Being accepted as a suitable Non-destructive Testing (NDT) method and monitoring technology by many engineers and technicians in different industries, Acoustic Emission Technique (AET) is not employed massively yet. The main reason is that AET is a relatively a new method of monitoring compared with already mature methods.

AE waves can be represented as very small amounts of energy, released suddenly by a crack or defect inside a material. Normally, the sources of AE waves are formed randomly and therefore AE signals are not reproducible. Since AE waves contain very small amount of energy, they require very sensitive sensors and cutting edge amplifiers to face the low signal-to-noise ratio. In this context, advanced signal processing is required to recognize the AE signals. In addition, the sources of acoustic emission waves, mode of travel of elastic waves across a material and how data acquisition equipment records these signals are matters which require adequate understanding in order to capture and interpret the AE data correctly.

Automatic signal recognition and classification of AE signals for machinery under study in order to identify different phenomena of interest is a matter of great importance for the maintenance engineers, and research engineers who spend much time processing the AE data.

Signals that cannot be described with a time-domain equation are called noise signals. However, this kind of signals can be quantified in a valuable way using the statistical measures of random sequences. These statistical indicators can be used as signal analysis tools.

The main objective of this research work is to contribute on the automatic decision making algorithm for the AE signals detected by AE sensors. For this purpose, different unsupervised and supervised classification algorithms are developed and their accuracy evaluated. AE signal classification by classifying algorithms using Pattern Recognition Technique (PRT) is discussed for mining applications.

Also a general approach to employ AET in combination with PRT is developed and described. The developed tool is used for equipment monitoring, process supervision in the industries like mining and also for processes involving AE phenomena.

Kenny Alexander Gárate Peñaranda • Pattern Recognition, Classification and Diagnosis of Acoustic Emission Signals in Applications for Mining

# Pattern Recognition, Classification and Diagnosis of Acoustic Emission Signals in Applications for Mining

**Kenny Alexander Gárate Peñaranda**

# Pattern Recognition, Classification and Diagnosis of Acoustic Emission Signals in Applications for Mining

From the Faculty of Georesources and Materials Engineering of the

RWTH Aachen University

Submitted by

**Kenny Alexander Gárate Peñaranda**

from Arica, Chile

in respect of the academic degree of

**Doctor of Engineering**

approved thesis

Advisors:

Univ.-Prof. Dr.-Ing. Karl Nienhaus

Univ.-Prof. Dr.-Ing. Thomas Pretz

# Abstract

In this work, Acoustic Emission Technique (AET) and Pattern Recognition Technique (PRT) are used in combination for classification of AE signals obtained from two experimental applications in the mining sector. AE signals are collected and used to form features or parameters in order to identify AE events of interest. Signal processing in time domain and in frequency domain are applied to extract these features from the AE signals.

The aim of this work is to identify and classify similar AE signals from experimental mining tests. It is achieved by means of using unsupervised and supervised methods of PRT. The main reason for performing these tests at a laboratory scale is to obtain a correlation of the AE signals with each mining process as well as to use the results as a guide for other specific mining applications.

Beginning with the formulation of the problem and data acquisition, each step of the pattern recognition process is carried out for classification of the AE signals, namely pre-processing of the data, feature selection and extraction, unsupervised or supervised classification, assessment and interpretation of the results.

In the pre-processing of the data, *Standard Score Normalization* is applied to the input data for further comparisons of AE features. For feature extraction, three algorithms are used, namely *Principal Components Analysis (PCA)*, *Linear Discrimination analysis (LDA)*, and *Multidimensional Scaling (MDS)*.

For *supervised classification* three methods are employed, these are *linear Support Vector Machine*, *non-linear Support Vector Machine* and *Back-propagation Neural Network*. These supervised classifiers are evaluated using *classification accuracy*.

For *unsupervised classification*, also three classification algorithms are used, namely *K-means Clustering*, *Fuzzy C-Means Clustering* and *Vector Quantization Clustering*. These unsupervised classifiers are evaluated using similarities between clusters by *Rand Index (RI)*.

Two experimental applications are studied using the algorithms and the selected AE features. First, AET in a laboratory column flotation cell is used to monitor the bubble activity and bubble size, as a means of improving the efficiency of the column flotation process. Next, in rock cutting, AET and PRT are employed to identify the rocks being cut as a means of automation in the operation of underground mining and tunneling.

These experiments were carried out in a laboratory scale, using AE wave measurements to analyze the processes. Pattern Recognition Technique is used in combination with classic and advanced signal processing techniques to characterize the collected AE signals.

# Zusammenfassung

In dieser Arbeit wird eine Mustererkennung (Pattern Recognition Technique PRT) für die Klassifikation von Acoustic Emission (AE) Signalen, die in Bergbauprozessen entstehen, vorgenommen. In zwei experimentellen Aufbauten werden die AE Signale durch geeignete Messtechnik erfasst. Anschließend werden Parameter und Merkmale in diesen Signalen gesucht, die als wiederkehrendes Muster zu erkennen sind. Für die Bestimmung dieser Parameter kann die Signalverarbeitung sowohl im Zeit- als auch im Frequenzbereich durchgeführt werden.

Ziel dieser Arbeit ist zunächst die Identifizierung von Merkmalen mit einer sich anschließenden Klassifizierung der AE Signale. Die Klassifizierung der Signale erfolgt sowohl durch eine überwachte als auch eine nicht überwachte Mustererkennung (PRT). Hierzu werden unterschiedliche Versuche durchgeführt, um Korrelation zwischen AE Signalen und Prozessabläufen im Bergbau herzustellen. Die Versuche finden zunächst im Labormaßstab statt, die Rückschlüsse auf reale Bergbauprozesse zulassen.

Im Verlauf dieser Arbeit wird zunächst eine Problemstellung formuliert. Anschließend werden die AE Daten in den Versuchen erfasst und die Mustererkennung für die Klassifikation der AE Signale durchgeführt. Die einzelnen Schritte hierbei sind die Datenvorverarbeitung, die Auswahl geeigneter Kennwerte, ein überwachte sowie nicht überwachte Klassifikation gefolgt von der Interpretation und Beurteilung der Ergebnisse.

Für die Vorverarbeitung der Daten werden die Daten zunächst mit Hilfe der Standard Score Normalization normalisiert. Die sich anschließende Auswahl der Kennwerte erfolgt nach der jeweiligen Anwendung drei unterschiedlicher Algorithmen, der *Principal Components Analysis (PCA)*, *Linear Discrimination analysis (LDA)* und *Multidimensional Scaling (MDS)*.

Die ausgewählten Kennwerte werden durch sechs verschiedene Methoden klassifiziert und auf wiederkehrende Muster untersucht.

Für die *überwachte Klassifikation* werden die Methoden linear *Support Vector Machine*, *nonlinear Support Vector Machine* and *Back-propagation Neural Network verwendet*. Die überwachten Cluster werden durch den *classification accuracy* untersucht.

Die *unüberwachte Klassifikation* erfolgt durch *K-means Clustering*, *Fuzzy C-Means Clustering* und *Vector Quantization Clustering*. Anschließend werden die Ergebnisse auf ihre Sinnhaftigkeit und Genauigkeit beurteilt. Durch den *Rand Index (RI)* lassen sich Gemeinsamkeiten innerhalb der gebildeten Cluster bestimmen.

Insgesamt werden zwei unterschiedliche Berbauprozesse betrachtet, die AE Signale während des jeweiligen Prozesses erfasst, und die Algorithmen auf die ausgewählten Kennwerte angewendet. In einem ersten Prozess wurden die AE Signale an einer Flotationszelle aufgezeichnet, um eine Aussage über die Blasenbildung und die Blasengröße innerhalb der Zelle treffen zu können. In einem zweiten Prozess wurde die AE Technologie für die Materialidentifikation während Schneidprozessen in der schneidenden Rohstoffgewinnung verwendet.

Die Versuche wurden im Labormaßstab durchgeführt und die AE Signale zur Beurteilung und Analyse des Prozesses verwendet. Die Mustererkennung (PRT) wurde in Kombination mit konventionellen und fortgeschrittenem Signalverarbeitungstechnologie angewendet, um die erfassten AE Signale zu charakterisieren.

# Acknowledgments

# List of symbols

The notation used in this work is given below.

$A$      Signal Amplitude

$c_1$      longitudinal velocity of an elastic wave

$c_2$      Transversal velocity of an elastic wave

$C$      number of classes

$E$      Young s modulus

$\mathcal{F}$      Fourier transformation

$f$      Frequency

$f_i$      The body force per unit mass

$K$      Bulk modulus

$N^o$      Is a random variable that represents the number of counts and can take the values $N_c^o$ related to the counts from AE coal signals and $N_d^o$ related to the counts from AE dead rock signals

$p$      Number of features

$Q$      Value that qualify the attenuation of elastic waves

$\mathbf{R}$      Is a random variable that represents the cut of a rock and can take the values $\mathbf{C}$ for coal and $\mathbf{D}$ for dead rock

$T_s$      Sampling rate

$\nu$      Poisson s ratio

$\mathbf{w}$      Is a weight vector

$w_0$      Radian frequency

$\boldsymbol{x}$      $= (x_1, ..., x_p)^T$ represents a pattern and is a vector containing $p$ features in a $d$-dimensional space

$\boldsymbol{X}$      $= (x_1, ..., x_n)^T$ $n \times p$ data matrix

$z$      Is a categorical variable that associate a pattern $\boldsymbol{x}$ with a class $\omega$. if $= i$ , the pattern pertains to $\omega_i$, with $i \ \epsilon \ \{1, ..., C \ \}$

$\Omega_i$      Are the regions in a $d$-dimensional space, separated by a decision rule, with $i \ \epsilon \ \{1, ..., C \ \}$

$\omega$      Refers to a class of an AE event, the researcher should determine $C$ classes before the design of the classifier

