A fuzzy expert system for automatic seismic signal classification

El Hassan Ait Laasri a,⁎, Es-Said Akhouayri a, Dris Agliz a, Daniele Zonta b, Abderrahman Atmani a

a Electronic, Signal processing and Physical Modelling Laboratory, Department of Physics, Faculty of Sciences, Ibn Zohr University, B.P. 8106 Agadir, Morocco
b Department of Civil, Environmental and Mechanical Engineering, University of Trento, via Mesiano 77, 38123 Trento, Italy

Keywords:
Seismic signal classification
Seismic signal feature extraction
Fuzzy rule based expert system
Fuzzy reasoning
Fuzzy set

Article info
Available online 27 August 2014

Abstract
Automatic classification of seismic events is of great importance due to the large amount of data received continuously. Seismic analysts classify events by visual inspection and calculation of event signal characteristics. This process is subjective and demands hard work as well as a significant amount of time and considerable experience. A reliable automatic classification task considerably reduces the effort required and makes classification faster and more objective. The aim of this study is to develop a fuzzy rule based expert system that is able to imitate human reasoning and incorporate the analyst’s knowledge of seismic event classification. The fundamental idea behind using this approach was motivated by the way in which human analysts classify seismic events based on a set of experiential rules. Additionally, this approach was chosen due to its interpretability and adjustability, as well as its ability to manage the complexity of real data. Relevant discriminant features are extracted from event signal. Using these features, the classification system was built based on the vote by multiple rule fuzzy reasoning method with three types of rules. Comparison of this method with the single winner classical fuzzy reasoning model was carried out. Classification results on real seismic data showed the robustness of the classifier and its capability to operate in on-line classification.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

The ease with which a trained human brain can classify objects, recognize faces or voices, understand spoken languages, read handwritten characters, etc., has inspired many scientists to seek to design and build machines that can mimic the human recognition system in recognizing patterns. Nowadays, with the rapid development in computer technology and digital systems, pattern recognition has gained more and more interest because of its practical importance in nearly all branches of applied science. In this study, we are particularly interested in signal classification, and more precisely seismic signal classification. In a signal classification problem, the aim is to design a minimum-error system for labeling an input signal with one of a given set of classes.

Seismic monitoring networks detect and record a huge number of seismic events continually. Such events can be produced from a wide variety of sources. In fact, seismic waves can be generated by natural sources such as earthquakes and ocean waves, or by man-made sources such as explosions and cultural activities (industry, traffic, etc.). Classification of seismic events is of great importance due to several reasons. Besides the need for recognizing earthquake signals to launch an early alarm, classification of seismic signals on a routine basis provides an organized database that could be easily exploited in seismic research. Location and investigation of active tectonics as well as analysis of seismic hazards in a region of interest are the main focuses of seismic studies. In these studies, identification of earthquake signals is the critical first step. Recognition of explosion signals is also a practical need in a wide range of applications. For example, detection and identification of nuclear explosion signals is of considerable interest in the context of nuclear test ban treaty verification (CTBT) (Hoffmann, Kebeasy, & Firbas, 1999).

Seismic signal classification is traditionally performed through manual analysis based on visual inspection and calculation of signal features such as spectral characteristics (Bormann et al., 2009; Bormann, Klinge, & Wendt, 2009). This process is subjective and demands hard work as well as considerable experience and a significant amount of time. Moreover, In view of the very high volume of seismic data received continuously, the considerable daily processing conducted by analysts is stressful and arduous. Therefore, constructing a reliable automatic classification system is crucial to considerably reduce the effort required as well as to make classification faster and more objective.

Due to the importance of an automatic classification task, many research efforts have been devoted to designing seismic
classification systems. Several methods have been developed and tested in many countries. The shape of recorded signals, the feature extraction method and the classifier are often different. This is mostly due to the fact that characteristics of seismic signals are closely related to properties of the medium, which vary significantly from area to area (Jenkins & Sereno, 2001; Zeller & Velasco, 2009). The simple classifier that has been mostly used in regions where repeating sources occur is based on the cross-correlation function (Harris, 1991; Joswig, 1990, 1995). In these regions, a set of signals are collected from previously occurred seismic events to represent the prototypes of classes. Cross-correlation function can then be used to recognize subsequent signals of each class. The drawback of this method is that it is very sensitive to the form of the signal, which is mostly distorted by the environment. Spectral ratios of seismic waves are commonly presented as good discriminants between earthquakes and mining blasts (Allmann, Shearer, & Hauksson, 2008; Kim, Park, & Kim, 1998; Walter, Mayeda, & Patton, 1995). Nevertheless, such information is not available all the time because of the fact that not all seismic signals present clear and separable seismic waves. In this case, the spectral ratio does not appear to be sufficient to discriminate these two classes of events. Limitation of these techniques has led to more sophisticated methods. Methods based on the Hidden Markov modelling were applied to volcanic event classification (Benitez et al., 2007). The capability of an unsupervised clustering technique based on the maximum likelihood estimation was tested on the seismic database of Iran (Ansari, Noorzad, & Zafarani, 2009). Various types of neural networks have been applied to seismic event discrimination between earthquakes and explosions (Ait Laasri, Akhouayri, Agliz, & Atmanli, 2011; Curilem, Vergara, Fuentelaiba, Acuña, & Chacón, 2009; Dysart & Pulli, 1990; Kuyuk, Yildirim, Dogan, & Horasan, 2011; Musil & Plesinger, 1996; Shimshoni & Intrator, 1998; Yildrima, Gulbag, Horasana, & Dogan, 2010; Zadeh and Nassery,1999), as well as classification of volcanic events (Langer, Falsaperla, Powell, & Thompson, 2006; Scarpette et al., 2005). A genetic algorithm-based boosting approach has been used to discriminate between earthquake and explosion signals (Orlic & Loncaric, 2010).

