

# Fuzzy Dissimilarity-Based Classification for Disaster Initial Assessment

J. Tinguaro Rodríguez<sup>1</sup> Begoña Vitoriano<sup>1</sup> Javier Montero<sup>1</sup>

<sup>1</sup>Faculty of Mathematics, Complutense University of Madrid, Spain  
{jtrodrig,bvitoriano,javier\_montero}@mat.ucm.es

## Abstract

A correct initial assessment of disaster consequences is crucial for an adequate decision-making in disaster and emergency management. However, such an initial assessment needs to be correct, but not necessarily fully precise, and thus it can be associated with a fuzzy classification problem in which the set of classes presents a relevant structure. This paper proposes the consideration of a dissimilarity operator in order to introduce such a structure in the classifier's learning and reasoning procedures, leading to an improvement in the classifiers adaptation to the disaster management context features and decision making requirements.

**Keywords:** Fuzzy rule based classification systems, disaster management, dissimilarity.

## 1. Introduction

Just after a disaster strikes somewhere in the world, international disaster relief agencies (as OCHA – UNDAC) and NGOs (as IFRC) start a decision process intended to reach a conclusion about the pertinence of a relief operation and about whether or not suitable conditions exist to initiate it. Therefore, the decisions to be made in a first stage have a strategic nature, more concerned with assessing the degree of involvement of the organization in a possible response operation than with the specific content of such an operation.

However, precisely because they determine the shape and guidelines of the actions to be done, strategic decisions have a major influence on the subsequent logistical and on terrain decision processes (see Fig. 1), which evaluate the amount of aid to send and how it will be delivered to the affected country and the suffering population (see [4]). As a consequence, strategic decision-making takes place in a highly time-pressured context, since any delay at this stage could affect the position of the organization in the international coalition delivering aid, thus affecting the organization's prestige and reputation, and slow down the subsequent decision processes, thus delaying the reception of aid by the affected population.

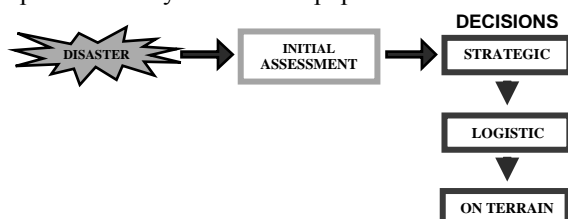


Figure 1: Hierarchy of decisions for disaster response.

Notice that this strategic decision process has to be flexible, in the sense that it has to be able to be carried out for every combination of disaster type and place, since international disaster relief agencies and NGOs are possibly specialized in covering some part of the relief tasks (as water sanitation, shelter and site management, health care, etc.) but are not specialized in response to, for instance, earthquakes in Haiti, floods in Pakistan or any other specific disaster scenario (despite the geo-strategic priorities an organization could have as a result of the interests of its donors, see [9]).

As shown in Fig. 1, strategic decision-making (and thus all the subsequent decision processes) is strongly dependent on a correct assessment of the consequences of a disaster and the resulting needs of the affected population. However, the available information just after the impact of a disaster uses to be affected by different kinds of uncertainty. The first reports are usually pretty incomplete, if not directly confusing and contradictory. Moreover, when it exists, relevant information is usually expressed linguistically, and thus it could be vague and imprecise. In fact, due to the effects an adverse phenomenon has on the informational system of a region, a more or less exhaustive and precise picture of the situation could not be obtained until some days (or even weeks) after the moment of the strike. This poses a strong difficulty in the development of the mentioned strategic decision-making process, since in this context the urgency of the decisions to be made in order to relief the people affected by a disaster clashes with the need of a correct estimation of the consequences of such a disaster, i.e. of the needs of the affected populations.

In the spirit of [17], all this complexity suggests as a promising alternative the development and application of inference techniques enabling a fast, flexible and correct assessment of disasters' consequences in the presence of uncertainty. Nevertheless, some constraints must be imposed on the nature of these techniques. For example, the procedures leading to such an assessment have to be understandable and interpretable by the decision makers, in order to guarantee the usability of such an inference tool. And, even more important, it is necessary to make realistic assumptions about the infrastructure and data requirements of a DSS to be used in organizations or countries where the operational infrastructure can not supply highly sophisticated data. It has to be remarked that this question is a critical issue when considering disaster management from the point of view of NGOs or developing countries (see [1][6]).

The long term objective of the project containing this work is to develop and implement a general, standard decision support system (DSS) for disaster

management, specifically designed to address a part of this complexity and help NGO's decision makers involved in the design of humanitarian relief operations. Particularly, this work focuses on SEDD (the Spanish acronym for Disaster Diagnostic and Evaluation System), which is the part of such a global DSS concerned with the assessment of the consequences of disasters with the very first information available after the strike (see [16] for a description of other parts of such a global DSS being developed in this project). As discussed in [13], SEDD constitutes one of the first proposals aiming to provide NGOs with disasters' consequences evaluation procedures specifically designed for them.