$\tau_{ij}$      Stress tensor

| | |
|---|---|
| $u_i$ | Displacement vector |
| $\mu$ | Lame constant |
| $\lambda$ | Elastic constant |
| $\rho$ | Mass density per unit of volume of the material |
| $\varepsilon_{ij}$ | Strain tensor |
| $\omega_{ij}$ | Rotation tensor |
| $()^T$ | Indicates vector transpose. In this work the terms *feature vectors*, *Patterns* and *data set* are used indistinctly |
| $\Delta$ | Dilatation of a material |
| $\bigtriangledown$ | Vector differential operator |
| $\varnothing$ | Phase shift in radians of a cosine signal |

# Abbreviations

| | |
|---|---|
| AET | Acoustic Emission Testing |
| CDF | Cumulative Distribution Function |
| DAS | Data Acquisition System |
| FCM | Fuzzy-C Means |
| FFT | Fast Fourier Transformation |
| KKT | Karush-Kuhn-Tucker |
| LDA | Linear Discriminant Analysis |
| MLP | Multilayer Perceptron |
| NDT | Non-Destructive Testing |
| PCA | Principal Component Analysis |
| PDF | Probability Density Function |
| PRT | Pattern Recognition Technique |
| PZT | Lead Zirconate Titanate |
| RI | Rand Index |
| RMS | Root Mean Square |
| SPR | Statistical Pattern Recognition |
| SSN | Standard Score Normalization |

VQ     Vector Quantization

# Contents

# 1. Introduction

## 1.1. Problem description

Acoustic emission (AE) is produced in materials when stored strain energy is released. This is known as AE events or emitting elastic waves (AE waves). The elastic waves propagate across a material and can be detected by an AE sensor. AE sensors are directly attached to the surface of the monitored material (structure borne). The study and signal processing of these elastic waves is known as Acoustic Emission Technique (AET). This technique is considered as a non-destructive and passive technique.

Nowadays, modern AET faces the problem of requirement of fast signal processing and recording equipment as well as recording techniques to manage the huge amount of data acquired in an AE measuring campaign. In this context, the selection of features from these AE signals and recording of them for further processing becomes an important issue in AET. Furthermore, this procedure could also cause a misinterpretation of the AE signals being monitored if the personal is not well-trained.

AET has a variety of industrial application in different fields of engineering. In each of these fields this technique has to face different difficulties. For example; in the mining sector the main problem for the practical application of AET is the lack of automatic signal recognition and signal classification for monitoring of mining machinery.

Specifically, in mining machinery many studies have found correlations between processes, failures mechanisms and their AE features [AT00], [NKBB12].The interpretation of acoustic emission signals in time domain and frequency domain is possible by means of AET. This signal processing can be supported by complex classification algorithm from Pattern Recognition Technique (PRT).

All investigations using AET have problems with data volume, false interpretation of AE signatures, background and electrical noise as well as signal processing among others [CH08], [RAC01], [IKAK12]. In this context PRT is a way to resolve these drawbacks.

AET is related to AE events produced in every mining machine or structure as well as the condition monitoring of them. AE waves can be considered unique since they cannot be reproduced identically in an arbitrary manner. As a result, recording of AE signals cannot be repeated.

The monitoring process can be performed by the observation of AE changes of the process conditions and further interpretation of these observations. Having high skilled personal interpreting these AE signals can signify increase of the operational cost. Although, the interpretation of crucial AE events, it is logically a task for a human being. However, more the automatic interpretations of the results, fewer the cost for manual interpretation by maintenance engineers and researchers. Automatic analysis techniques are expected to detect AE waves that are related to critical AE events of interest in mining machinery with a high reliability. Therefore all relevant AE signals must be recorded and the right AE features, threshold values and signal classification must be performed in a satisfactory manner.

## 1.2. Motivation

Being accepted as a suitable Non-destructive Testing (NDT) method and monitoring technology by many engineers and technicians in different industries, Acoustic Emission Technique (AET) is not employed massively yet. The main reason is that AET is a relatively a new method of monitoring compared with already mature methods. The other reason being the lack of knowledge of signal processing method required to interpret the acoustic emission (AE) signals. In the early fifties of the 20th century began the initial research in AET with the focus on its applicability and suitability. Nowadays, with the enormous advancements in material technology, computer science, software and signals processing, the AET is much more complex to be interpreted by engineers and researchers.

AE waves can be represented as very small amounts of energy, released suddenly by a crack or defect inside a material. Normally, the sources of AE waves are formed randomly and therefore AE signals are not reproducible. Since AE waves contain very small amount of energy, they require very sensitive sensors and cutting edge amplifiers to face the low signal-to-noise ratio. In this context, advanced signal processing is required to recognize the AE signals. In addition, the sources of acoustic emission waves, mode of travel of elastic waves across a material and how data acquisition equipment records these signals are matters which require adequate understanding in order to capture and interpret the AE data correctly.

Automatic signal recognition and classification of AE signals for machinery under study in order to identify different phenomena of interest is a matter of great importance for the maintenance engineers, and research engineers who spend much time processing the AE data.

Pattern Recognition Technique (PRT) is a multidisciplinary collection of techniques, which requires knowledge of probability theory (because AE signals are generated in a stochastic manner), linear algebra (because AE phenomena are treated as vector) and mathematical methods. PRT is currently a dynamic research area and is a mixture of information science and statistics. In fields such as image processing, biometrics, neural computing and many others, PRT technique is at the core of these subjects [Bis06].

Differentiation and classification of AE signals can be achieved by means of PRT, which utilizes

classifiers based on features of different AE phenomena. Pattern recognition analysis of AE signatures can be used for classifying the AE sources mechanisms, revealing the causes of generation and predicting new events for monitoring a process.

## 1.3. Objectives

The main objective of this research work is to contribute on the automatic decision making algorithm for the AE signals detected by AE sensors. For this purpose, different unsupervised and supervised classification algorithms are going to be developed and their accuracy evaluated.

AE signal classification by classifying algorithms using Pattern Recognition Technique will be discussed for mining applications.

Also a general approach to employ AET in combination with PRT will be developed and described. The developed tool will be used for equipment monitoring, process supervision in the industries like mining and also for processes involving AE phenomena.

## 1.4. Structure of the thesis

In this work, the general theory of the acoustic emission technique (AET) is analyzed and compared with the real condition existing in the mining industry, remarking its advantages and limitations for every case. Pattern Recognition Technique is used to process AE data with the aim of classifying AE signatures in an easy and automatic way. Including this first introductory chapter, this work is divided into six parts.

**Chapter 2** reviews the general theory of acoustic emission technique, focusing on important information required to measure, process and interpret the AE data which normally leads to a good or poor employment of this technique. Afterwards, the chapter describes how an AE wave travels from its origin to the AE sensor as well as the conversion of analog waveforms into digital signals for further signal processing. The generation process of AE phenomena and its classification is presented.

**Chapter 3** deals with the signal processing of AE signals and AE signal features. Conventional signal parameterization and signal based analysis are reviewed and compared. Moreover, advanced signal processing analysis in the time domain, frequency domain as well as in time-frequency domain is described.

**Chapter 4** is dedicated entirely to review the Pattern Recognition Technique (PRT) and it explains how this technique can be used for resolving problems related to AE signal processing analyzed in the chapter 3. The objective of this chapter is to explain the PRT theory and link it to AE signals. In others words, the whole process of PRT is made intelligible by means of examples using AE signals. As the first step, what is PRT and its different methods are described and analyzed. Secondly, PRT is reviewed in depth from an AE point of view and

its different stages are studied with examples of AE signals, preparing the definition of pattern classes, feature extraction, clustering analysis, classifier design and selection of training set sample. Finally, the limitations of using PRT in AE signals are highlighted and what should be taken into account for the appropriate use of this technique with AE signals is discussed.

**Chapter 5** describes two experimental cases of study related to mining applications. First, the vast amounts of applications of AET studied till now in the mining industry are reviewed and some interesting points of AET are addressed.

Next, the problem of estimation of the bubble sizes in a flotation column by means of Pattern recognition technique as well as the problem of turbulence are addressed and detailed in depth. Vibration sensors and AE sensors are used for the data acquisition.

Afterwards, the experimental case of rock cutting is analyzed with the aim of identifying signals emitted from different rock when they are cut and picked up by an AE sensor. In others words it recognizes the type of rock being cut. For this purpose parameters are extracted from every signal and processed with PRT.

**Chapter 6** is the final part of this research work, where the conclusions of the signal processing and pattern recognition of AE signals from the different experimental cases are presented. This chapter also summarizes the complete research work. The subsection "outlook"gives some recommendations to continue working with PRT and AET for future investigations.

# 2. Acoustic Emission Technique

## 2.1. Introduction

Acoustic Emission Technique (AET) can be considered to be a passive Non-Destructive Test (NDT). This method has its origin in the early fifties with the doctoral thesis of J. Kaiser [Kai50] who found that a material under load emits acoustic emission (AE) waves when its primary load level is overtaken. The name "Acoustic Emission"was first given in English by B. H. Schofield in USA who published since 1954 several reports entitled "Acoustic Emission under Applied Stress"[SBK58]. A well detailed history of the AET can be reviewed in [Dro94]. The term "Acoustic"is related in formal English to sound, hearing or even to a musical instrument. However, the term is used in AET to refer to inaudible mechanical waves.

Since then, developments in AE instrumentation, reinforcement of the AE fundamentals, formation of formal AE working groups, increment of its use in industrial applications and above all huge developments of the computer sciences have contributed to the formation of a mature and developed technology which nowadays is being vastly studied and utilized in many industrial applications.

A structure can be assessed with AET without any additional load or stimulus and without causing permanent damage. AET is a very sensible NDT method capable of detecting small amounts of energy (acoustic emission phenomena) at meters of distance far away from the AE source. The International Organization for Standardization (ISO) defines acoustic emission as a class of phenomena whereby transient elastic waves are generated by the rapid release of energy from localized sources within a material, or transient waves so generated [ISO01]. The ideal elastic waves are mechanical disturbances that propagate through a material causing oscillations of the particles of that material about their equilibrium positions without other change [WF05]. However, in reality this does not occur and dissipation of the energy take place causing attenuation of the elastic wave with the distance traversed. Usually, due to the high frequency nature of the AE waves, they can propagate just a few tens of meters. Non-elastic waves cannot propagate throughout a material due to their high attenuation. Hence, they cannot be picked up by an AE sensor. Transient refers to the short duration of the AE emissions in a temporary way.