In this study, we propose a Fuzzy Rule Based Expert System (FRBES) for automatic classification of seismic events. The fundamental idea behind this approach arises from the fact that seismic signal analysts are able, based on their experience of looking at many seismograms, to classify seismic events by visual inspection and calculation of signal characteristics. We are interested in transforming the analyst’s knowledge of seismic signal classification into a FRBES which can imitate the reasoning process of the analyst in solving the classification problem. The use of FRBES as a classifier for seismic signals is motivated by several reasons:

- FRBES can be built based on the training data as well as the heuristic knowledge and experience of people who have already mastered the classification procedure, in direct contrast to many other methods such as artificial neural networks which are entirely based on the training data. Due to years of experience and potential of human brain, it is clear that seismic signal analysts construct solid knowledge and ability to classify seismic events. These analysts usually employ experiential rules, which can be cast into a FRBES. In addition, the possibility of mixing different information such as that coming from expert knowledge and information coming from training data may significantly improve the generalization ability of the classifier and so its performance.
- FRBES has proved to be a powerful framework to incorporate imprecise knowledge using the concept of linguistic variables instead of precise quantitative values used in conventional mathematical tools. Such linguistic variables are more suitable for processing real world data and particularly seismic data, which are mostly imprecise, noisy and distorted.
- More importantly, FRBES employ linguistic variables and fuzzy rules which are human-understandable. These rules increase the transparency and interpretability of fuzzy systems and enable them to be easily comprehensible for humans, contrary to many other approaches, especially the artificial neural networks, which are considered as black-boxes. Furthermore, owing to its transparency, modifiability, comprehensibility and coherence with previous knowledge, a FRBES can be tuned, updated and adapted if necessary, thus enhancing the degree of freedom of the classifier adjustability.

These capabilities inspired many researchers to apply fuzzy algorithms to a variety of classification problems, including hydro-meteor classification (Marzano, Scaranari, & Vulpiani, 2007), speech/music classification (Reyes, Candeas, Galan, & Munoz, 2010), vehicle classification (Kim, Kim, Lee, & Cho, 2001), chemical process classification (Dash, Rengaswamy, & Venkatsubramania, 2003), and image classification (Sharma, Gupta, Kumar, & Kapoor, 2011). Many other investigators have sought to improve the performance and accuracy of FRBES for classification (Ho, Chen, Ho, & Chen, 2004; Mansoori, Zolghadri, & Katebi 2007; Sanz, Fernández, Bustince, & Herrera, 2010, 2011).

The rest of this paper is organized as follows. First, we briefly discuss seismic data characteristics in Section 2. In Section 3, the first part is devoted to feature extraction, and the second part presents the proposed classification system. The experimental results and performance of the system are shown in Section 4. Finally, Section 5 outlines conclusions.

2. Seismic data and their characteristics

2.1. Data

The data used in this work to evaluate our classification system were recorded by the local seismic network of Agadir. The latter belongs to the national seismic network of Morocco and consists of vertical-component short-period seismometers with an output proportional to ground velocity. The seismometers are deployed around Agadir city and linked with the data center (located in Agadir) via a radio-frequency FM modulated network. Seismic data are continuously acquired and transmitted in real-time to the data center where they are digitized and analyzed. Each detected event is recorded with the pre-event and post-event data in order to assure complete recording of seismic events. The employed detector in the local seismic network of Agadir, as in many other seismic networks around the world, is the STA/LTA energy-based detector (Akhouayri, Agliz, Fadel, & Ait Ouahman, 2001; Trnkoczy, 2009). This detector triggers whenever a sufficiently data segment energy is encountered. In this manner, the recorded seismicity is continually contaminated by various kinds of events that can trigger the detection algorithm. Because of many quarries located surrounding Agadir city, numerous quarry blast seismograms are recorded on a daily basis. Besides earthquake and quarry blast events, many additional sources are responsible for several other detections. Such sources are considered in this study as seismic noise sources, which incorporate wind, ocean waves and cultural activities (e.g., machinery). Typical vertical-component seismograms of four seismic sources are plotted in Fig. 1.

Classification of these events has been performed manually by human analysts who are sufficiently familiar with the various kinds of waveforms. However, the significant daily processing conducted by analysts is stressful and time consuming. It thus
becomes necessary to develop a task that can automatically classify seismic signals.

This work arose from the need to implement an automatic classification system that can allocate each detected seismic event to one of the following classes: local earthquake (LE), distant earthquake (DE), noise (NS) and quarry blast (QB). Recognition of local earthquake signals is indispensable to ensure accurate analysis of active faults and seismic hazards in the region. On the other hand, identification of quarry blast signals helps in other studies such as examination of the impact of explosion on structures. Distinguishing between local and distant earthquake signals is necessary because the latter type of events is not properly recorded by the local network.

2.2. Characteristics of seismic records and their sources

Understanding the fundamental characteristics of seismic signals and their sources is vital and essential for extracting the most relevant and discriminant features among classes.

The name “seismic source” assembles any generator of seismic waves. A seismic source can be simple or much sophisticated depending on many parameters such as its strength, its spatial and temporal characteristics. A tectonic earthquake source can be thought of as a sudden movement of rocks over a surface rupture that occurs along the fragile part of the Earth’s crust due to a force that exceeds its breaking strength. Explosions are usually shallow and mostly man-made and controlled (i.e., with known weight of explosive, location and explosion time) while tectonic earthquake sources can be deep, much larger and more complex. Due to the nature of forces acting at the source of each event type, the excited waves as well as their characteristics are different. Earthquake events radiate both P and S waves in different directions but with different amplitudes and polarities (Bormann et al., 2009; Bormann, Klinge, & Wendt, 2009). Whereas, explosion sources are instantaneous and produce almost a spherically expanding compressional P wave of approximately the same amplitude in all directions (Bormann et al., 2009; Bormann, Klinge, & Wendt, 2009). Seismic noise sources can be considered as the influence, for instance, of oceans waves, wind and human activities on the solid Earth.

It is obvious that the processes acting at the source of an explosion or seismic noise are very different from those acting at the source of a natural earthquake. These different processes may introduce recognizable characteristics to seismic signal, which can be readily used to identify the type of the source that produced the signal. It is expected that explosions generate much simpler signals, consisting of P waves and lacking or not significant S waves. On the contrary, earthquakes sources would create more complicated waveforms with P wave and significant shear wave component followed by longer coda waves. On the other hand, seismic noise sources are mainly superficial and engender predominately surface waves such as Love and Rayleigh waves.

Fig. 2 represents examples of vertical component seismogram of four seismic event types recorded by the local seismic network of Agadir, together with their corresponding envelopes and FFT. In a preliminary look at these seismograms, it seems that signal shape is a promising classification parameter. Compared to tectonic earthquakes, QB records are characterized by no clear S wave, less impulsive onset and very short duration of coda waves. These characteristics intervene jointly to give the signal shape a Gaussian-like envelope.

Signal associated with earthquake events differs appreciably from that of QB events as it involves larger S and surface waves, overlapping or isolated P and S waves; depending on the source-station distance, and an exponential decay of coda amplitude with time.