Let us recall with [14] that SEDD's inference capability is based on fuzzy machine learning procedures (see [5]), particularly on the methodology of descriptive fuzzy rule based classification systems (FRBCS, see [3]). This is, SEDD's methodology is a mixture between fuzzy inference systems, that enable inference to be carried out in terms of a linguistically expressed, interpretable set of rules and information, and data mining and machine learning (see for instance [5]), that allows these rules to be obtained from adequate databases. SEDD extracts its rules from EM-DAT (Emergency Database, see [www.em-dat.be](http://www.em-dat.be)), the most exhaustive and complete public database about disasters and emergencies. Moreover, as shown in [14], fuzzy rule based systems outperforms ordinary statistical techniques and most machine learning techniques in providing a simultaneously accurate and interpretable assessment of disaster consequences.

This paper studies the characteristic structure of the classification problem that underlies the disaster severity assessment provided by SEDD. Particularly, the relationships between the structure assumed on the set of classes and the requirements of the decision context are analyzed, showing that a significant improvement in the behavior of the classification methodology of SEDD in terms of its adaptation to such requirements is obtained when certain structures are introduced inside the classification models.

In this sense, as shown for instance in [11], in a supervised classification context it is possible to introduce relations between the classes by means of the notion of semantic antagonism or dissimilarity proposed in [10]. Thus, in this work the effect of different dissimilarity structures is discussed in terms of their ability to replicate and adapt to some of the requirements of the disaster management context.

## 2. Strategic disaster severity assessment as an structured classification problem

As explained above, the strategic decision-making about the involvement of an NGO on a disaster response operation is strongly dependent on the initial assessment of the consequences of such a disaster. However, a fully precise numerical evaluation of disasters' effects, as casualties, homeless people or the extension of the material damage, is unrealistic in such a decision context. In fact, even a more or less complete

and precise description of these consequences is usually not available by the time in which such (urgent) strategic decisions have to be already taken.

This is mainly due to the uncertainty and the referred features of the available information just after a disaster strike, but it is also because of the imprecise nature of some of the relevant categories. For example, the notion of *affected people* shows such an imprecision, since it could be not always clear whether a person has been affected or not. As a consequence, the number of affected people is usually stated through an implicitly imprecise quantity, as happens when it is said that a disaster produced, for instance, 40.000 affected people.

Nevertheless, it is important to notice that a totally precise and exhaustive evaluation of consequences is not actually needed in order to perform the above described strategic decision-making. As pointed out above, strategic decisions determine the shape of an operation but not its specific contents. As the decision process in Fig. 1 develops, decisions need information to be more and more precise, since decisions become more and more concrete. In this sense, NGOs usually deliver experts on the affected location in order to be able to acquire such a more precise evaluation for its logistical and on terrain decisions. But the decision of acquiring such a further evaluation is a strategic decision that has to be taken in the first moments after the strike, when little information is available, i.e. on the basis of the initial reports of the disaster. However, such an initial assessment of consequences needs to be correct or accurate, but not necessarily fully precise.

For example, consider the estimation of the variable *number of homeless people*. This variable measures the number of people that become homeless as a consequence of a disaster. For NGOs, such a quantity constitutes a key indicator of the size of the efforts a potential relief operation should place in matter of temporary shelter and site management. This also provides an idea of the efforts to be placed in the water sanitation area, for instance. In practice, at a first stage, for an NGO decision maker it is not important at all to distinguish whether 50.000 or 70.000 people became homeless as a consequence of a disaster, since anyway such a number is going to be considered *large* and the strategic decisions in terms of both the subsequent actions and the size and nature of the required relief operation are going to be similar.

Therefore, in the first moments after a disaster strike, NGO decision makers assess disaster severity (and thus also the needs of a potential intervention) in a *qualitative* (rather than *quantitative*) way. In other words, their problem consists on evaluating, in a context of highly uncertain and imprecise information, the *magnitude* of the consequences of a disaster in relation with the relevant scenarios and decisions that can arise regarding the implementation of a relief operation. In such a context, it is even possible that an assessment stated in crisp and precise terms could result little trustable to those decision makers.

However, a linguistic description of the magnitude order of the consequences coming from a reliable source, for example stating that there are *a lot of*

casualties or that buildings took a *several* damage, will be much more trustable to decision makers, despite its implicit imprecision. This kind of linguistic information is enough relevant to elaborate a first perception of the disaster scenario, providing a base for the subsequent strategic decision making.

Thus, in order to obtain such an initial assessment giving rise to an adequate strategic decision making, we consider that instead of a numerical evaluation, it is rather more plausible and realistic to *classify* the severity of the consequences of a disaster in terms of the relevant scenarios for the NGO's decision makers. Therefore, the abovementioned practical problem of evaluation of disaster consequences leads to a *classification* problem in which the *classes* are identified with the linguistic terms that describe those relevant scenarios, as *no casualties* or *a lot of injured people*. These linguistic labels or *classes*, assessing the magnitude of the different relevant consequences of a disaster, have to be assigned on the basis of the description or *attributes* of such a disaster given by the first available information, as the type of disaster, its intensity and the features of the affected location (e.g. its vulnerability, see [8]).