Being a passive method, AET can only record information about AE phenomena occurring during the measurements and no further information can be interpreted immediately. The great amount of possible AE source makes every AE signal unique and containing many char-

acteristics. In general two types of AE signals can be distinguished namely burst signals and continuous signals.

## 2.2. Overview of the AE technique

AE waves can be considered to be microseismic waves generated by a process of interest, normally during a failure process. AE can be defined as the spontaneous release of localized strain energy in stressed material [GO10]. Usually the frequencies range that AET deals with starts from 20 kHz up to several MHz.

Due to the compact size of the measurement equipment required to undertake AE tests, it is not laborious to measure in industrial environments as well as in controlled laboratory environments.

The process of measurement with an AE sensor is carried out with the help of different equipment. Nowadays, there is a variety of commercial AE measurement equipment available with different features. The general process of AET is showed in figure 2.1. When a mechanical wave starts its way to the AE sensor, it is subjected to different obstacles that cause it to attenuate and sometimes this mechanical wave changes its frequency (Absorption or dissipation)[OJ99].

In the **first zone** the mechanical wave is originated. AE sources are all types of rapid microscopic displacements inside or on the surface of a material. The wave propagates outward in a spherical way. This mechanical wave consists in different types of waves, which travel almost together across the material (almost because they do not have the same velocity) and their velocities depend on the composition of the material. From the energy point of view that is released from the AE source, only the elastic waves are capable of traversing the material to reach the AE sensor. This elastic wave moves backwards and forwards causing a transfer of energy or momentum from one point to another inside the material. The heterogeneity of the material is a very important issue for the wave propagation because it determines how much energy the wave loss [MLMC05]. The problem of wave propagation in non-heterogeneous material has been vastly studied and different theories and propagation models can be found in the literature, usually from a mathematical point of view, different scales and all of them with common points and contradictions [BSS13], [BBE09], [Mou09].

Since AE waves are vibrational motions of atoms about its equilibrium point in a material, they can be treated from a microscopic point of view by thermodynamics [Mou09], however normally AE waves are treated from a macroscopic point of view, supposing that the medium in which an AE wave travels is continuous.

Next, in **the second zone** if the AE wave manages to overcome the attenuation suffered inside the material, it gets to the surface and can be picked up by an AE sensor. Normally, a piezoelectric element (inside of a protective housing) is used for this task. A gel-based couplant should be used between the measurement surface and the AE sensor to remove any air from the interface and improving so the transmission up ten times for a 500 kHz AE wave [TZA08]
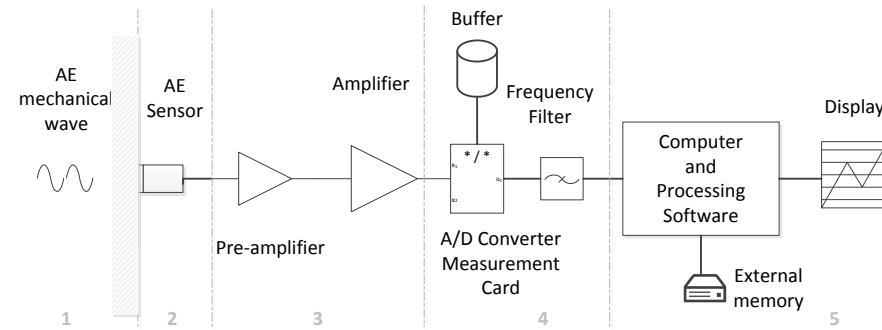
Figure 2.1.: Overview of the data acquisition equipment required for the AET.

(due to its low acoustic impedance). The most commonly used AE sensor are made of ceramic such as lead zirconate titanate (PZT) Pb $[Zr_xTi_{1-x}] O_3$ [SMTD01]. The piezoelectric effect is based on the occurrence of electric dipole moments within the piezoelectric material. Using this effect, AE sensors respond to dynamic motion on the surface of a material caused by an AE event, transforming mechanical motions into electrical signals. Usually the sensitivity based on voltage output per unit pressure input of an AE sensor is 1 V/mbar [GO10]. AE sensors can be classified into two categories according to the frequencies range in which they are more sensitive (frequency response). These types are *broadband AE sensors*, which respond uniformly to a very broad band of exciting frequencies and *resonant AE sensor* which are more sensitive at its own resonance frequency. Once an AE sensor has converted the micro surface motion into an electrical signal, a cable, usually coaxial is used to transmit the electrical signal to the pre-amplifier. This cable should not be more than two meters to avoid the internal electro-magnetic noise and attenuation.

**The third zone** corresponds to the pre-amplifier and amplifier. Some AE sensors have included in its structure a pre-amplifier circuit and only an external (main) amplifier is required. AE sensors with integral pre-amplifier are larger and more suitable for measurements on site because the spare cables and the measurement equipment configuration is easier to carry out. The amplifier usually has a configurable gain for different situations. The amplifiers enlarge an electrical signal and then send it to a measurement card. This device are supplied with AC power (typically). Normally, the gain of an amplifier is given in decibels (dB), a 40 dB gain is typical for many applications. A cable transmits the signal from the amplifier to the measurement card. This cable can be several meters long.

**The fourth zone** is where the analog signal is converted into a digital signal. In this zone the signal is filtered to remove the low and high frequency content. Usually, these measurement systems are modular and portable with only a few input channels or larger embedded systems

in a rugged chassis with multiples channels and more connectivity. In a ragged chassis-based system the measurement card is inside as a plug-in card. Additionally, this system can act as processing unit capable of processing the AE signals.

**The fifth zone** is where the signal processing takes place by means of the processing software to record and extract properties of the measurement. A graphical software displays signal features and records this features or the whole waveform. A processing software provides a graphical interface for data manipulation and generation of reports. Some common AE analysis methods are waveform analysis, FFT analysis, wavelet transform and feature extraction. Due to the large amount of data generated during a typical measurement, normally an external memory is required to record the data.

## 2.3. Advantages and disadvantages

Due to the non-invasive nature of AET, it can be applied in many mining systems and in operation without stopping it and under normal operation to gather information about the process and structure health. This technique is suitable for monitoring AE phenomena undergoing during the test, providing a real time monitoring of the process. The table 2.1 summarizes the main advantages and disadvantages of AET.

Table 2.1.: Advantages and disadvantages of AET.

| Advantages | Disadvantages |
|---|---|
| AE waves can be measured without stopping the process. On-line monitoring of machine components and mining systems is possible. External noise and mechanical noise have normally frequency outside of the range of AE waves. | Sometimes low rate of signal-to-noise ratio is obtained from a measurement. Consequently, AET is more suitable for AE waves with a burst form. Nevertheless, the use of frequency filter can help with this drawback. |
| Global monitoring of a structure is possible, requiring only a few sensors to monitor the process. AE waves propagate in all directions spherically. | Source location and AE phenomena recognition require advanced signal processing. |
| AE data acquisition equipment is easy to set up and AE sensors are simple to mount, requiring only superficial inspection. | No information about the shape of the AE source is given by AET. |
| AET is suitable for monitoring large structures of steel, concrete, rock and wood among others. | Usually, just the irreversible AE phenomena occurring in the plastic phase of a material gives valuable information. However, AET can give valuable information in the reversible phase in some processes. |
| Using adequate signal processing, a great amount of information can be obtained from an AE test. | Repeatability is not perfect, due to the random nature of AE sources; as a result, AE recordings are not exactly reproducible. AE is stress unique and each loading or source generator is different. |
| Real-time evaluation can be performed with AET. | A large amount of data needs to be managed and recorded. |
| Small releases of energy from a machine or process can be addressed by AET. Early-stage failure detection is possible. | Very sensitive and reliable data acquisition equipment is required. |
| AE data acquisition equipment has a good performance/price ratio. | Personal with knowledge in signal processing is required for interpreting the AE signal acquired. |
| Material anisotropy does not affect significantly the AE measurements. | The structure that transports the AE stress wave attenuates it. |
| Many types of machine elements and mining systems can be monitored since AET is not intrusive. | AE waves are sensitive to interfaces of different materials when they propagate across a machine. The AE waves suffer reflection and dispersion; as a result only part of the original wave reaches the AE sensor. |

## 2.4. Acoustic emission waves

One important use of elastic waves in ultrasound range is its use in nondestructive testing (NDT). AE waves are seen as mechanical waves containing small amounts of energy and generated by micro- and macroscopic sources. These elastic waves are mechanical disturbances that propagate through the space and time in a spherical form across a medium in which deformation (caused by the waves) is reinstated after the waves pass. This deformation of the medium results in an increment in the temperature (minor to be considered) and provokes attenuation of the waves.

Based on the shape (after analog-to-digital processing), AE waves can be classified in burst or continuous type as is showed in figure 2.2.

Burst types are discrete events of AE phenomena as rapid crack formation or rapid crack growth (microscopic sources) as well as impact or a mining wire rope break (macroscopic sources). In the same way, continuous type AE waves are produced by AE phenomena that lasts more in the time as friction inside a wire rope or internal deformation in a rock. Once, an AE phenomenon occurs this can only propagate through a material (structure borne wave). AE waves are inaudible waves with frequencies from 20 kHz to several megahertz. The higher the frequency of the wave the higher is its attenuation in the medium.

In solid and fluid mechanics, the medium is considered to be continuous, so that the properties such as density or elastic constants are considered to be uniform functions representing averages of microscopic quantities [Gra77]. In case of a continuous media, the mass and elastic parameters are distributed in terms of mass density and the elastic moduli. The disturbance spreads outward in a three dimensional form.

In a solid, an elastic wave performs three types of stress, namely tensile and compressive stress (transmitted to the particles of the media which move in the direction of the wave) and shear stress, where the motion of the particles is perpendicular to the direction of the propagation. In fluids, only the tensile and compressive stresses are possible in a wave.