Analysis of many QB signals shows that most of them have the same shape, and can be recognized using only the envelope. Unfortunately, not only QBs signals show this feature, LE and cultural activities may also generate signals with the same envelop as QB. In such situations, other discriminant features should be employed. Initial observation of seismograms FFT reveals distinctive characteristics among the four types of events. It is evident that DE events will be easily identified in the frequency domain as their associated signals generally present low frequency content. This is principally due to the strong heterogeneity within the medium as well as the high frequency absorption along propagation path (the frequency content decreases rapidly as the source-receiver distance increases). This effect was successfully utilized in recognizing this.
type of events. Fig. 2b depicts an example of such events. After traversing a long propagation path, the high frequencies are filtered by the medium which acts as a low-pass filter (Fig. 2f). Examination of many seismograms shows that the dominant frequency of this type of events is generally below 6 Hz.

Frequency domain is also an appropriate representation for NS events identification. It was observed that records of these events are often narrow band (Fig. 2l). The combination of this information with that extracted from time domain such as signal duration and skewness can effectively classify such records.

Inspection of many event signals indicates that seismograms of QB are poorer in high frequencies than seismograms of LE. The same results demonstrate that this type of events present mostly narrower frequency content which is generally above 1 Hz. Examples of a typical QB signal together with its frequency content are presented in Fig. 2g and k.

Another important parameter to be considered in this work is time duration of events records. It is influenced by many factors, including the characteristic dimensions of the source. Thus, it should be expected that time duration of natural earthquakes would be longer than time duration of explosions. In this study, it was found that QB records have durations of less than 40 s while tectonic earthquake records may last for several minutes.

The simplest variable that can be used to exclude QB events from the others is the hour of detection. This parameter relays on the fact that QB are launched during specific hours. This technique is used in many local networks for discrimination between local earthquakes and explosions (e.g., Agnew (1990)).

3. Methodology

Signal classification problems have been addressed in almost all cases as a three-stage process (see Fig. 3): signal pre-processing stage, signal feature extraction stage and signal classification stage. The former two stages are used to produce more consistent data and then extract a minimum number of features where the most relevant information for signal classification is presented. The latter stage uses the features extracted in the previous stage to identify the class of the signal.

Fig. 2. Vertical component seismogram of four seismic events generated by different sources and recorded by the seismic local network of Agadir: (a) local earthquake LE, (b) distant earthquake DE, (g) quarry blast QB, (h) machinery NS, together with their corresponding envelope (c, d, i, j) and FFT (e, f, k, l), respectively.
3.1. Signal processing and feature extraction

The performance of the classification system mainly depends on the quality of the feature set used to represent the high dimensionality of seismogram. Therefore, extraction of the most relevant features contributes to the reliability of the classification system. Additionally, it is highly suitable to reduce noise effect.

The feature set mostly used by seismic signal analysts corresponds to:

3.1.1. Envelop similarity $E_s$

In order to extract the signal envelope $E$ of the vertical component seismogram $z(t)$, we use the amplitude of the analytical signal. $E$ can be expressed as follows:

$$E(t) = \sqrt{z(t)^2 + HT[z(t)]^2}$$

where $HT$ is the Hilbert Transform. Suitable

As described in the previous section, almost all QB events possess the same signal shape. The high rate of resemblance among QB signal envelopes allows us to use this characteristic to identify a large percentage of their seismograms. The envelop similarity $E_s$ is measured using Manhattan distance $D$ (Miśkiewicz, 2012) between a normalized reference envelope $E^r$, determined from previously recorded explosion events, and the normalized envelope $E$ of each incoming event:

$$E_s = \frac{D(E, E^r)}{\sum E}$$

Before measuring $D$, the maxima of the two envelopes are coincided. A finite impulse response filter (FIR) was designed to minimize the rapid variation of the envelope.

$E_s$ of an incoming event with approximately the same envelop as QB one tends towards 0. In contrast, $E_s$ of an incoming event with larger envelop tends towards 1. $E_s$ can be greater than 1 if the incoming event envelop is very shorter than the reference one. This is the case when the event belongs to NS class.

3.1.2. Duration $T_d$

It is defined as the total duration in seconds of the event record from the $P$ wave onset $t_p$ to the end of the signal $t_{end}$ defined as the point where the signal is no longer seen above the noise.

$$T_d = t_{end} - t_p$$

$t_p$ and $t_{end}$ are estimated using STA/LTA (Trnkoczy, 2009) algorithm with the signal envelop as a characteristic function. In order to take into account the effect of noise on the signal, thresholds are estimated according to the noise energy before $P$ wave.

3.1.3. Hour $H$

Day time event distribution reveals that quarry explosions are generally launched between 11:00 a.m. and 02:00 p.m. and between 05:00 p.m. and 06:00 p.m. GMT. Beyond this time intervals, explosion are absent and the seismicity pattern should not be affected by this type of event. Parameter hour $H$ can be computed as follows:

$$H = \text{hour} + \text{minute}/60 + \text{second}/3600$$

3.1.4. Spectral centroid $S_c$

It indicates the barycenter of the signal spectrum. This measure is obtained using the normalized amplitude of FFT envelope weighted by its corresponding frequencies.

$$S_c = \frac{\sum_{i=0}^{N} f(i)e(i)}{\sum_{i=0}^{N} e(i)}$$

$e(i)$ represents the amplitude of the FFT envelop of the bin number $i$, and $f(i)$ represents its frequency.

3.1.5. Spectral length $S_l$

$S_l$ of each event is estimated by applying thresholds on its FFT envelop.

$$S_l = f_n - f_0$$

where $f_0$ and $f_n$ are the first and last selected frequency bins.

3.1.6. Skewness $S$

It is used here to characterize the degree of symmetry or asymmetry of a signal around its mean. $S$ for a roughly symmetrical signal is near zero. $S$ of a real signal is given by:

$$S = \frac{1}{N} \sum_{i=1}^{N} (z(i) - \bar{z})^3 \quad \left(\frac{1}{N} \sum_{i=1}^{N} (z(i) - \bar{z})^2\right)^{3/2}$$

$N$ is the length of signal $z$ and $\bar{z}$ its mean.

The real data are inevitably affected by noise. Such noise could significantly alter the extracted features and result in unsatisfactory classification performance. Hence, applying a noise reduction algorithm to original signals may enhance the system performance. In this study, an adaptive filter has been implemented in the frequency domain (Anderson & Mcmechan, 2005; Vaseghi, 2006) (see Fig. 4). This technique aims to restore the spectrum of a signal, through subtraction of an estimate of the noise spectrum from the noisy signal spectrum. The noise spectrum is estimated from the pre-event noise.

This method is based on the assumption that noise is stationary and additive. Since the noise and noisy signal segments are not of the same length, the noise amplitudes are scaled by a factor equal to the number of samples in the noisy signal window divided by the number of samples in the noise window (Anderson & Mcmechan, 2005). After padding, the spectra of the noise and noisy signal are calculated and subtracted. Negative amplitudes are replaced with zeros.