## 2.1. Structure of the set of classes

Consider now one of the variables that have to be linguistically evaluated in order to obtain such a first initial assessment. For instance, let us focus on the variable *number of casualties* (CAS). This variable estimates the number of people that were killed as a result of the strike of an adverse phenomenon. As just explained, in a first stage a fully precise estimation of such a number is not strictly necessary, but just a qualitative, linguistic assessment. In this way, for instance we can measure the magnitude of a disaster scenario, in terms of the casualties it produced, by means of the labels *no casualties*, *very few*, *few*, *quite a lot* and *a lot of casualties*. These labels represent the classes in which such a disaster scenario has to be classified in order to provide an initial assessment of the relevant consequence CAS.

The particular meaning of these labels has to be specified by means of intervals, or more generally, through fuzzy subsets of the range of the underlying numerical variable, in this case the positive integers. Moreover, such meanings have to be related to the different scenarios that are relevant in terms of the decisions to be made. For example, each label can be associated with a different order of magnitude of the number of casualties, in an increasing way, as shown in Table 1.

Class	Label	Interval
CAS1	No casualties	[0,10)
CAS2	Very few	[10,100)
CAS3	Few	[100,1000)
CAS4	Quite a lot	[1000,10000)
CAS5	A lot	[10.000,+∞)

**Table 1.** Intervals associated to the linguistic labels defined for the variable *number of casualties* (CAS).

Notice that, as they are associated to different orders of magnitude of the consequences of a disaster, these labels can be considered as ordered from the lowest (*no casualties*) to the greatest (*a lot of casualties*) level of magnitude of such consequences. This is, the classes associated to the variable CAS (in the first column of Table 1) are *linearly ordered*, i.e.  $CAS_i < CAS_j$  whenever  $i < j$ .

This assumption of *linearity* on the effects of a disaster is quite general, since what decision makers try to assess in a first stage is precisely the order of magnitude or severity of these consequences, or equivalently the order of magnitude of the efforts that a relief operations should consider in order to adequately alleviate such consequences. As explained above, these orders of magnitude are not necessarily powers of ten of the underlying numerical variable, but rather they are associated to the different relevant scenarios that can arise, which can be also considered as ordered attending to the *gravity* of the humanitarian crisis taking place in each of such scenarios.

Thus, in this setting classes are not independent, unrelated items, but they conform a valuation structure, in which some relationships hold between the valuation states given by the classes. This is, following [7], in this context the set of classes presents a relevant *structure*. For example, we do not commit the same error when a disaster scenario having *no casualties* (CAS1) is assessed as one with *very few casualties* (CAS2), than when it is evaluated as having produced *a lot of casualties* (CAS5). Therefore, the abovementioned problem of disaster severity assessment can be understood as a classification problem with a structured set of classes, i.e., as a *structured classification problem*.

## 2.2. Context requirements

As stated above, the notion of structure of the set of classes is introduced in order to capture some relevant relationships that hold between the concepts represented by the classes. These relationships are given by the features of each particular application context, which provides the specific meaning or semantics attributed to the classes. However, it is important to notice that these relationships not only depend on the semantics of the classes, but they also have to reflect the criteria and requirements of the decision context.

For instance, the set of classes introduced before for the variable CAS in principle fits into a linear structure, since classes are semantically associated to orders of magnitude of the consequences. As a consequence of this assumption, the classifier could be required to show a gradable, smooth behavior, in the sense that small variations on the attributes that describe a disaster scenario should not produce a large variation on the predicted consequences. Some classes are *closer* than others to a given class, and thus different error levels can be distinguished according to such a *distance*.

However, notice that the error committed by assessing a CAS1 scenario as a CAS5 one is also different from that committed when a CAS5 scenario is

evaluated as verifying *CASI*. In the first case, overestimation error is committed, while in the second case scenarios are underestimated. Though the two types of errors are relevant, notice that underestimation of disaster consequences could lead to much more dangerous situations than overestimation in terms of the prestige of an NGO and the relief of the affected population.

In this sense, overestimation of consequences may lead to an initial overreaction, but as soon as observers are deployed on terrain and further information is available the scenario can be reassessed and the decisions reconsidered without too many difficulties. However, when a disaster scenario is underestimated, it uses to attract less attention and to be considered as less important, which could lead to not properly ask for further information or even to ignore it in a first moment, thus potentially affecting the timing of the strategic decision-making stage (with the resultant delays on the subsequent logistical and operational decision phases) as well as the NGO's reputation.

Consequently, NGO decision makers usually tend to avoid the risk of underestimation of disaster effects, for instance by carrying out a worst-case analysis of the scenarios under study. In this way, initial assessments of a disaster scenario could be required to be developed under the assumption of avoiding underestimation risk. In this sense, such a decision-related requirement entails introducing a somehow asymmetric configuration in the linear structure of the classes, since different error levels are then attained depending on whether a disaster scenario under study is underestimated or overestimated. Therefore, as pointed out above, the structure of the set of classes has to capture both the relevant aspects of the semantics of the classes as well as the objectives and requirements related with the decision context in which the classifier is used.

### 2.3. Dissimilarity structures

Notice that the assumption of asymmetry on the linear structure of the set of classes forces to look for more general structures than orders. In this work, we adopt the notion of *dissimilarity structure* proposed in [10] to provide a formal definition of the structure of the set of classes in an structured classification problem. Recall that dissimilarity structures are based on the notion of semantic antagonism (also proposed in [10]), that provide a formal framework to model the opposition relationships between a set of concepts in which such an opposition is allowed to be asymmetric. Therefore, by adopting the notion of dissimilarity or antagonism instead of that of linear order, we somehow translate the semantic *distance* between two classes, coming from the linear ordering of the consequences, into the *degree of opposition* among them, which is however allowed to be asymmetric in order to reflect the requirement of underestimation risk avoidance.