The propagation of elastic waves in an infinite medium can be described by the elasticity equations. For a homogeneous isotropic elastic solid, the equation can be summarized in Cartesian tensor notation as

$$\tau_{ij,j} + \rho f_i = \rho \ddot{u}_i \tag{2.1}$$

$$\tau_{ij,j} = \lambda \varepsilon_{kk} \delta_{ij} + 2u\varepsilon_{ij}$$
$$\varepsilon_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i})$$
$$\omega_{ij} = \tfrac{1}{2}(u_{i,j} - u_{j,i})$$

where $\tau_{ij,j}$ is the stress tensor at a point and $u_i$ is the displacement vector of a material point. The stress tensor is symmetric ($\tau_{i,j} = \tau_{j,i}$). $\rho$ is the mass density per unit of volume of the
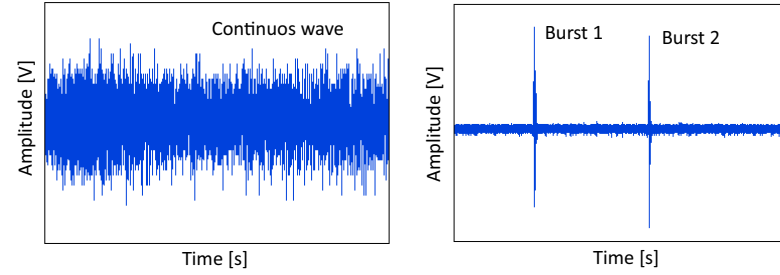
Figure 2.2.: Example of a continuous wave signal (left) and burst wave signal (right).

material and $f_i$ is the body force per unit mass of the material. $\varepsilon_{ij}$ is the strain tensor and $\omega_{ij}$ is the rotation tensor.$\lambda$ is the elastic constant (usually is the shear modulus) and $\mu$ is known as the Lamé constants for the material. $\lambda$ and $\mu$ may be expressed in terms of other elastic constants, such as Young´s modulus $E$, Poisson´s ratio $\nu$ and the Bulk modulus $K$. This representation is convenient under special loadings, such as simple tension or simple shear.

Youngs modulus $\qquad\qquad\qquad E = \mu(3\lambda + 2\mu)/(\lambda + \mu)$

Poissons ratio $\qquad\qquad\qquad \nu = \frac{1}{2}\lambda/(\lambda + \mu)$

Bulk modulus $\qquad\qquad\qquad K = \lambda + \frac{2}{3}\mu$

To obtain an equation in term of displacements we must substitute the expression for strain into the stress-strain relation.

$$(\lambda + \mu)u_{j,ji} + \mu u_{i,jj} + \rho f_i = \rho \ddot{u}_i$$

where $u_i$ is the displacement vector of a material point. The vector equivalent of this expression is

$$(\lambda + \mu) \bigtriangledown \bigtriangledown . \mathbf{u} + \mu \bigtriangledown^2 \mathbf{u} + \rho \mathbf{f} = \rho \ddot{\mathbf{u}} \qquad\qquad (2.2)$$

The equation 2.2 represents in terms of rectangular scalar notation [Gra77], the following equations

$$(\lambda + \mu) \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 \nu}{\partial x \partial y} + \frac{\partial^2 w}{\partial x \partial z} \right) + \mu \bigtriangledown^2 u + \rho f_x = \rho \frac{\partial^2 u}{\partial t^2}$$

$$(\lambda + \mu) \left( \frac{\partial^2 u}{\partial y \partial x} + \frac{\partial^2 \nu}{\partial y^2} + \frac{\partial^2 w}{\partial y \partial z} \right) + \mu \bigtriangledown^2 \nu + \rho f_y = \rho \frac{\partial^2 \nu}{\partial t^2}$$

$$(\lambda + \mu) \left( \frac{\partial^2 u}{\partial z \partial x} + \frac{\partial^2 \nu}{\partial z \partial y} + \frac{\partial^2 w}{\partial z^2} \right) + \mu \bigtriangledown^2 w + \rho f_z = \rho \frac{\partial^2 w}{\partial t^2}$$

where $u, \nu, w$ are the particle displacements in the $x, y, z$ directions. The dilatation of the material is defined by

$$\Delta = \bigtriangledown . \mathbf{u} = \varepsilon_x + \varepsilon_y + \varepsilon_z = \varepsilon_{kk}$$

And so the equation 2.2 can be written as

$$(\lambda + \mu) \bigtriangledown \Delta + \mu \bigtriangledown^2 \mathbf{u} + \rho \mathbf{f} = \rho \ddot{\mathbf{u}} \tag{2.3}$$

Wave equation for isotropic elastic solid. Now, considering the governing displacement without body forces

$$(\lambda + \mu) \bigtriangledown \bigtriangledown . \mathbf{u} + \mu \bigtriangledown^2 \mathbf{u} = \rho \ddot{\mathbf{u}} \tag{2.4}$$

Performing the vector operation in the above equation [Gra77], the wave equation is expressible in the form

$$\bigtriangledown^2 \Delta = \frac{1}{c_1^2} \frac{\partial^2 \triangle}{\partial t^2}$$

where the longitudinal propagation velocity $c_1$ is given by

$$c_1 = \left( \frac{\lambda + 2\mu}{\rho} \right)^{\frac{1}{2}} \tag{2.5}$$

**longitudinal velocity of an elastic wave**. Equation 2.5 says us that a change in volume, or dilatational disturbance, will propagate at velocity $c_1$.

The equation 2.4 can be expressed in terms of $E, K, \nu$

$$c_1 = \left( \frac{E(1-\nu)}{\rho(1+\nu)(1-2\nu)} \right)^{\frac{1}{2}} = \left( \frac{3K(3K+E)}{\rho(9K-E)} \right)^{\frac{1}{2}} = \left( \frac{\mu(4\mu+E)}{\rho(3\mu-E)} \right)^{\frac{1}{2}}$$

If we perform the operation of curl on the equation 2.4 (considering that the curl of a gradient of a scalar is zero) this gives

$$\mu \triangle^2 \mathbf{w} = \rho \frac{\partial^2 w}{\partial t^2}$$

where $\mathbf{w} = \bigtriangledown \times \mathbf{u}/2$ is the rotation vector. As a result, we have the form of the vector wave equation, expressed as

$$\triangle^2 \mathbf{w} = \frac{1}{c_2^2} \frac{\partial^2 \mathbf{w}}{\partial t^2}$$

where $c_2$ is the transversal propagation velocity and is given by

$$c_2 = \left( \frac{\mu}{\rho} \right)^{\frac{1}{2}} \tag{2.6}$$

**Transversal velocity of an elastic wave**. The equation 2.6 can be expressed in terms of $E, \nu$

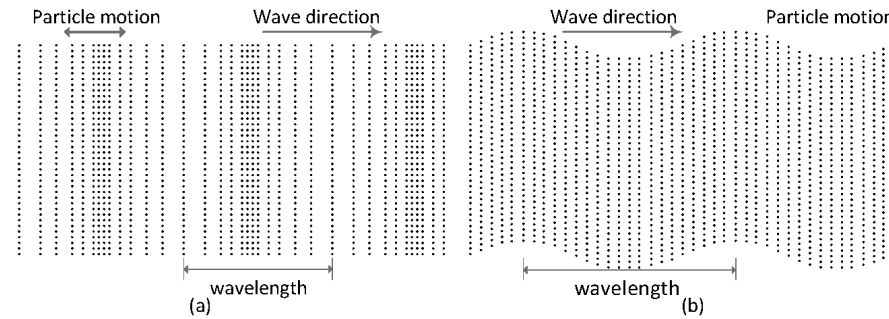$$c_2 = \left( \frac{E}{2\rho(1+\nu)} \right)^{\frac{1}{2}}$$

Figure 2.3.: directions of particle motion. (a) AE longitudinal wave, and (b) AE transversal wave.

The waves can propagate inside of an elastic isotropic solid at two different speeds $c_1$ and $c_2$. The longitudinal or volumetric waves involve no rotation and propagate at velocity $c_1$. On the contrary, transversal or rotational waves involve no volume changes and propagate at velocity $c_2$.

Comparing equations 2.5 and 2.6 is evident that $c_1 > c_2$ and so longitudinal waves are faster than transversal. In seismology, longitudinal waves are termed $P$ waves (for primary) and transversal $S$ waves (for secondary).

Supposing that simple harmonic plane waves are propagating in a media. The basic nature of the particle motion for the two types of waves related to velocities $c_1$ and $c_2$ can be shown in figure 2.3.

## 2.5. AE wave propagation and attenuation

The physical basis for the propagation of AE waves in various media lie in the interaction of the discrete atoms of a solid. As mentioned in the previous topic, in solid and fluids mechanics, the medium is considered to be continuous, and so the properties of the medium are considered continuous functions representing averages of microscopic quantities.

In real cases, the objects to be measured have boundaries (surfaces). This cause a modification of the behavior of the elastic waves and so new wave properties are found near or on the surface of these objects. When an AE wave travels across a material and encounters other medium, the change in the wave depend on three possible mechanical interfaces [CD77]. ($i$) Solid in contact with a vacuum, the surface is stress free. ($ii$) Solid in contact with a fluid, the shear stress is zero, but the transversal stress is continuous. ($iii$) Solid in contact with other solid (firmly adhered), the particle motion on the surface is the same. If there is a gap between the two solids, the stress on each surface is given be the ratio of their elastic mechanical impedances.

In contrast with the equation 2.3 for continuous media, in real objects, the AE wave suffer

attenuation due to different factors. Additional complexity is added when an elastic wave travels across a porous fluid-sutured media as a rock. In inhomogeneous media, we have to use the Biot`s equations for poroelasticity media [WF05].

If the material at the surface is the same as the bulk material (no layers) the surface wave is termed as *Rayleigh wave*, which is a combination of the shear wave and longitudinal wave [WF05]. However, an elastic wave propagates in a media in various modes [Hud80], apart from the above derived *longitudinal wave* (or P-wave, dilatational wave) and the *transversal wave* (or S-wave, shear wave, distortional wave). In semi-infinite media, surface waves such as the Rayleigh wave, which performs elliptic particle motion on planes perpendicular to the surface. Other surface waves are called with the name of who analyzed them such as; *Love waves* which occur when a substrate has a single layer, parallel to the plane of the layer, but perpendicular to the wave motion [Hud80]. *Lamb waves*, which occur in plates with ticks less than the wavelength, the particle motion here is perpendicular to the plate. *Stoneley waves*, which propagate between the interfaces of two elastic media [Hud80].