3.2. Classifier design

Our classification system involves assigning a class $C_j$ from a predefined class set $C = \{LE, DE, QB, NS\}$, described by a set of features or attributes $X = \{E, Td, H, S_c, S_l, S\}$, to each incoming event $E^x$ represented by a feature vector $x^f = (x^f_1, x^f_2, x^f_3, x^f_4, x^f_5, x^f_6)$, where $x^f_5$ represents the value of the nth attribute $X_n$ associated with the event $E^x(x^f_i \in X_n)$. Thus, the problem of designing the classifier is to define an optimal mapping $F$, which could label unseen feature vectors with the smallest possible error, such as:

$$F : E_x \times Td \times H \times S_c \times S_l \times S \rightarrow C$$

Inspired by the capability of seismic analysts to recognize seismic signals, the aim is to develop a fuzzy expert system that is able to imitate human reasoning.

3.2.1. A fuzzy rule based expert system for seismic signal classification

Expert systems (Durkin, 1994; Jackson, 1999; Krishnamoorthy & Rajeev, 1996; Todd, 1992) are computer programs that attempt
to simulate the reasoning and behavior of an expert or an experienced analyst in solving a particular problem. There is a variety of approaches that have been proposed and employed to develop expert systems (Todd, 1992). The most common approach is fuzzy rule-based programs which use a collection of fuzzy rules to model and represent the qualitative aspects of human knowledge and reasoning process (Buchanan & Shortliffe, 1984; Buckley, Siler, & Tucker, 1986; Buckley & Tucker, 1989; Kandel, 1991, 1992; Siler & Buckley, 2005; Turksen, 1992). Actually, the FRBES is one of the most active research areas in artificial intelligence, and there are many scientists who believe that human reasoning and cognitive skills can be expressed and captured using fuzzy rules (e.g., (Anderson, 1993; Da & Kerre, 2000)). For example, Anderson interprets a production rule in his book ‘Rules of the Mind’ (Anderson, 1993) as “Each production rule is thought of as a modular piece of knowledge in that it represents a well defined step of cognition”.

A FRBES, in addition to capturing human knowledge, can handle the imprecision and uncertainty of real world data by using imprecise variables (Zadeh, 1975a, 1975b, 1975c); the term fuzzy refers to vagueness of the variables that can be managed by the system. These variables can accept linguistic expressions (Zadeh, 1999) which are commonly used by human to cope with uncertainty of natural phenomena. In order to manipulate such variables, fuzzy set and fuzzy logic (Klir & Yuan, 1995; Zadeh, 1965) are employed.

Fuzzy logic originates from the concept of fuzzy set theory, which was introduced by Zadeh (1965) as a means of representing and manipulating data that are not precise and exact, but rather fuzzy and approximate. In classical theory of sets, each element in a set is described by binary statements: an element must either definitely belong or not belong to the set. Therefore, the transition from one set to another is always abrupt. By contrast, the boundary of a fuzzy set is smooth. That is, the move from one set to its neighbors may be gradual rather than sharp. This gradual change allows an element to belong to a set to some partial degree by means of a membership function valued within the real unit interval \([0 \, 1]\). More formally, if \(X\) is the universe of discourse and its elements are denoted by \(x\), then a fuzzy set \(A\) in \(X\) is defined as a set of ordered pairs:

\[
A = \{x, \mu_A(x) | x \in X\} \quad \text{where} \quad 0 \leq \mu_A(x) \leq 1
\]

\(\mu_A(x)\) symbolizes the membership function of \(x\) in \(A\). The membership function maps each element of \(X\) to a membership value between 0 and 1 which represents the degree to which \(x\) belongs to the set \(A\).

The linguistic variables are usually defined as fuzzy sets with appropriate membership functions (Dombi, 1990). Boundaries of these concepts are vague and hence more suitable to deal with classes of objects encountered in the real physical world, which do not have precisely or sharply defined boundaries. These characteristics lend additional interest to FRBES. Indeed, FRBESs have found successful applications in a wide variety of disciplines including control, decision making and pattern recognition (Castanho, Hernandez, Re, Rautenberg, & Billis, 2013; Erik, Allahverdi, Sert, & Saritas, 2009; Jeongsu, Jeongsam, & Lee, 2012; Llata, Sarabia, & Oria, 2001).

A FRBES implements a nonlinear mapping from its input space to output space. This mapping is established by a set of fuzzy rules. The antecedents of the fuzzy rules partition the input space into a number of fuzzy regions, while the consequents specify the corresponding outputs (Jang & Sun, 1995). The basic structure of a fuzzy inference system can be split into two main conceptual parts: knowledge base and fuzzy reasoning method. The knowledge base consists of a set of if-then fuzzy rules, and defines the membership functions of fuzzy sets used in the fuzzy rules. The reasoning method provides the mechanism to perform the inference procedure. Due to various kinds of applications to which fuzzy systems are applied, different types of fuzzy inference were introduced. The three most popular and widely employed ones are Mamdani fuzzy model (Mamdani & Assilian, 1975), Takagi–Sugeno model (Takagi & Sugeno, 1985) and Tsukamoto fuzzy model (Tsukamoto, 1979). In a fuzzy classification system, the reasoning models are also different and depend on the type of fuzzy rule used. The employed reasoning method in this study is based on the general model proposed by Cordón, Jesus, and Herrera (1999). Fig. 5 illustrates the architecture of our FRBES.

In a FRBES, the performance level mainly depends on both the fuzzy rule structure employed and the fuzzy reasoning method. As is well known in fuzzy systems, the classical method, ‘single winner’ or ‘maximum matching’, is widely employed in the fuzzy reasoning process (Abe & Thawonmas, 1997; Ishibuchi & Nakashima, 2001; Marzano et al., 2007). This inference method classifies an input pattern with the consequent class of the single winner rule. In this way, we lose the information provided by all other fuzzy rules which may also represent relevant information.

In this paper, we aim to design a classification system that operates on the vote by multiple fuzzy rule-based reasoning process (Ishibuchi, Nakashima, & Morisawa, 1999), where all significant rules fired by an input event participate in the classification process. Such behavior seems more like human reasoning and may improve the classification result. In order to find the highest performance level, three types of fuzzy rules have been tested.