Moreover, as we shall see in next section, dissimilarity structures provide an easy and effective method of introducing the relationships between the classes into the classification models, i.e. in the learning

and reasoning processes of the classifiers. In this context, the opposition between classes represented in the dissimilarity structure enable to distinguish *significant exceptions* to a classification rule from simple, logical counterexamples, which leads to introduce a *negative* confidence degree of the rules. Such a negative degree together with the usual, positive confidence degree, constitute then a bipolar evidence pair for the evaluation of classification rules, which leads to a bipolar fuzzy rule-based classification framework (see [11] for further details).

Therefore, let us denote by  $\zeta = \{C_1, \dots, C_N\}$  the set of concepts or classes into consideration. Recall that a dissimilarity structure can be built upon this set by means of a *dissimilarity matrix*  $\Delta = (d_{ij})_{N \times N}$ , such that the value  $d_{ij} \in [0,1]$  expresses the degree up to which the class  $C_j$  is opposite, antagonistic or dissimilar to the class  $C_i$ ,  $i, j = 1, \dots, N_C$ . Notice that  $\Delta$  is allowed to be non-symmetric, thus enabling the underlying notion of dissimilarity to be asymmetric.

This way, for example, a matrix  $\Delta' = 0$  represents a situation in which no class is opposite to any other. On the other side, a matrix  $\Delta'' = 1 - Id$  describes a situation in which every class is totally opposite to each other. As we shall see in the next section, the consideration of a dissimilarity matrix allows to introduce and take into account some of the requirements and constraints of the application context *inside of* the classification model. Therefore, there exist a wide range of possibilities, lying between these extreme cases, in order to model specific dissimilarity conditions. The choice of a particular dissimilarity matrix will of course depend on the specific semantic requirements to be fulfilled.

## 3. SEDD's bipolar classification methodology

Recall that, as it was described in [14][12], SEDD's fuzzy rule-based classification methodology does not explicitly consider any structure on the set of classes. Therefore, in this section we illustrate the ideas above by adapting the methodology of SEDD to an structured framework. To this end, we adopt the bipolar fuzzy rule-based classification framework proposed in [11]. Thus, here we study the ability of different dissimilarity structures (i.e. of different matrices  $\Delta$ ) to capture the semantics and the requirements of the disaster response NGO strategic decision context as well as to produce a correct classification result, i.e. a correct assessment of disaster consequences, through the bipolar classifiers proposed in [11].

### 3.1. Some basics about SEDD

For the sake of an adequate understanding of the example we will propose in next section in order to illustrate the feasibility of the proposed approach, let us first recall some basics about SEDD.

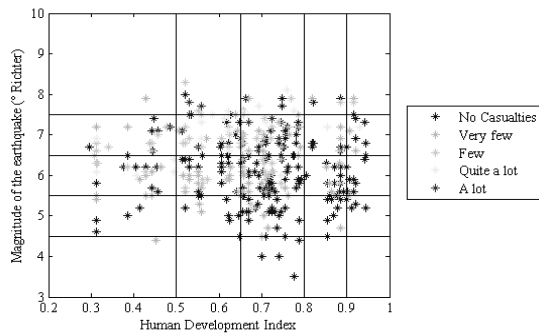
Firstly, recall that SEDD can be understood as a fuzzy rule based classification systems (see [3]). As such, the knowledge or rules that guide the

classification process have to be learned from training examples, and the reasoning process that assigns a class to a query needs such a query to be described in the same terms as the learning examples. Training examples for SEDD are provided by EM-DAT (Emergency Database, see [www.em-dat.be](http://www.em-dat.be)), the most exhaustive and complete public database about disasters and emergencies. However, EM-DAT provides a rather incomplete description of the locations affected by disasters. For this reason, EM-DAT has been merged with both UNDP data about the Human Development Index (*HDI*), as well as with US Census historical data on population densities (*POP*). Also, EM-DAT informs on the type of adverse phenomenon that produced each registered disaster, its *magnitude* (*MAG*) or intensity as well as on the effects it produced in terms of a set of consequences, that range from the number of casualties (*CAS*) to the number of homeless people or the extent of the material damages (see Table 2 below).

Variable	Description
Magnitude	Intensity of the adverse phenomena: degrees on the Richter scale for earthquakes, inundated area in km <sup>2</sup> for floods, etc.
HDI	Human Development Index: an estimation of the affected country's <i>vulnerability</i> at the moment of the strike.
Population Density	Population density of the affected country at the moment of the strike: an estimation of the affected place's <i>population-at-risk</i> .
Casualties	Number of casualties produced by the disaster.
Injured	Number of injured people.
Homeless	Number of homeless people.
Affected	Number of affected people.
Damage	An estimation of the amount of infrastructural damage in thousands of US dollars.