When AE waves propagate through a medium, encounter cracks, cavities, inclusions, and as a result, scattering and diffraction problems take place generating attenuation of the AE waves.

Attenuation problem can be attributed to the following phenomena when an AE wave propagates.

- **Geometric attenuation** is the phenomenon in which the amplitude of an AE wave decreases as the wavefront spread out over a volume as it propagates away from its source. This phenomenon is the most important and the main reason of attenuation of the AE wave. There are various mathematical models to quantify the geometric attenuation for different AE wave types, all of them are based on the decrease of amplitude with the inverse propagation distance $r$. For example, in *Love waves* the amplitude attenuation have an inverse proportion to $r^\eta$, where $\eta$ is the geometric attenuation coefficient, which depend on the type of wave front [EJ57].

- **Absorption or dissipation** is a result of internal friction due to the work done at material interfaces when two materials are not elastic bounded [OJ99].

- **Scattering** is a complex process, which is dependent upon the intrinsic length scale of the scatterer (cracks, inclusions, etc.), the number of scatterers per unit volume, its distribution and the acoustic properties of these scatterters in relationship to the base material. This scatterers cause deflection and dispersion of the AE waves [GK78].

- **Diffraction** occurs when an AE wave encounters an obstacle and it is a sum of many phenomena. It is the bending of an AE wave passing through an aperture or around the edge of irregularities such as crack [GK78].

- **Thermoelastic dissipation** occurs in interactions between fluid (internal gap filled with air) and solid phase. The AE wave establishes different temperatures in the fluid and the solid, thus it initiates interfacial heat flow, causing dissipation of energy [Bio62].

The attenuation of elastic waves can be qualitatively represented by means of $Q$ values.

$$Q = \frac{2\pi E}{\triangle E} \tag{2.7}$$

Initially the wave has an energy $E$ which is attenuated by $\triangle E$ over one wavelength propagation [GO10]. When an AE wave propagates for a distance $D$, the amplitude $A(f)$ of the frequency components $f$ is attenuated from $A_0$ to

$$A(f) = A_0 \times exp\left(\frac{-\pi f D}{vQ}\right) \tag{2.8}$$

where $v$ is the AE wave velocity in the medium. From the equation, above we can deduce that the higher the frequency components of the AE wave, the higher is its attenuation.

## 2.6. AET in the mining industry

Condition monitoring of different structures and machines as well as components, is the area in which AET has been mainly applied. Since the beginning of the AET some practical uses for mining has been developed such as underground mine stability [EMM98], these studies have been expanded to monitor the geomechanics of gas and petroleum installations as well as different mines [FM90]. Nowadays, AET is applied on numerous areas in mining and its applications are continuously growing.

The main challenge for this method is to be robust enough to bear the extreme mining conditions, also the high signal to noise ratio present in a mining environment is an important issue to overcome. The table 2.2 summarizes some applications of AET in mining.

Table 2.2.: Some applications of AET.

| Application | Process/use | Description |
|---|---|---|
| Rock stability | Mining safety | To avoid mine collapses, AE activity is monitoring high concentrate stress areas in mining structures [EMM98]. Monitoring of radioactivity waste disposal [SHEE05]. |
| Monitoring of gas and oil platforms and reservoir | Condition monitoring/ mining safety | Detection of sand in flow lines of offshore gas and oil producing platforms. Study of hydraulic fracturing of oil and gas reservoir rock [SK75]. |
| Monitoring of tunneling projects | Mining safety | Examination of the effects of different stress regimes and excavation techniques [FY98]. Assessment of the excavation damages zones [MYR98]. |
| Determination of rock mechanical properties | Rock cutting/ mining safety | Study of the fundamental deformation and failure behavior of rocks [Ron79]. Understanding the bursting mechanism of rock under polyaxial stress conditions [HMF10]. |
| Rock cutting | Material extraction/rock cutting | Recognition of the type of rock being cut for mining and tunneling machinery. Improvement of efficiency and productivity in the operation [GNKB12], [NKBR13]. |
| Monitoring of SAG Mill process | Grinding | Monitoring of the internal state of SAG mill and operation conditions such as particle size, charge mass, pulp density, amount of grinding balls [AT00], [SCWL99]. |
| Monitoring of mechanical integrity of pressure vessels and pipelines | Maintenance/ condition monitoring | Detection and location of pipeline leaks. Characterization of source and leak mechanisms [M$^+$99]. Monitoring of crack propagation and correlation with AE signals [ELHC06]. |
| Monitoring of bubble size distribution | Column flotation process | Determination of the bubble size distribution in column flotation using the bubble signal spectrum. Improvement and efficiency of the operational conditions [S$^+$11]. |
| Monitoring of wire ropes | Material extraction/condition monitoring | Detection of internal wire rearrangement and wire breaks produced by over tensile load in wire ropes. Applications on electric shovels and draglines [Z$^+$06]. |

| Application | Process/use | Description |
|---|---|---|
| Monitoring of gear-boxes in mining machinery | Diagnose, condition monitoring | Monitoring and diagnose of large gearboxes based on external non-invasive AE measurement on the ring gear. Applicability to high torque gearboxes [MGN12]. |
| Condition monitoring of large diesel engines | Diagnose, condition monitoring | Monitoring of large diesel motors by means of acoustic emission energy revolution signals [PS05]. Identification of injector faults in diesel engines [EGB10]. |
| Diagnose of power transformer machinery | Diagnose, condition monitoring | Assessment of the condition of power transformers based on partial discharge location using AE sensors. AE signal amplitude and time differences of arrival are used [SW13]. |
| Monitoring of rock bolting in underground mine | Mining safety/ condition monitoring | Evaluation of rock bolt integrity by means of guided AE signals. Fourier and wavelet transforms are applied to evaluate the signals [Lo12]. |
| Diagnose of bearing in mining machinery | Diagnose, condition monitoring | Monitoring of low rotational speed of bearing by means of AE analysis allow real time evaluation and a more accurate method of diagnose than other traditional ones [BNGB12] |

## 2.7. AE sources and properties

Acoustic emissions are produced when material is stressed enough by an external force. Depending on the direction of the force the stress produced may be tensile, compressive, shear or torsion. Normally, the stress produced in a material is a combination of these classes of stress. As a result of this stress, the material is deformed elastically and if the stress is high enough, plastically too. From a microscopic point of view plastic deformation involve the sliding of atomic planes over another one. AE waves can be seen as short pulse of elastic and kinetic energy that propagate through a material. The quantity of energy released and the amplitude of the AE wave depend on the size of the source and the speed that the source was developed.

Typical AE sources include leaks, rubbing, impacts, phase transformation, crack nucleation and plastic deformation. Mechanisms as a wire rope break or brittle fracture produce AE bursts and mechanisms as rock cutting and leaks produce continuous AE waves. AE waves propagate through a material in all directions (equation 2.3 ) in a spherical way and attenuate proportionally to their frequencies and traversed distance (equation 2.8).

The material properties are dependent on the particular dimension of the material. That is because, a material can be considered as homogenous if the particular dimensions of the heterogeneities are smaller than the wavelength of the AE waves. The wavelength of an AE

wave is a function of its velocity and its frequency. The AE wave velocity is particular for each material and the propagation velocity can be measured [Gra77]. On the other hand, AE waves contain frequencies in the range from 20 kHz to several megahertz. Materials such as rock and concrete can be considered as homogeneous, because AE wave velocities are over 1000 m/s and AE waves in this material have frequencies around 100 kHz, as a result the AE wavelengths in this material have a few centimeters (larger than the heterogeneities).

When a crack initiates in a material, it produces a large amplitude AE wave. However, as the material deformation on the crack tip progresses, it produces small amplitude AE waves. The amplitude of the AE wave is proportional to the area created by the crack and the crack propagation velocity.

Most of the time, unwanted signals (noise) come together with AE signals, as a product of AE events as environment AE phenomena (from process around the measurement). This noise must be identified and removed by means of digital filter from the recorded data.

## 2.8. Summary

In this chapter, the main topics about acoustic emission technique are discussed. An overview of the physical AE phenomena was described in five zones, namely mechanical wave propagation in the material, the collection of the signal by means of a AE sensor, the (pre)amplification of the analog signal, the transformation of the analog signal to a digital signal by means of a signal conditioning equipment, and finally, the signal processing of the digital signal to calculate the appropriate features of the signal by means of a software.

Next, the principal advantages and disadvantages of the AET were reviewed from a mining point of view. The main advantages of this technique are; continuous online monitoring, global monitoring of large structures, easy setup of the DAQ system, great amount of information can be processed, it is not intrusive, etc. The main disadvantages correspond to low rate of signal-to-noise, advanced signal processing is required, no information about shape of the AE source is provided, repeatability is not possible, large amount of data is recorded, etc.

Then, the theory of acoustic emission waves is described, from a physical point of view. The wave equation for isotropic elastic solid is derived as well as the longitudinal and transversal velocity of an elastic wave.

The issue of wave propagation and attenuation is explained. The obstacles of an elastic wave when it propagates are cracks, cavities, inclusions. As a result, scattering and diffraction takes place. These imperfections generate attenuation of the AE waves. The main reasons of the attenuation are detailed. Also, a mathematical expression to qualify the attenuation is given as a function of the $Q$-values.

The descriptions of the principle uses of the AET in mining industry are also reviewed. Finally, the main AE sources are mentioned as well the main AE properties.

# 3. Signal processing and AE signal features

## 3.1. Introduction

In this section, the signal analysis used in chapter 5 is reviewed in depth to provide a mathematical background of the parameters used in this chapter. AE signals have different characteristics depending on the AE source. Monitoring of AE signals is used for a variety of process control, with the background noise, being the most important difficulty to overcome. Data acquisition systems convert analog signals to digital signals using antialiasing filters. The most common AE parameters are number of peak, or bursts, rise time, duration, energy and RMS. Normally, these parameters are calculated from AE signal, after the signal have exceed an arbitrary threshold.