### 3.2.2. Knowledge base

#### 3.2.2.1. Fuzzy rule structure

Different types of rule have been used for fuzzy classification problems (Abe & Thawonmas, 1997; Ishibuchi & Nakashima, 2001; Ishibuchi & Yamamoto, 2005; Ishibuchi et al., 1999). In this work, we will consider a fuzzy expert system with a rule base of the following fuzzy rule form:

\[
R_i: \text{if } x_1 \text{ is } A_{i1} \text{ and } x_2 \text{ is } A_{i2} \text{ and } \ldots \text{ and } x_n \text{ is } A_{in} \text{ then Class is } C_i \text{ with } \omega_{i1}, \ldots, \text{ and } \omega_{im}\]

where \(R_i\), \(i = 1, \ldots, L\) is the rule label, \(i\) is the rule index and \(x = (x_1, x_2, \ldots, x_n)\) is a feature vector. \(A_{in}\), \(n = 1, \ldots, N\) is an antecedent fuzzy linguistic value represented as a fuzzy set by a membership function \(\mu_{in}\). \(\omega_j = 1, \ldots, M\) is the certainty degree (or rule weight) for the rule \(R_i\) to select the class \(C_i\). Usually \(\omega_j\) is real number in the unit interval \((0 \leq \omega_j \leq 1)\). Our fuzzy rule set can be written as the following:

![A block diagram illustration of spectral subtraction.](image-url)


Fig. 5. Architecture of the fuzzy seismic event classification system.

\[ R_1: \text{if } e_i \text{ is } A_{11} \text{ and } t_i \text{ is } A_{12} \text{ and } h \text{ is } A_{13} \text{ and } s_i \text{ is } A_{14} \text{ and } s_i \text{ is } A_{15} \text{ and } s \text{ is } A_{16} \text{ then Class is } LE \text{ with } \omega_{11} \text{ and } RE \text{ with } \omega_{12} \text{ and QB with } \omega_{13} \text{ and NS with } \omega_{14} \]

…

\[ R_2: \text{if } e_i \text{ is } A_{31} \text{ and } t_i \text{ is } A_{32} \text{ and } h \text{ is } A_{33} \text{ and } s_i \text{ is } A_{34} \text{ and } s_i \text{ is } A_{35} \text{ and } s \text{ is } A_{36} \text{ then Class is } LE \text{ with } \omega_{31} \text{ and } RE \text{ with } \omega_{32} \text{ and QB with } \omega_{33} \text{ and NS with } \omega_{34} \]

Three cases can be considered depending on the given values to \( \omega_{ij} \).

Case 1: Fuzzy rules without certainty degrees and only a class in the consequent.

Each rule selects a single consequent class. This type of rule has been used in many classification problems (Abe & Thawonmas, 1997).

Case 2: Fuzzy rules with certainty degrees and only a class in the consequent.

Many studies have shown the important role of the certainty degree in increasing both the interpretability and performance of fuzzy rule-based classification systems (Ishibuchi & Nakashima, 2001; Mansoor & Eghbal, 2007; Mansoori et al., 2007; Nauck & Kruse, 1998). It was demonstrated that weighting rules can be an alternative way to modifying MFs. In other words, the classifier can be tuned without altering the position of the fuzzy sets given by experts, but simply by adjusting the certainty degrees of rules. Since the latter is a single real number, its adjustment is much easier and simpler than changing the parameters of MFs. Furthermore, this approach maintains the MFs of given linguistic values as determined by experts, and thus the comprehensibility of the classifier is not deteriorated. Additionally, the necessity of using certainty grades becomes apparent when it comes to handling a mixture of general and specific rules (Ishibuchi & Nakashima, 2001). General rules can be thought of as rules with only few antecedent conditions, while specific rules could have many antecedent conditions and used for describing complicated classifications boundaries especially in the overlapped region (Ishibuchi & Nakashima, 2001). Several other studies (e.g., Ishibuchi and Nakashima (2001), Kuncheva (2000)) illustrated that the classification boundaries can be of different shapes when the certainty degree is employed. Conversely, the decision areas created by rules without certainty degree could only be rectangular or hyperrectangular unless fuzzy rule base is incomplete (there are missing rules in the rule base) (Ishibuchi & Nakashima, 2001). Such observation demonstrated the necessity of using the certainty degree for handling complex classification issues.

Case 3: Fuzzy rules with certainty degrees and all classes in the consequent.

In this case, a single rule can select multiple consequent classes with different certainty degrees. This type of rule presents a general model that extends the first and second types. We use this type of rule because we believe that it plays a crucial role in helping the classifier decide which class is the most appropriate, especially when the event falls in the overlapping region. That is, fuzzy rules usually present overlapping regions between neighboring sets, and hence it is highly probable that there will be some conflicting rules which possess the same IF part, yet select different classes. Such rules can be combined in one rule which simultaneously select multiple classes with distinct grades. In this way not only the conflict problem is resolved, but also the number of rules is reduced.

3.2.2.2. Generating fuzzy rules. In this study, fuzzy rules are derived from both expert knowledge and training data. This is because each of the two kinds of information alone could be incomplete. Although the classification is successfully performed by human analysts, the latter could fail to provide and express all their accumulated experience in the form of linguistic rules. On the other hand, the information learnt from training data is also mostly not enough alone for a successful classification. This is because the training data set is limited and could not cover all situations that the classifier can face. Furthermore, the generalization capability highly depends on the training data quality. Therefore, we believe that combination of these two kinds of information could dramatically enhance the generalization capability and performance of the classifier. The key idea behind this combination is to exploit the human analyst knowledge to enhance the generalization capability of the classifier, and at the same time generate some fuzzy rules (especially type 2 and 3) from data to improve the ability of the system to classify events that present overlap region in the feature space.

A. Analyst knowledge

As shown in Fig. 6, analyst knowledge intervenes in many stages. First, it is used to partition fuzzy space for input and output. Secondly, analyst knowledge is expressed in the form of linguistic if-then rules of type 1 that states in what situations a specific class should be chosen (a typical subset of rules is listed in Table 1). Finally, analysts interpret the derived rules from data to remove inappropriate rules and construct the final knowledge base.

B. Generating fuzzy rules from data

To lean fuzzy rules from data, we have used an extension of the heuristic technique proposed firstly by Wang and Mendel (1992) and then used by Chi, Wu, and Yan (1995), Chi, Yan, and Pham (1996) as well as Cordón et al. (1999).

Suppose we are given a training data set with the following structure:
The steps to classify an incoming seismic event of rules is determined by the ratio \( \frac{q}{1} \) analyst knowledge. Different for the different input variables. A membership function is identified training data set, which describes the mapping feature space \((\text{LE, DE, QB, NS})\) (input of the classifier) and the class set \((\text{LE, DE, QB, NS})\) (output of the classifier).

\[ F : (e_i, t_d, h, s_l, s_h) \rightarrow (\text{LE, DE, QB, NS}) \]

The learning algorithm consists of the following steps:

**Step 1: Partition the input space and output space into fuzzy regions.**

In this step, the domain interval (universe of discourse) of each feature variable is divided into \( n \) fuzzy regions (sets) \((n \text{ can be different for the different input variables)}\). A membership function is then assigned to each region. This step is performed based on the analyst knowledge.