**Table 2:** Relevant variables contained in EM-DAT for each disaster type.

The descriptive variables *MAG*, *HDI* and *POP* are taken as explanatory or independent, while those related to effects or consequences, as *CAS*, are taken as dependent variables to be assessed. Thus, SEDD need a scenario to be described in terms of these independent variables in order to produce an assessment of its consequences. It is important to notice that EM-DAT poses a difficult classification problem, since the variability of the consequences is huge for similar values of the independent variables. Also, the sample is highly unbalanced, as can be observed in Fig. 2 below.



**Figure 2.** Fuzzy partition and linguistic labels for the variables *Magnitude of an earthquake (MAG)* and *Human Development Index (HDI)*.

Thirdly, a learning procedure has to be applied in order to extract the rules from the data. Therefore, let us denote by  $X_1, \dots, X_n$  the  $n$  attributes or independent variables that are used to describe a disaster scenario, and assume that a set of  $m$  historical disaster scenarios  $(x_1^p, \dots, x_n^p; C^p)$  is available as learning sample, where for each  $p = 1, \dots, m$ ,  $C^p$  is one of the classes in  $\zeta = \{C_1, \dots, C_N\}$  defined to (linguistically) assess the consequence or dependent classification variable  $Y$ . Assume also that a dissimilarity matrix  $\Delta = (d_{ij})_{N \times N}$  has been defined modeling the opposition relationships between the classes in  $\zeta$ . For all  $i = 1, \dots, n$  let us also denote by  $A_{i1}, \dots, A_{ic_i}$  the  $c_i$  linguistic labels defined for each attribute  $X_i$ , in such a way that  $\mu_{A_{ij}}(x_i) \in [0, 1]$  represents the degree up to which the value  $x_i$  fulfills the label  $A_{ij}$ . A premise  $A$  of a rule is given by the combination of a label  $A_{ij_i}$  of each attribute, i.e.

$A = A_{i1} \times \dots \times A_{in}$ . The degree of fulfillment of a premise  $A$  by  $x = (x_1, \dots, x_n)$  is usually obtained through a t-norm (see [15])  $T$ , i.e.  $\mu_A(x) = T(\mu_{A_{i1}}(x_1), \dots, \mu_{A_{in}}(x_n))$ .

In these conditions, following [11], a rule with premise  $A$  and having as a consequent a class  $C_j \in \zeta$  is evaluated by means of a pair of confidence degrees

$$r_j^+(A) = \frac{\sum_{p=1, \dots, m/C^p=C_j} \mu_A(x^p)}{\sum_{p=1}^m \mu_A(x^p)}, \quad r_j^-(A) = \frac{\sum_{p=1}^m \mu_A(x^p) \cdot \mu_{\Delta C_j}(C^p)}{\sum_{p=1}^m \mu_A(x^p)}$$

where  $\mu_{\Delta C_j}(C_i) = d_{ji}$ . Notice that  $r^+, r^- \in [0, 1]$  and that it holds that  $r^+ + r^- \leq 1$ . As explained in [11],  $r_j^+(A)$  and  $r_j^-(A)$  respectively estimate the proportion of positive examples and significant exceptions of the rule  $A \Rightarrow C_j$  out of the total number of training patterns fulfilling the premise  $A$ . Therefore, the dissimilarity structure is introduced in the evaluation of the rules by means of the negative confidence degree. Different procedures (see for instance [2]) can be used to obtain the set of premises  $A$  for which rules  $A \Rightarrow C_j$  ( $j=1, \dots, N$ ) have to be built, basically ensuring that each training pattern is *covered* by at least one rule. Once this learning stage is finished, a set of rules or *rule base* is available, which represents the knowledge of the classifier. It is important to remark that a main advantage of descriptive fuzzy classifiers (like SEDD) is that they provide rules expressed in terms of a natural language, so their knowledge is explicit and interpretable.

Lastly, a fuzzy reasoning method (see [3]) has to be applied in order to produce an assessment of a disaster scenario  $x = (x_1, \dots, x_n)$  under study. Following [11], here we apply the VA1 (*one-dimensional additive veracity*) reasoning method, that computes the *degree of veracity*  $t_j(A) = \max\{r_j^+(A) - r_j^-(A), 0\}$  for each of the  $R$  available rules  $R^q: A^q \Rightarrow C_j$ , and then obtains the

evidence degree  $t_j(x)$  for the classification of the scenario  $x$  in each class  $C_j$  by means of the expression

$$t_j(x) = \frac{\sum_{q=1, \dots, R} \mu_{A^q}(x) \cdot t_j(A^q)}{\sum_{q=1, \dots, R} \mu_{A^q}(x)}.$$

Therefore, the vector  $t(x) = (t_1(x), \dots, t_N(x))$  constitutes the final output of SEDD, assessing the degree of evidence for each class or level of consequences. If a crisp prediction is needed, then it is usual to assign the scenario  $x$  to the class with maximum evidence, i.e. to the class  $C_h$  such that  $t_h(x) = \max_j t_j(x)$ .