It is important to keep in mind that before signal processing there are four stages, which have an impact on shape and amplitude of the AE signals. These stages are; the AE source, the propagation of the wave, the AE sensor, and the signal conditioning. The shape of the original mechanical wave is enormously different from the final AE signal. This difference exists also in the frequency content, which is analyzed by means of the Fourier analysis. Wave propagation determines the signal shape. Thus the shape-dependent signal features are influenced. When a piezoelectric AE sensor is reached by a mechanical wave, it produces a corresponding oscillating voltage at the frequency of the mechanical wave. The sensitivity of the AE sensor depends on the frequency of the mechanical wave, which is greatest at its resonant frequency. The signal conditioning refers to the amplification, filtering, detection and measurement of the signal. The filter defines the frequency range to be used and leaves out others frequencies corresponding to the background noise.

AE parameters are obtained after signal processing with the aim of detecting AE events of interest. Thus, the first step is the correlation of the AE parameters with these AE events. AE signals contain high frequency components and AE events arise in a rapid and randomly manner.

Two approaches can be defined for signal processing analysis, namely *parameter-based* and *signal based*. The first, corresponds to the classic parameter processing where only some features of AE signals are recorded. This was the only possibility decades ago, when data acquisition equipment was limited. On the other hand, *signal based* processing consider the recording of the whole signal or AE event for further analysis. Nowadays, the latter approach is possible

thanks to the huge technological progress in signal processing equipment.

## 3.2. Time domain analysis

Measurement of relevant features in time domain is made typically based on a threshold trigger. The most basic signal feature in this domain is the amplitude, which is the largest voltage present in the signal waveform. Its importance resides in the fact that it must exceed a predetermined threshold to trigger the detection of the signal. Also, an automatic amplitude threshold can be employed, using a bandpass filtering of the signal [Lyo04].

The number of waveform peaks or bursts is detected in a signal using a specific signal length. For this task, it should specify the threshold height difference between a burst and its neighboring samples. The algorithm returns the number of burst that exceeds their neighbors by at least the value of the specified threshold.

AE signals are basically random signals. Thus, some features as root mean square, variance, kurtosis value and standard deviation can represent this type of signals in a statistical manner based on same distributions. These features help us to understand the nature of the signals.

**The Root Mean Square** (RMS) of an AE signal with a specific length can be calculated by

$$RMS = \sqrt{\frac{1}{n}\sum_{i=0}^{n}|x_i|^2} \tag{3.1}$$

where $n$ is the number of samples in the signal length. From a physical point of view, it can be considered as the mean power of a stochastic signal. Also, it is an estimation of the signals` deviation.

**The variance** $\sigma^2$ is also computed from an AE signal or set of input data samples with a specific signal length. It is a measurement of how spread out the different samples of the AE signal might be considering a random process. Mathematically, it can be expressed by

$$\sigma^2 = \sum_{i=1}^{n}\frac{(x_i - \mu)^2}{n} \tag{3.2}$$

where $n$ is the number of samples in the signal length and $\mu$ is the mean. From the variance, another statistical signal feature can be obtained, namely **the standard deviation**, which is just the square root of the variance, $\sigma$.

$$\sigma = \sqrt{\sum_{i=1}^{n}\frac{(x_i - \mu)^2}{n}} \tag{3.3}$$

so a variance of zero indicates that the signal samples are identical and small variances indicate that the signals samples are close to each other and consequently, close to the mean value too.

**Kurtosis value** is related to the normal distribution of the AE samples. A normal distribution has a kurtosis value of 3. A kurtosis value of less than 3 indicates a flatter distribution than normal and a kurtosis value greater than 3 indicates a sharper distribution than normal. In other words kurtosis value is an indicator of the shape or peakedness of a probability density. Mathematically, for a discrete number of samples or finite AE signal, it can be expressed by

$$Kurt = \frac{\frac{1}{n}\sum_{i=0}^{n-1}(x_t(i) - \mu)^4}{\sigma^4} \tag{3.4}$$

where $n$ is the number of samples in the signal length and $\mu$ is the mean. To show in a graphical manner, the effect of the type of signal in the kurtosis value, a histogram can be used.

In this context, **Entropy** is an indicator of uncertainty of the random sample. In this work, it is calculated the Shannon`s entropy of the wavelet packet coefficients, which quantifies the expected values of the samples contained in a discrete AE signal. Also, it is an average unpredictability of the AE signal. AE signals with high entropy indicate they are fairly unpredictable. This definition was first established by C. E. Shannon [Sha48]. Mathematically, it is calculated by the following expression

$$WPE = \int_{-\infty}^{\infty} f(t)^2 dt = \sum_{j=1}^{n} E_j = \sum_{j=1}^{n} [x_t^2(i)log(x_t^2(i))] \tag{3.5}$$

where $n$ is the number of samples of the input series.

## 3.3. Frequency domain analysis

### 3.3.1. Frequency domain representation

Transformation of AE signals to the frequency domain is commonly used by signal processing engineers to analyze and monitor processes. This analysis provides sometimes a more effective representation of the AE signals as well as efficient processing of the data, when compared to time domain analysis.

One of the most commonly used frequency domain transformation is the Fast Fourier transformation (FFT). Displaying and interpreting FFT results in the extraction of important frequency information of the AE signal [Lyo04].

FFT illustrates time domain functions into frequency domain representations of the frequency content of these functions. Mathematically, continuous functions can be expressed as

$$X(f) = \mathcal{F}\{x(t)\} = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft}dt \tag{3.6}$$

where $x(t)$ represents the time domain signal and $X(f)$ represents the Fourier transformation of the signal.

For discrete functions, the discrete Fourier Transformation is used. It is defined by

$$X_k = \sum_{i=0}^{n-1} \mathsf{x}_i e^{-j2\pi ik/n} \tag{3.7}$$

for $k = 0, 1, 2, ..., n - 1$. Where $\mathsf{x}$ represents the time domain signal, $X$ represents the Fourier transformation of the signal and n is the number of samples in the discrete signal and in the discrete frequency domain.

The direct application of the equation 3.8 on the discrete signal requires $n^2$ operations. On the contrary, computational efficiency algorithms require as little as $nlog_2(n)$ operations. These types of algorithms are called *fast Fourier transformation*.

The Fourier transform of any discrete sequence of samples results always in a complex output sequence $X$ of the following form

$$\mathcal{F} = X = X_{Re} + jX_{lm} = Re[X] + jIm[X] \tag{3.8}$$

which has the property of $X_{(n-i)} = X_{(-i)}$, thus the $(n-1)^{th}$ element of $X$ contains the results of the $-i^{th}$ harmonic.

Furthermore, if $x$ is real, then the $\pm i^{th}$ harmonic are complex conjugates,

$$X_{n-1} = X_{-i} = X_i^* \tag{3.9}$$

Therefore, $ReX_i = Re\{X_{n-i}\}$ and $ImX_i = -Im\{X_{n-i}\}$. These symmetrical Fourier properties of real sequences of samples are referred to as conjugate symmetric in equation 3.9.

Sinusoidal signals have a simple mathematical representation and they are the most basic signals in the theory of signals and systems.

$$x(t) = A\cos(\omega t + \phi) = A\cos(2\pi f_0 t + \phi) \tag{3.10}$$

where $f = \frac{1}{T}$ is the frequency, with $\omega T = 2\pi$. The parameters (A) amplitude, $(\omega_0)$ the radian frequency and $(\phi)$, phase shift [rad] of the cosine signal are fixed numbers for a particular cosine signal.

Many physical systems generate signals that can be modeled (represented mathematically) as a sine cosine functions versus time. Sinusoidal signals are defined in terms of familiar sine and cosine functions of trigonometry and their properties determine the properties of sinusoidal signals.
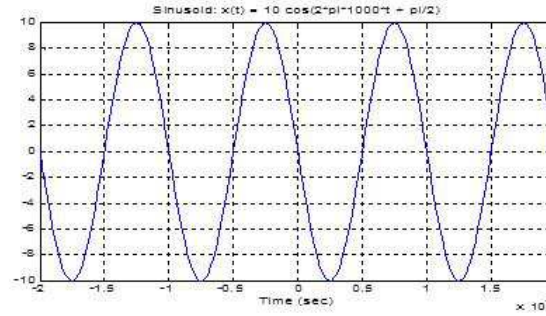
Figure 3.1.: Sinusoidal signal.

How large can the sampling spacing be such that the cosine signal can be reconstructed accurately from the samples? The sampling spacing must be less than half the period $(T_s < T/2)$, that is, the average number of samples per cycles need to be only slightly more than two.

Analysis and manipulation of sinusoidal signals is often greatly simplified by dealing with related signals called **complex exponential signals**. Complex number $z = x + jy$ is often represented as a point in a Cartesian complex plane.

Another way to represent a complex number is the polar form $z = re^{j\theta}$. By using Euler`s formula ,it can be expressed as

$$z = re^{j\theta} = r\cos\theta + jr\sin\theta \tag{3.11}$$

Complex exponential signals are defined as

$$z(t) = Ae^{(\omega_0 t + \phi)} = A\cos(\omega_0 t + \phi) + jA\sin(\omega_0 t + \phi) \tag{3.12}$$

Plotting a complex signal as a function of the time requires two graphs. Both are real sinusoidal signal, and they differ by only a phase shift of $\pi/2$ rad (Figure 3.2).