**Step 2: Generate fuzzy rules**

First, for each event \( E^q \) in the training data set, determine the membership degrees of its features and class in the fuzzy sets determined in the previous step. Secondly, assign each feature to the label (linguistic term) of the fuzzy region where the maximum membership degree is achieved. Finally, combine the linguistic terms associated with the selected fuzzy regions by AND operators to form the IF-part of the fuzzy rule. The THEN-part is then determined by the class \( C^l \). In this manner, each seismic event in the training data set produces one fuzzy rule.

**Step 3: Revise the determined rules**

All generated rules are revised by an experienced seismic analyst. This revision aims to remove redundant and inappropriate rules as well as find the conflicting rules.

**Step 4: Determine the certainty degree \( \omega_{ij} \) of rules**

As was discussed above, rule weight can be used as a simple mechanism to improve the classification performance of the constructed rule-base. Many studies have been devoted to the certainty degree evaluation and improvement (Ishibuchi & Yamamoto, 2005; Mansoor & Eghbal, 2007; Mansoori et al., 2007). In this study, a frequently used heuristic method was employed (Ishibuchi & Yamamoto, 2005). Let us assume that \( Q \) labeled events \( E^q \), \( q = 1, 2, \ldots, Q \), are given from the four classes. The certainty degree for the rule \( R_i \) to select the class \( C_j \) is determined by the ratio between the sum of the matching degrees \( z_q \) for the events that belong to class \( C_j \) and the sum of the matching degrees for all the events according to the rule \( R_i \):

\[ \omega_{ij} = \frac{\sum_{q=1}^{Q} z_q(E^q)}{\sum_{q=1}^{Q} z_q(E^q)} \]

Examples of generated rules are presented in Table 2.

### 3.2.2.3. Membership function

Determination of the right membership function either in term of shape or boundary is one of the most critical first steps in constructing fuzzy systems. This function has a clear influence on the classification result since it determines the correspondence between the input data and linguistic terms, upon further fuzzy inference is done. The membership functions employed in this study to partition the feature space (step 1 in the previous section ‘Generating fuzzy rules’) are defined by intuition and deep studying as well as qualitative analysis of the input variables and output classes. The statistical measures estimation of each input parameter for each class population, including mean, standard deviation and minimum and maximum values also help determine the membership functions boundaries.

In this study, trapezoidal function was used, which mathematically defined by:

\[ f(x; a, b, c, d) = \max \left( \min \left( \frac{x - a}{b - a}, 1, \frac{d - x}{d - c} \right), 0 \right) \]

where the constants \( a, b, c \) and \( d \) determine the \( x \) coordinates of the four corners of the function. The final MFs for each input variable and output class are depicted in Figs. 7 and 8.

### 3.2.3. Fuzzy reasoning model

#### 3.2.3.1. Algorithm

The steps to classify an incoming seismic event using the proposed fuzzy reasoning model are as follows:

1. Satisfaction degrees of the clauses \((x_0^k \in A_{0n})\) \( n = 1, \ldots, 6 \).
The degree to which each variable input \( x_i \) belongs to the fuzzy set \( A_{\text{th}} \) is determined by its membership function:

\[
\mu_{A_{\text{th}}}(x_i^2)
\]

### 2- Matching degree

The strength of activation of the \( if\)-part of the rule \( R_i \) with the event \( E^j \) is defined by:

\[
\alpha_i(E^j) = \min \left( \mu_{A_{\text{th}}}(e_i^2), \mu_{A_{\text{th}}}(t_i^2), \mu_{A_{\text{th}}}(h_i^2), \mu_{A_{\text{th}}}(s_i^2), \mu_{A_{\text{th}}}(e_i^8), \mu_{A_{\text{th}}}(s_i^8) \right)
\]

### 3- Association degree

The association degree of the event \( E^j \) with the four classes according to the rule \( R_i \) is evaluated by:

\[
a_{ij} = \alpha_i(E^j) \cdot \omega_{ij}, \quad j = 1, \ldots, 4
\]

For each class \( C_j \), \( j = 1, \ldots, 4 \), calculate:

### 4- Aggregation

All rules consequents must be combined in some manner in order to make a decision. Such process can be performed by an aggregation operator \( g \) as:

\[
b_j = g(a_j), \quad i = 1, \ldots, L
\]

The aggregation operator is discussed in the following section.

### 5- Decision making

The final decision is made as follows: For each event \( E^j \), we primary determine the first and second winner class label \( h \) and \( k \) corresponding to the maximum and second highest value of the aggregation. This is achieved by the following equations:

\[
C_h = F(b_j) = \max(b_j), \quad i = 1, \ldots, M
\]

\[
C_k = F(b_j) = \max(b_j), \quad k \neq h
\]

Then, we compute the difference \( \Delta \) between the two classes:

\[
\Delta = C_h - C_k
\]

This gives the quality of decision; the higher \( \Delta \), the better the quality is. \( \Delta = 0 \) corresponds to an indeterminate class. In this case, we cannot decide which class is the most certain: the rule base has not enough knowledge to discriminate between the two classes.

#### 3.2.3.2. Aggregation function

Depending on the aggregation function \( g \) used in the step 4, different reasoning model can be designed. The most popular and classical fuzzy reasoning method employed in the fuzzy rule-based classification systems is the single winner method. In these systems, the maximum operator is employed as an aggregation operator (Abe & Thawonmas, 1997; Ishibuchi & Nakashima, 2001). Using this inference procedure, the final decision is made by the single winner rule corresponding to the highest association degree. Working in this way, the information associated with all other rules is unexploited as their association degrees with the input event are more or less inferior than the association degree of the selected rule. Thus, the aggregation value of each class \( C_j \) is determined as follows:

\[
b_j = g(a_j) = \max(a_j), \quad i = 1, \ldots, L
\]

The event is classified in the class \( C_j \) with the maximum aggregation value.