### 3.2. Dissimilarity structures for disaster assessment

Here we illustrate the effect of different dissimilarity structures on the assessment provided by SEDD. Particularly, in order to be able to produce a *picture* of the assessments obtained by using each dissimilarity matrix  $\Delta$ , we drop the population density *POP* from the set of independent variables, which leaves the variables *MAG* and *HDI* as the only explanatory variables to be used. Similarly, in this example we will focus on just one consequence variable, the number of casualties *CAS*, and on one type of disaster, *earthquakes*. Therefore, it is  $n=2$  and  $N=5$  (the same classes as in Table 1 are used for the variable *CAS*). The training sample for these explanatory and dependent variables is shown in Fig. 3, and it is  $m=386$ . As the maximum number of premises is quite small ( $5 \cdot 5 = 25$ ), we adopt a grid-based learning procedure, i.e. rules are built for all possible premises. However, a support threshold ( $\delta=0.01$ ) is defined in order to avoid those rules built from a too small sample.

In order to illustrate and compare the effects of each matrix  $\Delta$ , the behavior of the resulting VAI classifiers is simulated in a dense mesh of points of the input space of the attributes *MAG* and *HDI*, in such a way that a picture of the predictions and class boundaries produced by each dissimilarity structure is obtained. Also, two error measures are used in order to measure the performance of the different classifiers: 1) %CC represents the rate of correct classifications obtained over the training sample, thus evaluating the predictive accuracy of each classifier; 2) to measure both the deviation of the predictions from the real classes and the risk of underestimation, the average cost *AVCOST* of the predictions is computed over the training sample, where the cost of classifying a instance from the class  $i$  in the class  $j$  is given by the element  $COST_{ij}$  of the matrix

$$COST = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 2 & 0 & 1 & 2 & 3 \\ 4 & 2 & 0 & 1 & 2 \\ 6 & 4 & 2 & 0 & 1 \\ 8 & 6 & 4 & 2 & 0 \end{pmatrix}.$$

#### 3.2.1 No dissimilarity

Let us start by assuming that no opposition relationships hold between the classes, i.e. by taking

$\Delta \equiv 0$ . In this case, it is  $r_j^-(A) = 0$  for every premise  $A$  and consequent  $C_j$ , so it is  $t_j(A) = r_j^+(A)$  for all rules  $A \Rightarrow C_j$ . Therefore, this case corresponds to a non-structured classification framework, in which a usual non-bipolar fuzzy classifier, identical to that used by SEDD in [12], is obtained. As no structure is assumed on the set of classes, the resulting classifier treats all the classes as independent items. Consequently, the classifier will be biased towards the classes with more training examples. This is clearly shown in Fig. 3, where the results of the simulation are depicted. Note that a huge part of the input space of the attributes is assigned to the lowest class *CAS1=no casualties*, i.e. that with the highest proportion (54.4%) of training patterns. This entails a great risk of underestimation of consequences. Furthermore, the behavior of the classifier is not smooth at all, since predictions present sharp variations. Table 3 presents the performance measures for this non-bipolar classifier. Though we will use these results for comparison with the rest of classifiers, it is important to remark that almost all the correct classification rate (%CC=54.15) is due to examples of class *CAS1*. In fact, notice that %CC is in this case almost equal to the proportion of examples from the class *CAS1*.

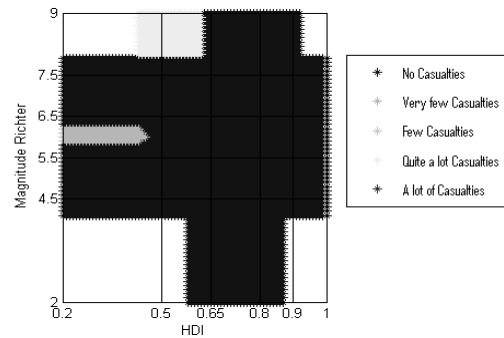


Figure 3. Simulation result of the classifier with  $\Delta \equiv 0$ .

%CC	AVCOST
54,15	1,453

**Table 3:** Performance measures of the non-bipolar classifier ( $\Delta \equiv 0$ ).

#### 3.2.2 Total opposition

As explained above, dissimilarity matrices range between the extremes given by  $\Delta \equiv 0$  and  $\Delta = 1 - Id$ . Let us analyze now this last case, that corresponds to a situation in which each class is totally opposite to the others. Notice that now it is  $r_j^-(A) = 1 - r_j^+(A)$ , and thus  $t_j(A) = \max\{2r_j^+(A) - 1, 0\}$ . Therefore, since all classes are equally related, the same bias as before towards the more abundant classes is obtained. However, a rule  $A \Rightarrow C_j$  will obtain  $t_j(A) = 0$  unless  $r_j^+(A) > 0.5$ , i.e. unless more than a half of the training examples compatible with the premise  $A$  belong to class  $C_j$ . If no class fulfills this condition, then it is  $t_j(A) = 0$  for all  $j$ . As a result, it is possible for a query  $x$  to keep unclassified if no activated rule have a positive veracity degree. The simulation of this classifier clearly

illustrates this point, as shown in Fig 4. A half of the input space is left unclassified, while the other half is assigned to the majority class CAS1, the only one that reaches the 0.5 threshold for some premises. In this sense, if the previous classifier predicts the class with the greatest (positive) evidence, then it is possible to say that the present classifier only give a prediction if the evidence for a class is much larger than for the others. To some extent, this classifier can be associated with a requirement of not giving an assessment unless robust, strongly supported predictions are feasible. In this sense, it can be used for a first assessment of whether a scenario have no consequences at all or not. Table 4 shows the performance measures of this classifiers. Notice that 43.78% and 38.86% of the  $m=386$  training examples are respectively left unclassified and correctly classified. This gives an error rate of only 17.36% (the previous was 45.85%), though at the price of not-classifying almost a half of the sample. Similarly, the average cost is significantly lower, since a great part of the instances from the highest classes are left unclassified.