Any sinusoid can be written in the following form

$$A\cos(\omega_0 t + \phi) = \Re\left\{Ae^{(\omega_0 t + \phi)}\right\} = \Re\left\{Ae^{j\theta}e^{j\omega_0 t}\right\} \tag{3.13}$$

Phasor addition rule states that the sum of two cosine signals each having the **same frequency**, but having different amplitude and phase shift, can be expressed as

$$x(t) = \sum_{k=1}^{N} A_k \cos(\omega_0 t + \phi_k) = \cos(\omega_0 t + \phi)$$
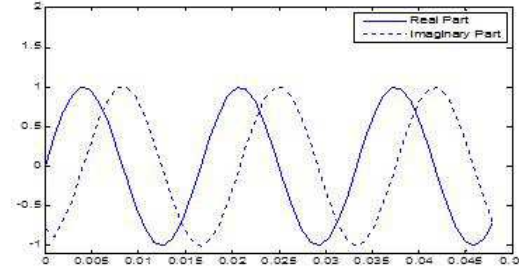
Figure 3.2.: Real and imaginary parts of a sinusoid signal.

$$\sum_{k=1}^{N} A_k e^{j\varphi_k} = A e^{j\phi}$$

Knowing this, the spectrum of a signal can be represented. It is a compact representation of the frequency content of a signal that is composed of sinusoids. It is the simple collection of amplitude, phase, and frequency information that allow us to express the signal in the form;

$$x(t) = A_0 + \sum_{k=1}^{N} A_k \cos(2\pi f_k t + \phi_k) = X_0 + \Re\left\{\sum_{k=1}^{N} X_k e^{j2\pi f_k t}\right\}$$

where $X_0 = A_0$ is a real constant and $X_k = A_k e^{j\phi k}$ is the complex amplitude (i.e. phasor).

This information in graphical representation allows us to see interrelationships among the different frequency components and their relative amplitudes quickly and easily.

Some extraordinarily complicated waveform can be constructed from rather simple combinations of basic cosine waves. The inverse Euler formula gives a way to represent $x(t)$ in the alternative form as

$$x(t) = X_0 + \Re\left\{\sum_{k=1}^{N} X_k e^{j2\pi f_k t}\right\}$$

$$\text{The inverse Euler form} \Longrightarrow x(t) = X_0 + \sum_{k=1}^{N}\left\{\frac{X_k}{2} e^{j2\pi f_k t} + \frac{X_k^*}{2} e^{-j2\pi f_k t}\right\}$$

It is the set of $2N + 1$ complex amplitude and the $2N+1$ frequencies that specify the signal.

$$\left\{ (0, X_0), (f_1, \frac{1}{2}X_1), (-f_1, \frac{1}{2}X_1^*), \ldots (f_k, \frac{1}{2}X_k), (-f_k, \frac{k}{2}X_k^*) \right\}$$

Each pair $(f_k, \frac{1}{2}X_k)$ indicates the magnitude and relative phase of the signal component contributing at frequency $f_k$. It is common to refer to the spectrum as the frequency-domain representation of the signal.

### 3.3.2. Power spectrum

The power spectrum function is related to the Fourier transform. It calculates the harmonic power in a signal. It can be calculated by the following expression

$$PowerSpectrum = \frac{FFT^*(signal) \times FFT(signal)}{n^2} \tag{3.14}$$

where $n$ is the number of samples of the signal and the symbol $*$ represents the complex conjugate [FP93]. Since the power spectrum format is identical to the real part of the FFT, single-sided formats apply to the power spectrum. With a single-sided format, less memory is needed as well as it eliminates redundancy while retaining the whole information. The power spectrum shows the harmonic power in discrete-time, real-valued signal samples.

From the power spectrum, it can be calculated the frequency and power peak. The frequency peak is an estimated frequency of the peak in the input Power Spectrum and can be calculated as

$$Fre.Peak = \frac{\sum Powerspectrum_j * (j * df)}{\sum (Powerspectrum_j)} \tag{3.15}$$

for $j = i - \frac{span}{2}, \ldots i + \frac{span}{2}$ where $i$ is the peak index, power spectrum $(j)$ = power in bin $j$, and $df$ = the frequency bin width.

The power peak refers to the estimated power of the peak in the input Power Spectrum of the discrete signal. It can be calculated as

$$Pow.Peak = \frac{\sum (Powerspectrum_{(j)})}{ENBW} \tag{3.16}$$

for $j = i - \frac{span}{2}, \ldots i + \frac{span}{2}$ where $i$ is the peak index, power spectrum $(j)$ = power in bin $j$, and $ENBW$ = Equivalent Noise Bandwidth of the Window [FP93].

## 3.4. Joint time-frequency analysis

### 3.4.1. Time-Frequency domain representation

Time-frequency signal analysis is related to signals with time varying frequency content. It is one way to represent a signal, which changes its frequency over the time. This representation is called *time-frequency distribution*. Normally, a signal contains parts with valuable information and other parts are just noise. A way to separate this valuable part of the signal from the noise part is to characterize its frequency content. This task is possible if fast Fourier transform (FFT) is applied to convert the time waveform to a frequency spectrum.

The use of FFT provides frequency information that may be hidden in the time domain. As seen before the square of the FFT components are called power spectrum, which represents the energy distribution of a signal in the frequency domain. This distribution of the energy refers to the relative intensity of energy of a signal at each frequency for the complete signal. Nevertheless, the power spectrum for a shorter time can be more useful.

In time domain, the frequency representation is averaged over the values of the time representation. In the same way, in frequency domain, the time representation is averaged over the values of the frequency representation. On the contrary, in time-frequency representation these two variables are presented together.

Some of the common applications of time-frequency representation are; the analysis of the raw data in $(t, f)$ domain in order to find properties as time variation, frequency variation, relative amplitudes, separation of components from each other and from the background noise, among others. One method to provide power spectra for a short time scale is the spectrogram. Figure
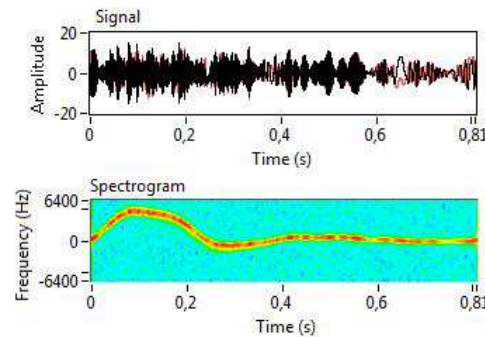


Figure 3.3.: STFT Spectrogram of signal [Lab11].

3.3 illustrates the spectrogram of a signal, where the frequency content is depicted as function of time.

There are two categories of time-frequency analysis, namely linear methods and quadratic methods. Linear methods are used to reduce noise and find signal components. Quadratic methods are used to select features in a signal.

### 3.4.2. Wavelet analysis

Wavelet functions are used to decompose signals. The wavelet transform are based on the inner products of the signal and a family of wavelets. They are localized in both time and frequency domain, thus wavelets are suitable for non-stationary signals, whose frequency content vary with the time. Multiresolution analysis of non-stationary signals is possible using the adaptive time-frequency resolution of the wavelet signal processing [DDJ94].

Wavelets analysis is employed in many applications such as analysis of signal in different scales, noise reduction, data compression and features extraction of data.

Multiscale analysis refers to the observation of a signal at different time and frequency scales in order to understand both the long-term trends and the short-term variations of a signal at the same time.

Noise reduction is one of the most common used applications of wavelets in signal processing. It is much more effective than traditional methods and retains the details of a signal after denoising.

Data compression is used when a high amount of data is stored or/and transmitted. The wavelet transform generates large coefficients only around discontinuities, thus it is a useful method to convert signals to sparse representation.

The extraction of relevant features from the data using wavelets is a useful tool to analyze and interpret signals. It permits the signal characterization by means of local features such as peaks.

Wavelet transform is divided into two classes, namely Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). The main differences between these two types of wavelets is the fact that using continuous wavelets coefficients, it cannot use the resulting wavelets coefficients to recover the original data sample. Thus, for applications that requires signal reconstruction, discrete wavelets are recommended.

For non-stationary signals, the scale is a basic variable of the wavelet transform. Small values of scaling indicate high values of frequency and vice-versa.

Wavelet analysis is based on the decomposition of the original signal into varying wavelet functions. This functions are called "mother wavelet". A fundamental condition of a mother wavelet is the "admissibility condition", which enables the existence of inverse transformation. Each mother wavelet has to satisfy the following condition.

$$\int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{\omega} < \infty$$

where $\Psi(\omega)$ is the Fourier transform of the mother wavelet. There are many mother wavelets for different applications. The most well-known wavelets are; Morlet, Daubechies, haar, symlet, biorthogonal and coiflet wavelet.

Daubechies wavelets are one of the most used wavelets. These compactly supported orthonormal wavelets make discrete wavelet analysis possible. The names of this family of wavelets are denoted by dbN, where N is the wavelet order, and db is an abbreviation of Daubechies. Higher the order, smoother the wavelet.

Figure 3.4 illustrates the discrete wavelet transform. Two channel perfect reconstructions are commonly used to implement wavelet transform. The signal is first filtered by filter banks $G_x(z)$, and then, the outputs are downsampled by a factor of 2, and so forth, for each out until processing.

Lowpass filters remove high-frequency fluctuations from the signals and preserve slow trends. Then the outputs of lowpass filters are used as an approximation of the original signal. Highpass filters take the slow trends from the signal and keep the high frequency fluctuations.

The outputs of highpass filters provide detail information about the signals. The outputs of the lowpass filters and highpass filters define the approximation coefficients "A" and the detail coefficients "D" respectively as can be seen in figure 3.4. Detail coefficients are also called
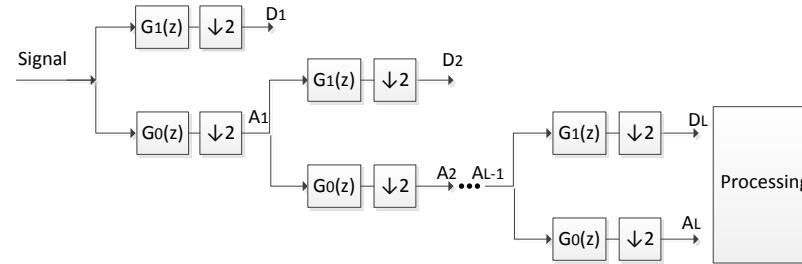


Figure 3.4.: Discrete wavelet transformation.

wavelet coefficients and the subscripts 0 and 1 indicates the decomposition path. 0 indicates lowpass filtering and 1 indicates highpass filtering.

Afterwards, the modified signals or outputs are upsampled by a factor of 2 and filtered by another filter bank.