As decision making based on the fuzzy reasoning model that takes into account the information provided by a set of rules instead of only one may improve the performance of the classifier, it is of great importance to find out an alternative aggregation function that would allow all rules fired by an input event cast their votes and participate in the classification process. A wide variety of aggregation operators have been introduced and improved during the few past years (Calvo, Kolesarova, Komornikova, & Mesiar, 2002; Calvo & Mesia, 2003; Chiclana, Herrera, Herrera, & Alonso, 2007; Pelaez & Dona, 2003; Salido & Murakami, 2003). The most common goals among these aggregators are to provide a way of implementing operators that satisfy many criteria and have the property of lying between the two extremes Min and Max. Many of these operators were also developed in such a way to be generalized and parameterized so that they can be adjusted to become more like 'or' (tend toward maximum) and 'and' (tend toward minimum), as well as other popular operators. Among these aggregators, the ordered weighted averaging (OWA) operators have attracted much interest among researchers (Pelaez & Dona, 2003; Salido & Murakami, 2003). These operators have shown to be suitable for many diverse types of aggregation problems. The definition of the OWA aggregators is similar to the definition of the weighted mean of a set of ordered elements (in a decreasing way). It could be defined for the class \( C_j \) by:

\[
g(a_{ij}, \ldots, a_{kj}) = \sum_{i=1}^{L} \theta_i \cdot a_{ij}^{(i)}
\]

where \( \theta = (\theta_1, \ldots, \theta_L) \) is a weighting vector verifying \( \theta_i \in [0, 1] \) and \( \sum_{i=1}^{L} \theta_i = 1 \); \( \sigma \) is a permutation in which elements are sorted in a decreasing way, such that \( a_{ij}^{(i)} \geq a_{ij}^{(i+1)}, \quad \forall i = 1, \ldots, L - 1 \), i.e., \( a_{ij}^{(i)} \) is the \( i \)th highest value and \( a_{ij}^{(l)} \) is the maximum value in the set \( \{a_{ij}, \ldots, a_{kj}\} \). According to the above definition, OWA operators provide a general class of aggregators which include min, max and average, depending on their parameters values. In addition to their important characteristic which is the use of weights in the values to be aggregated, they satisfy many interesting properties such as monotonicity and continuity.

---

**Table 2**

Examples of rules extracted from data.

<table>
<thead>
<tr>
<th>Fuzzy rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>IF ( t_d ) is long and ( s_i ) is medium and ( s_l ) is low and ( e_i ) is non Gaussian and ( s ) is low THEN class is ( LE ) with 0.018 and ( RE ) with 0.982</td>
</tr>
<tr>
<td>IF ( t_s ) is medium and ( s_i ) is medium and ( s_i ) is medium and ( e_i ) is Gaussian and ( s ) is low and ( h ) is evening THEN class is ( QB ) with 0.821 and ( LE ) with 0.179</td>
</tr>
<tr>
<td>IF ( t_d ) is medium and ( s_i ) is medium and ( s_i ) is medium and ( e_i ) is Gaussian and ( s ) is low and ( h ) is noon THEN class is ( QB ) with 0.821 and ( LE ) with 0.982</td>
</tr>
</tbody>
</table>

**Fig. 7.** Membership functions of the variable class.
A relevant issue in using OWA operators is the determination of the associated weighting vector. In this paper we use two different weighting vectors. We first use the so called BADD (Basic Defuzzification Distribution) OWA (Chiclana et al., 2007; Pelaez & Dona, 2003), where the weights are function of the aggregates and given by the following equation:

$$\hat{\theta}_i = \frac{\sum_{j=1}^n \bar{a}_{ij}}{\sum_{j=1}^n \bar{a}_{ij}}, \quad p \in R$$

In this case, no ordering is necessary. Thus, the OWA aggregator can be rewritten as the following:
In this way, the decision is made neither according to only one rule, nor according to all rules, but based on the subset of rules that are more compatible with the event to classify. Such action is more comparable to the reasoning of seismic analysts, who never makes a final decision based only on one rule, but based on the set of significant rules.

### 4. Experimental results and discussion

In this section, the performance of the FRBES for seismic signal classification is examined. To do so, a data set of 343 seismograms was chosen and classified by seismic analysts. These seismograms are distributed among the four classes as depicted in Table 3. To assess the accuracy and generalization capability of the system, we used 2-fold cross-validation. In this way, the entire data set is divided into two subsets of approximately the same size. In each iteration of the cross-validation, one subset is used as training data for generating fuzzy rules, and the other subset is used as test data for evaluating the system. The classification result of each test is summarized in a confusion matrix that shows the number of correctly classified and misclassified events of each class. The output of the cross-validation is the sum of the two confusion matrices. Thus, it indicates the final performance of the classifier on the entire data set. Table 4 summarizes the classification results (average accuracy; correctly classified signals/total) of the system for the two fuzzy reasoning methods (classical fuzzy reasoning method and vote by multiple rule fuzzy reasoning method), and for the three discussed fuzzy rules structures.

Fig. 9 illustrates in more details the classification results obtained in case of using vote by multiple rule reasoning method with rule type 3 or 2, and in case of using classical reasoning method with rule type 1. Each confusion matrix shows in the rows the target classes and in the columns the predicted classes. The diagonal indicates agreement and the other cells indicate the misclassified events.

From a preliminary look at Table 4, it can be seen that the classifier shows generally good classification results on the seismic data. Such results reveal that the classifier performance is influenced by both the rule structure and aggregator type.

Based on the comparison of the classification rates, we could point out that:

- The best classification results (the bold values) are achieved when we use the second and the third fuzzy rule structure together with vote by multiple rule fuzzy reasoning method that operates either on the arithmetic mean or OWA operators.

#### Table 3

Distribution of seismograms in the four classes.

<table>
<thead>
<tr>
<th>Class</th>
<th>LE</th>
<th>NS</th>
<th>QB</th>
<th>DE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of event</td>
<td>83</td>
<td>115</td>
<td>100</td>
<td>45</td>
</tr>
</tbody>
</table>

#### Table 4

Average classification results (%).

<table>
<thead>
<tr>
<th>Rule type</th>
<th>Max</th>
<th>Arithmetic mean (p=0)</th>
<th>Badd operator (p=4)</th>
<th>OWA</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>92.41</td>
<td>93.52</td>
<td>93.29</td>
<td>93.52</td>
</tr>
<tr>
<td>(2)</td>
<td>95.32</td>
<td>97.95</td>
<td>94.74</td>
<td>97.95</td>
</tr>
<tr>
<td>(3)</td>
<td>95.32</td>
<td>97.95</td>
<td>94.74</td>
<td>97.95</td>
</tr>
</tbody>
</table>

\[
g(a_{ij}, \ldots, a_{kl}) = \frac{\sum_{i=1}^{n} a_{ij}^{p}}{\sum_{i=1}^{n} y_{i}}
\]

If \( p = 0, \quad g(a_{ij}, \ldots, a_{kl}) = \frac{\sum_{i=1}^{n} a_{ij}}{L}
\]