%NC	%CC	AVCOST
43.78	38.86	0.811

**Table 4:** Performance measures under a total opposition between classes ( $\Delta \equiv 1 - Id$ ).

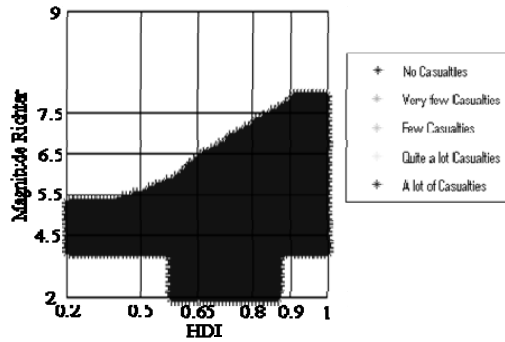


Figure 4. Simulation result of the classifier with  $\Delta \equiv 1 - Id$ .

### 3.2.3 Restricted asymmetric linear order

In order to reproduce the semantics of linear order associated with the classes, let us now introduce the dissimilarity matrix

$$\Delta_1 = \begin{pmatrix} 0 & 0.2 & 0.2 & 0.5 & 1 \\ 0 & 0 & 0.2 & 0.5 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

This matrix produces a structure in which the higher classes are gradually more and more dissimilar to the two lowest classes, CAS1 and CAS2, but not symmetrically, i.e., the lower classes are not dissimilar to the higher ones. Moreover, classes CAS3-CAS5 are completely unrelated between them. Therefore, in this situation the classes CAS1 and CAS2 receive negative information from the higher ones in a progressive way, but not conversely, representing the linear structure of the classes together with the requirement of avoid underestimation risk. As a consequence, classes CAS1 and CAS2 will obtain a lower veracity degree in the presence of the higher classes, thus requiring more

evidence for the former classes in order to be predicted. However, if it is estimated that such a risk can be disregarded for the higher classes (e.g. if CAS3 and higher scenarios are always further assessed), then it is possible to restrict the linear order structure to the first two classes and allow the classes to compete freely between them, similarly to what happened for all the classes when  $\Delta \equiv 0$ . These are the assumptions behind the matrix  $\Delta_1$  above. Compared to those of the non-bipolar classifier in Fig. 3, the simulation results now show a smoother behavior, in which the class CAS3 (in light blue in Fig. 5) appears in the transition zone between the lower and the upper classes. Notice that, in general, the upper classes obtain a greater portion of the input space than before. In fact, as shown in Table 5, this classifier obtains an average cost of 1.069, thus reducing the underestimation risk of the non-structured case. Furthermore, this classifier obtains a better classification rate (%CC=55.18) than the non-bipolar one, i.e. the consideration of a dissimilarity structure leads in this case to a more accurate classifier than without it.

%CC	AVCOST
55.18	1.069

**Table 5:** Performance measures for the restricted asymmetric linear structure ( $\Delta \equiv \Delta_1$ ).

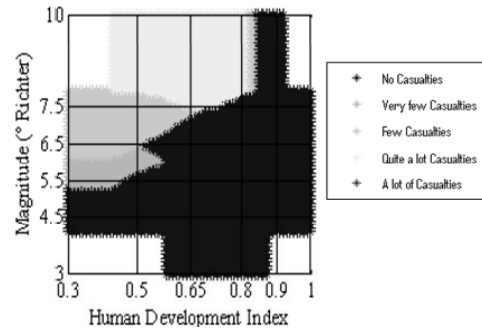


Figure 5. Simulation result of the classifier with  $\Delta \equiv \Delta_1$ .

### 3.2.4 Worst-scenario analysis

Finally, let us consider the requirement of a total avoidance of the underestimation risk. To this end, consider now the dissimilarity matrix

$$\Delta_2 = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

This leads to a structure in which each class is totally opposite to all the classes lower than it, but not conversely. Therefore, in this setting there is not an explicit linear structure, but just the assumption of a total asymmetry between the lower and the higher classes. Each class receives the positive confidence of all the higher classes as negative information. Thus, the lower the class, the harder it is for such class to be predicted. In this sense, matrix  $\Delta_2$  fits to the requirement of performing a worst case analysis of the disaster scenario under study. Notice that in this setting, no class is guaranteed to be predicted unless it attains a positive confidence  $r^+ > 2/3$ , except the highest,

which is predicted whenever its confidence is bigger than 1/3. Consequently, the class *CAS5* appears for the first time in the simulation results of this classifier, as shown in Fig. 6. Note also the improved smooth behavior of the classifier, producing a soft transition between classes. In fact, a straight line could be drawn in the input space passing through all the classes in order. More importantly, this classifier enable to distinguish a clear trend in the consequences, in such a way that worst consequences are associated with lower HDI values (and thus with a greater vulnerability) and greater intensities of earthquakes. This trend is logically expected, but notice that no one of the previous classifiers could express it so clearly. The performance measures of this classifier, shown in Table 6, presents a further reduction of the average cost (AVCOST=0.968) and thus of the underestimation risk, that could be even more important since the rate of correct classification (%CC=50.52) is lower than before (with the subsequent increment of non-zero costs).

