
<td>50</td>
<td>W</td>
<td>V</td>
<td>W</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Top and last five positions of the global direct/indirect influence rankings, according to MICMAC and FLMICMAC.
Godet’s and FLMICMAC rankings are roughly similar but the latter are more accurate. This is the first important result. This roughly similarity confirms the correctness of our approach and the good quality of the obtained rankings. It has been formally assessed using a correlation Kendall’s tau-b test whose results are shown in Table 4. We can check that no significant differences are found between MICMAC and FLMICMAC’s rankings. The null hypothesis of rank independence is thus rejected in all cases, according to the obtained p-values. Furthermore, there is a significant positive correlation between original MICMAC method and our FLMICMAC method.

However, as we explain next, FLMICMAC provides more accurate rankings since it takes into account the underlying vagueness of the aggregated experts’ judgments. Let us focus on the five top ranked variables of Table 3(a). Looking at the central values $b_i$ of the TFNs of column $I_k$, we see that they perfectly match the crisp global influence given by Godet (column $In_{f_G}$). However, the two sides of the triangles contain additional information to measure the vagueness of the fuzzy global influence of the variable, leading to a more accurate ranking. For example, variables $V_4$ and $V_{32}$ are indistinguishable according to MICMAC. Their fuzzy influences are respectively $[32, 46, 70]$ and $[34, 46, 80]$. Hence $V_{32}$ is located to the right (is greater), i.e. in the TFNs we see that $a_{32} > a_4$ and $c_{32} > c_4$. For this reason, $V_{32}$ has to be ranked higher than $V_4$, as done by our method. Another interesting comparison arises between $V_5$ and $V_{32}$. Although the $In_{f_G}$ value of $V_5$ is greater than that of $V_4$, our method ranks $V_4$ first. This is due to the much greater right tail of $I_{32}$ which goes from 46 to 80 whereas in $I_4$ it ranges from 49 to 63. This can be understood as that $V_{32}$ has a slightly greater possibility of being more influent than $V_4$, as shown by their defuzzified values. Similar situations arise between variables $V_{25}$ and $V_{32}$, and between $V_{38}$ and $V_{48}$ for the global indirect influence.

In addition to this, we should highlight that top ranked variables pose more uncertainty than those in the last positions. If we measure the width of the base of the triangle of the TFNs, $c_k - a_k$, it can be seen that it is greater in the top ranked variables, and this happens in both the direct and the indirect method. The reason is the following. Most likely, a variable $k$ ranked among the first positions exerts influence on a lot of variables, hence its fuzzy influence $I_k$ is the result of the aggregation of a higher number of influence values than that of the variables in the low part of the ranking. Since each of these influence values (TFNs) represents an inherently vague assessment, the higher the number of influence values aggregated, the more the accumulated uncertainty, and thus the wider the TFN. Therefore, the output TFNs are able to properly capture such uncertainty, making our method especially robust for the most important variables of the system.

---

6The mutual influence assessment and its corresponding TFN may be provided either by one expert or agreed among a panel of experts. FLMICMAC is independent from how such linguistic assessment and TFN have been obtained.
Figure 5: Number of variables with each possible combination of linguistic influence and dependence for different granularity levels at the output.
Figure 6: Fuzzy influence/dependence planes showing the defuzzified influence-dependence pairs and the output linguistic term sets on the axes.
**Absolute, linguistic influence/dependence measures.** Since the linguistic information contained in columns $L_3$, $L_5$, and $L_9$ of Table 3 is absolute, it allows to know whether the top-ranked variables really have a high degree of influence or dependence. Notice that although a variable may be placed on the top of the ranking, this means only that the influence or dependence of the variable is higher than the others, but not necessarily high in absolute terms. This could not be determined in the original MICMAC method because an integer number is not enough to assess the influence/dependence in an absolute manner. Such issue is solved in FLMICMAC. In addition, linguistic values are easier to understand and interpret by humans. This does not happen when the influence/dependence are expressed only numerically.

**Overall linguistic results depiction in heat maps.** As mentioned before, in order to have a clear idea of the absolute influence/dependence of all the variables, it is possible to summarize all the linguistic information in 2D heat maps (Fig. 5) showing the number of variables with each combination of absolute linguistic influence/dependence labels. In our problem, the heat maps show the central linguistic terms (those around Moderate) appearing more often in the results, both in the direct and indirect method. The number of different labels used at the output of the indirect method is actually very small, and they are concentrated around central terms: the majority of the squares are left blank with an intensely red central area consisting of 3 or at most 4 colored squares. This is due to the fact that the indirect method tends to fade the output after long computations, leading eventually to more uniform results, which is a common drawback of many linguistic procedures. The normalization applied here by taking the $\beta$-th root tries to overcome such disadvantage.

In our case study, these plots allow the decision maker checking at a first glance that, in this system, most of the variables present a low-intermediate global influence and dependence. Moreover, in a broad sense, heat maps also serve as a way of clustering the variables, since those with the same combination of linguistic influence and dependence may play a similar role in the system.

**Linguistic influence-dependence plane combining both absolute and relative information.** As mentioned before, we propose improving Godet’s influence-dependence chart with linguistic capabilities that make it more informative and easier to interpret than a raw categorization of the variables, as happens in Fig. 2.