This corresponds to the arithmetic mean. In this case each rule casts a vote for a single class (rule types 1 & 2) or to all classes (rule type 3). The strength of the vote is defined by the association degree. The total strength for each class is determined by averaging all the given association degrees for each class. In this way, the reasoning process takes into consideration the information given by all rules even those presenting very low association degrees. Such behavior, overemphasizes the outcome of the rules presenting low association degrees at the expense of those having higher degrees, can lead the system towards a mistaken classification, especially, when most of rules present low association degrees (Cordon, Jesus, & Herrera, 1998). In order to overcome this drawback, an intelligent aggregator should select only the subset of rules that present significant association degree. This selection can also be carried out by means of the ordinary OWA operators where, unlike the BADD OWA operator, the aggregated elements must be ordered. Indeed, a fundamental aspect of the ordinary OWA operators is the reordering of the values to be aggregated based upon their magnitude. Such a characteristic can be used to rank the association degrees. Furthermore, due to the use of weights, the OWA aggregator can take into consideration only the first \( k \) significant rules on the list. In this manner, the weighting vector is static and can be evaluated as follows:

\[
\theta_i = \begin{cases} \frac{1}{k}, & i \leq k \\ 0, & i > k \end{cases}
\]

\[
\text{Classification results for vote by multiple rule fuzzy reasoning method with rule type 3 or 2.}
\]

\[
\text{Classification results for classical reasoning method with rule type 1.}
\]
- Rule type 2 generally improves the classification results from rule type 1. In fact, fuzzy rules of type 1 are not preferable for seismic classification. This is due to the fact that these rules solely select a single consequent class, and all rules are assumed of equal importance, whereas, in reality, seismic classes are not perfectly separable, but present overlapping regions which need to be interpreted as a matter of degree. Therefore, using weights offers an opportunity to control the strength of each rule in choosing the final class.

- Rule type 3 combines conflicting rules in one rule of multiple consequent classes, hence reduces the rule base and makes the classifier more transparent.

- The voting reasoning method based on the arithmetic mean or OWA operators shows generally good behavior for the three types of rule, it generally demonstrates enhancement of the classification results. Conversely, the classical fuzzy reasoning method achieves less performance on seismic data as it relies solely on the outcome of a single fuzzy rule, while multiple rules can be fired by an input pattern.

- Badd operator \((p = 4)\) achieves approximately comparable classification rates with the classical method \((\text{Max})\) for the three types of rule. This is due to the fact that this operator is becoming more like Max aggregation operator when we increase the value of the parameter \(p\).

To estimate separately the classifier’s performance on each class, other measures are employed (Gu, Zhu, & Cai, 2009; Sokolova & Lapalme, 2009). These include: sensitivity, specificity, accuracy, error and precision. Sensitivity describes the classifier ability to correctly identify signals belonging to a particular class. Specificity characterizes the ability of the classifier to recognize signals that are not of a particular class. Precision represents the capability of the classifier to not include signals of other classes in the considered class. Table 5 shows the performance evaluation.
of the classifier for the four classes when vote by multiple rule fuzzy reasoning method with rule type 3 or 2 is used. These measures were extracted from the confusion matrix presented on the left in Fig. 9. The results show that the system is more accurate for DE and QB events. The sensitivity of the classifier for the classes LE and NS is better than its sensitivity.

In order to better understand the two reasoning methods and the effect of rule weights on the classification results, we examined in more details their behavior on the misclassified events by the classical method when only rules without weights (type 1) are employed (see Fig. 9). We should point out that these events generally lie in the overlapping regions in the feature space. Table 6 illustrates a simple comparison between the classical method and the voting method (using OWA) on two misclassified QB events (Fig. 10) when the classical method with rule type 1 is used. In this comparison, rules of type 3 are employed (Table 7).

The two events belong to QB–LE overlapping region. Such a region may hardly be interpreted by only a single rule (winner rule). This is why the classical method fails to identify the two events. From the Table 6, we can see that event (a) is properly classified by the classical method when weights are applied on the rules. This result reveals the importance of using rule weight in such a situation, as it gives the significance of the rule in contributing to select the final class. One can see that applying the voting reasoning method improves the quality of classification.

Event (b) shows how the classical method can lead to a mistaken classification result based solely on the outcome of the single rule \( R_k \). In this case, the relevant information provided by all other fuzzy rules is not exploited. From these results, we can conclude that fuzzy system based on the voting reasoning method with rule type 3 reflects reality much better. It describes input attributes and output classes more intuitively using overlapping fuzzy subsets with certainties degree and approximate reasoning. In addition, it combines the information provided by various fired rules. Events that belong to the overlapping regions are treated in all classes using the degree of matching. On the contrary, the reasoning method that considers solely the winner rule or rules without weights fails to interpret the information provided by the regions of overlapping fuzzy subsets.

5. Conclusions

In this study, a fuzzy rule based expert system was successfully utilized to serve as a classification system of seismic events. The system was applied to the local seismic network of Agadir, however it can be tuned for other seismic networks. This system provides an appropriate nonlinear tool to manipulate the uncertainty and complexity of real data. More importantly, this system presents several interesting advantages over many other methods. For example, fuzzy rules and specially rules type 3 employed in this study make the classifier more transparent and human-comprehensible as well as adjustable. In addition, the vote by multiple fuzzy rule-based reasoning process used by the system seems more like human reasoning and permits separation of overlapping regions. Moreover, the incorporation of rules derived from both training data and expert knowledge offers the possibility of exploiting information obtained from diverse sources, hence expanding the generalization capabilities of the classifier. The relevant discriminant features are obtained from vertical seismogram component and are very simple to compute. The classification results achieved showed that the fuzzy expert system generally achieves high performance with low complexity and it could be operated as an online classifier. The main disadvantage of this approach is the fact that FRBES is a multi-parameter optimization problem.

Further research should be conducted in multiple directions. Firstly, we believe that performing further exploratory analysis of the seismic signals, using only vertical seismogram component or the three components, in order to identify more relevant features deserves more investigation. Secondly, the system described in this paper was built to classify among four classes, considered that only these four types of events are recorded here. However, it can be easily extended to identify signals of more than four classes (e.g., nuclear explosions, landslides, and volcanoes). Thirdly, the excellent performance achieved on seismic signal classification encourages applying fuzzy expert systems to other problems in seismology such as signal detection. More generally, this technique can be extended to other type of signal classification problems such as speech recognition and EEG signal classification.

Acknowledgments

The authors are very grateful to the editors and anonymous reviewers for their deep and thorough review as well as for their thoughtful comments and constructive suggestions, which help to improve the quality of this paper.

The first author received a Moroccan PhD excellence grant (CNRT).

References