%CC	AVCOST
50.52	0.968

**Table 6:** Performance measures for the worst case scenario analysis ( $\Delta \equiv \Delta_2$ ).

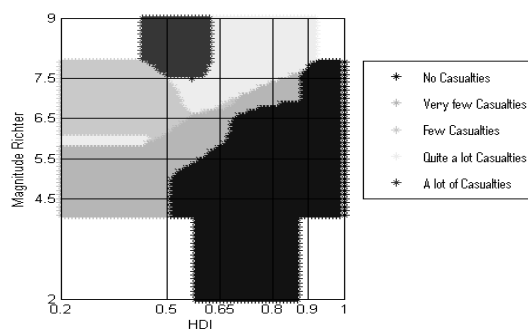


Figure 6. Simulation result of the classifier with  $\Delta \equiv \Delta_2$ .

#### 4. Conclusions

When considering the production of an initial assessment of disaster consequences as a structured classification problem, the structure of the set of classes gets closely related with the semantic features and decision requirements of the disaster management context. For this reason, an adequate characterization of this structure is an important step towards the adaptation of a classifier to those features and requirements. In this paper, we have shown that the introduction of a dissimilarity operator over the set of classes enables the consideration of several structures and their introduction in the classifier's learning and reasoning procedures. As a consequence, a significant improvement in the accuracy and adaptation of the classifiers to the decision requirements of the disaster management context is achieved.

#### References

- [1] Aleskerov, F., Iseri Say, A., Toker, A., Akin, H. L., Altay, G. (2005) A cluster-based decision support system for estimating earthquake damage and casualties, *Disasters*, 29 (3) 255-276
- [2] Casillas J, Cordon O, Herrera F (2002) COR: A Methodology to Improve Ad Hoc Data-Driven Linguistic Rule Learning Methods by Inducing Cooperation Among Rules. *IEEE Trans. on Systems, Man and Cybernetics-Part B: Cybernetics* 32 (4): 526- 537
- [3] Cordón, O., del Jesús, M. J., Herrera, F. (1999) A proposal on reasoning methods in fuzzy rule-based classification systems, *International Journal of Approximate Reasoning*, 20 (1) 21-45
- [4] Drabek, T. E., Hoetmer, G. J. (1991) Emergency management: principles and practice for local government, *International City Management Association*, Washington DC
- [5] Hüllermeier, E. (2005) Fuzzy methods in machine learning and data mining: Status and prospects, *Fuzzy Sets and Systems*, 156 (3) 387-406
- [6] Mendonça, D., Beroggi, E. G., Wallace, W. A. (2001) Decision support for improvisation during emergency response operations, *International Journal of Emergency Management*, 1 30-38
- [7] Montero, J., Gomez, D., Bustince, H. (2007) On the relevance of some families of fuzzy Sets, *Fuzzy Sets and Systems*, 158 (22) 2429-2442
- [8] Morrow, B. H. (1999) Identifying and mapping community vulnerability, *Disasters*, 23 (1) 1-18
- [9] Olsen, G. R., Carstensen, N., Høyen, K. (2003) Humanitarian crises: What determines the level of emergency assistance? Media coverage, donor interests and the aid business, *Disasters*, 27 (2) 109-126
- [10] Rodríguez J. T., Franco C.A., Vitoriano B., Montero J. (2011) An axiomatic approach to the notion of semantic antagonism, *Procs IFSA-AFSS'11 FT104-1/6*
- [11] Rodríguez, J. T., Vitoriano, B., Montero, J. (2011) Rule-based classification by means of bipolar criteria. 2011 IEEE Symposium on Multicriteria Decision Making (SSCI-MCDM) 197-204
- [12] Rodríguez, J. T., Vitoriano, B., Montero, J. (2012) A general methodology for data-based rule building and its application to natural disaster management, *Computers & Operations Research*, 39 (4) 863-873
- [13] Rodríguez, J. T., Vitoriano, B., Montero, J., Omaña, A. (2008) A decision support tool for humanitarian operations in natural disaster relief, *Computational Intelligence in Decision and Control*, 1 805-810
- [14] Rodríguez, J. T., Vitoriano, B., Montero, J., Kecman V. (2011) A disaster-severity assessment DSS comparative analysis. *OR Spectrum* 33(3) 451-479
- [15] Schweizer B., Sklar A. (1983) *Probabilistic Metric Spaces*. North-Holland/Elsevier, New York
- [16] Vitoriano, B., Ortuño, M. T., Tirado, G., Montero, J. (2010) A multi-criteria optimization model for humanitarian aid distribution, *Journal of Global Optimization (JOGO)*, 51 189-208
- [17] Wallace, W. A., De Balogh, F. (1985) *Decision Support Systems for Disaster Management*, *Public Administration Review*, 45 134-146