In this way, for the direct method, the decision maker can see in Fig. 6(e) that both the input and output variables (top left area and lower right area, respectively) have in general a Weak-Moderate degree of direct influence and dependence, whereas the stake variables (top right area) present a Moderate-Strong influence and dependence. If we measure the indirect relationships, then the input, output and stake variables all pose a Weak-Moderate influence/dependence of the system.

5. Conclusions and further work

In this work we have presented a novel methodology for structural analysis in the context of Scenario Planning. Concretely, we have developed a fuzzy linguistic version of MICMAC using CW techniques, called FLMICMAC, that maintains the essence of the original MICMAC but adds new stages to define the input and output in a linguistic way. The main novelties are the following:
• Allow for linguistic assessment of mutual influence between variables.

• Capture and handle the vagueness of the subjective experts’ opinions using linguistic labels and TFNs.

• Provide linguistic measurements of the global direct/indirect influence/dependence, leading to more interpretable results which, differently from MICMAC, express the information in absolute terms.

• Present two new techniques to visualize the results (Fig. 4): first, heat maps that summarize the linguistic information of the global influence and dependence of the variables; second, an improvement of the classic influence/dependence plane that incorporates linguistic information to the plot axes. As this information reflects the global influence/dependence in absolute terms, this new graphic shows both relative and absolute information to the user.

The new method has been compared versus the original MICMAC in a real case study [44]. The key findings of this comparison are listed below.

• The rankings obtained with MICMAC and FLMICMAC are roughly similar, which confirms that our new method obtains reasonable results.

• The use of linguistic labels with underlying TFNs successfully captures the accumulated vagueness that results from aggregating linguistic judgments when computing the global influence or dependence of a variable. This ultimately leads to more accurate and robust rankings than those calculated by MICMAC.

• Such accumulated uncertainty is greater for the most influent/dependent variables, making them specially prone to misplacements in the rankings. This has important implications, since the objective of the structural analysis is to determine the most relevant variables of the system, which are usually the top ranked variables. Relevant variables are usually the basis of posterior analyses in Scenario Planning so their correct identification is crucial. This fact gives even more value to the robustness and accurateness of the rankings computed by FLMICMAC.

• The heat maps allow the decision maker to see a snapshot of the strength, in absolute terms, of the relationships among the variables of the system. They also cluster the variables in a visual way, since variables with the same combination of influence/dependence may play a similar role in the system.

• The fuzzy linguistic influence/dependence planes, apart from categorizing the variables in input, output, excluded or stakes (recall Fig. 2), also establish the influence/dependence of the variables in each category in absolute terms, leading to a better description of the role of the variables in the system.

Although our method has been tested only on this concrete case study, in our opinion the size and complexity of the problem are enough to consider that the advantages discussed above are generalizable to other real studies.

The improvements done to MICMAC are not “for free” and entail some drawbacks that should be considered when applying our method. First of all, addressing the uncertainty and imprecision involves a higher complexity that may not worth in two cases:
• When the vagueness in the input data (relationships between variables) is low, our method would probably provide the same results as the original MICMAC.

• Understanding and taking advantage of all the information given by FLMICMAC requires basic notions of fuzzy sets which may make FLMICMAC unattractive for some decision-makers who are not familiar with these techniques.

Secondly, the results of FLMICMAC also depend on how the input labels have been defined, hence it has a higher number of parameters (extremes of the TFNs) to be set. Although the setting used in this paper can be employed in a wide range of cases, some particular problems may require a different configuration of the TFNs. If the definitions of the TFNs associated to the input labels do not capture the uncertainty of the experts opinion properly, FLMICMAC results might be biased and lead to inexact or wrong conclusions. The sensitivity to this setting is low, but it is important to ensure that the input labels are correctly defined so as to avoid these problems.

Finally, the higher conceptual complexity of FLMICMAC, which also involves a more difficult implementation, might be a drawback towards its generalized adoption in the community. With this in mind, we plan to release a web version of FLMICMAC shortly that will be freely available so other researchers and practitioners can use our method with their own data\footnote{This tool will be available at \url{http://modo.ugr.es:8080/FLMICMAC/}}. Despite the apparently high number of operations, FLMICMAC running time is negligible in practice even in a 50-variable problem like the case study presented here, hence time complexity is not a problem.

A point we have not addressed in this work is the handling of potential influence relations between variables, a feature that MICMAC does take into account. Potential relations are those that do not exist in the present, but may become effective (with any strength) in the future. Although they do not require a new methodology (MICMAC just substitutes them by a positive influence degree and then checks the variations that take place in the system), we plan to develop specific methods for determining to what extent potential relations can affect the whole system. A preliminary study in this direction can be found in [45].

Apart from this, in future works we intend to apply CW and Soft Computing techniques to other Scenario Planning methods, such as Morphological analysis, for which very promising preliminary results have been achieved recently [46]. We consider that robust management of the vagueness and uncertainty present in experts’ judgments -often used as input in foresight studies- is essential to improve existing methods in Scenario Planning and hence, Soft Computing tools can be very helpful.

6. Acknowledgments

This work has been partially funded by the projects TIN2011-27696-C02-01 from the Spanish Ministry of Economy and Competitiveness, TIC-08001 from the Andalusian Government and FEDER funds from the European Union. The first author acknowledges a FPU scholarship from the Spanish Ministry of Education.
References