From the wavelets decomposition (approximate and details signals), the amount of energy contained can be determined. Mathematically, this parameter can be expressed as

$$\int_{-\infty}^{\infty} f(t)^2 dt = \sum_{j=1}^{n} E_j$$

## 3.5. Automatic AE event detection

Acoustic emission testing is fundamentally related to the recording of AE phenomena from a monitored process. The onset detection of an AE phenomena of interest can be determined by an operator or automatically by a picking algorithm. The onset time is usually defined as a point where the energy of an AE event first reaches the AE sensor [Leo00]. This definition is preferred as the point in the time where the difference from the background noise first occurs. Based on experience, an analyst can find the time arrival with exactitude as can an automatic picking algorithm. However, in an AE test, normally thousands of pick arrivals are observed. Hence automatic method of detection and recording is necessary.

Methods of reducing noise to detect the pick arrival must be applied carefully in order not to distort the arrival signal. The determination of an onset AE event is an important issue for AE source location. However, sometimes, it is preferred a simple detection of an AE event without timing information of the onset of the first arrival. To solve problems such as data reduction and automatic signal classification, a simple detector is required. When the difference between noise and signal are slight it is incorrect to apply a general detector algorithm for every situation but specific or even multiple detector algorithms should be applied. For example, if the noise and signal have the same frequency content, a bandpass filter is not appropriate and other method should be applied.

A threshold is a simple AE event detector, however is not suitable with a low signal to noise ratio or AE events with continuous waveforms. In these situations, most sophisticated methods should be applied[IKAK12]

An automatic amplitude thresholding detector uses a bandpass filtering of the signal. The bandpass filter can be arbitrarily set and suitable filter can be used. Mathematically, the threshold can be written as

$$Threshold = 4 \times \mu_{\frac{1}{2}} \left\{ \frac{|BPF|}{0,6745} \right\}$$

where $\mu_{\frac{1}{2}}$ represents the median of the sample values, $BPF$ is the bandpass filtered signal and the value 0,6745 is a computational estimation [DDJ94].

The right part of this equation represents four times the standard deviation of the background noise [DDJ94]. Note that this threshold is set in function of the background noise and it is different for each AE signal (with different background noise levels).

## 3.6. Summary

In this chapter the main topics of signal processing is reviewed. Beginning with the time domain analysis, a discrete signal with finite number of samples is considered. Waveform features such as number of peaks are reviewed along with the statistical parameters such as root mean square, variance, kurtosis value and standard deviation. Also the Entropy of a discrete signal is described, which is an indicator of the uncertainty of the random samples of the signal. In this case, the Shannon`s Entropy was selected to be used.

Next, Fourier and Spectral analysis are reviewed. The mathematical background of Fourier transformation is described for continuous and discrete signals. Also, the frequency domain representation is described, starting at the sinusoids signals, where their mathematical representation is given. The main parameters of sinusoid signal are presented. The complex exponential representation and the polar representation of the sinusoidal signals are detailed.

The mathematical spectrum representation are presented as the spectrum of a sum of sinusoids, by which complicated waveforms can be constructed from rather simple combinations of basic cosine waves.

The *Power spectrum* of a discrete signal is described as harmonic power in a signal and its relationship with the fast Fourier transformation is mentioned. Also *frequency peak* and *Power peak* are described as features extracted from the power spectrum.

Time-Frequency representation is reviewed. It is related to the processing of signals with time varying frequency content and it can be considered as a way to represent a signal, which changes its frequency over the time. Also, an overview of the wavelet analysis is given, especially, the daubechies wavelets, which are orthonormal. The wavelet decomposition process is described and the wavelet coefficients are deduced. A mathematical expression for the cumulative energy of the wavelet coefficients is presented.

Finally, the fundamentals of automatic AE event detection are described. An automatic amplitude thresholding detector, which uses a bandpass filter, is presented.

# 4. Pattern recognition applied to AE Technique

## 4.1. Introduction

Pattern Recognition Technique (PRT) is concerned with the problem of how a machine or a device learns about its surrounding or about available data, then distinguishes patterns of these data, categorizes the patterns and finally makes decisions by itself. In case of AE signals, have to be picked up, filtered, classified, evaluated and interpreted automatically (make a decision). In the mining industry, following techniques and methods use PRT (figure 4.1).
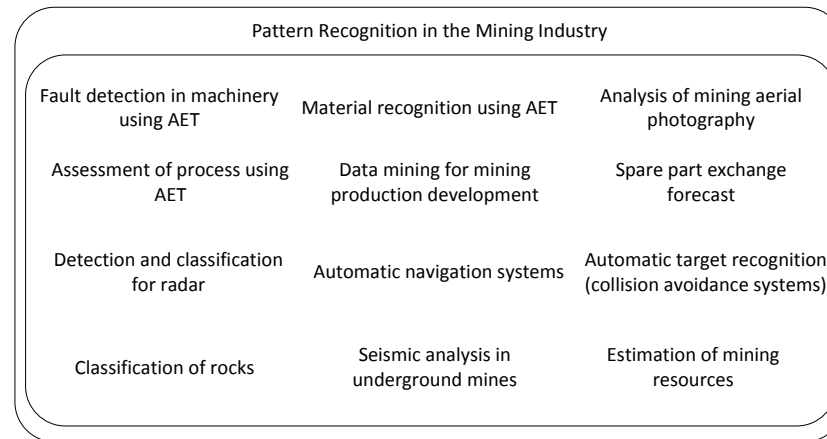


Figure 4.1.: Some uses of the Pattern Recognition Technique in the mining industry.

### 4.1.1. AE testing in the context of pattern recognition technique

Engineers that work with AET have to process a large amount of AE data to identify characteristics and diagnose the origin of AE phenomena.

In pattern recognition, a pattern is represented by a $d$-dimensional feature vector $\boldsymbol{x}$. An AE pattern corresponds to a group of features from an AE signal (a specific AE phenomenon).

When a pattern is defined, the next step is its recognition (classification). This task can be made by *Supervised classification*, which use a predefined pattern classes to classify the input patterns or *unsupervised classification*, which allocates the input pattern in different clusters. **This method is based on the similarities between patterns**, without the use of a predefined pattern class.

In AE technique, depending on the process to be measured, many times previous AE data build a predefined pattern class, but sometimes systems completely unknown are faced, in these cases the clustering approach has to be used.

Also, in some cases a combination of approaches of pattern recognition and multiples methods is more appropriate to perform an optimal classification.

The process of Pattern Recognition can be divided in data acquisition, preprocessing, classification and finally data representation and decision-making.

## 4.1.2. Patterns in the feature space

In the **classic pattern recognition**, a pattern $\boldsymbol{x}$ represents a d-dimensional vector containing $p$ features and it signifies a point in a d-dimensional space (figure 4.2). If adequate features are chosen, compact group of points (patterns) are obtained and separated between each other in this $d$-dimensional space
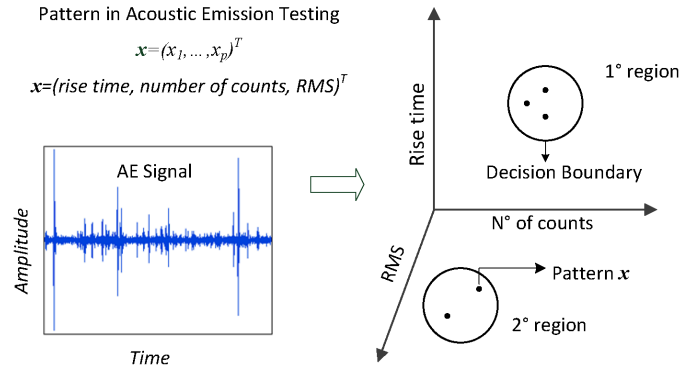


Figure 4.2.: Concept of Pattern in Acoustic Emission Technique.

If training set of patterns for each class is available, the objective is to formulate **decision boundaries (classifier)** [DGL96].

## 4.2. Overview of the Pattern Recognition Process

For the recognition we assign $C$ classes of events associated with our study and are denoted by $\omega_1, \omega_2, ..., \omega_c$. In AET, each class is related to the identification need e.g. $\omega_1$ can be related to a specific fault condition, or coal cut. Associated with each pattern $\boldsymbol{x}$ there is a categorical variable, $z$, which indicates to which class $\omega$ pertains. If $z = i$, the pattern pertains to $\omega_i$, with $i \in \{1, ...C\}$. A decision rule separates the $d$-dimensional space into $C$ regions $\Omega_i$ with $i \in \{1, ...C\}$ [Bis06]. The boundaries between each region are called the decision boundaries as shown in the figure 4.3.



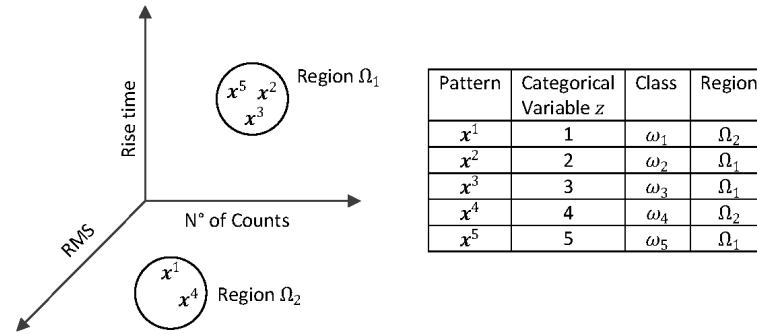| Pattern | Categorical Variable $z$ | Class | Region |
|---------|--------------------------|-------|--------|
| $\boldsymbol{x}^1$ | 1 | $\omega_1$ | $\Omega_2$ |
| $\boldsymbol{x}^2$ | 2 | $\omega_2$ | $\Omega_1$ |
| $\boldsymbol{x}^3$ | 3 | $\omega_3$ | $\Omega_1$ |
| $\boldsymbol{x}^4$ | 4 | $\omega_4$ | $\Omega_2$ |
| $\boldsymbol{x}^5$ | 5 | $\omega_5$ | $\Omega_1$ |

Figure 4.3.: Patterns and classes.

Table 4.1 shows the steps of a pattern recognition processing of AE signals.

Table 4.1.: Pattern Recognition Process for AET.

| Stages for PRT | Description and observations |
|----------------|------------------------------|
| 1. Formulation of the problem | Selection of most important variables of measurement. Description of the measurement equipment and calibration values. |
| 2. Data collection | The data is divided into training data set and data for test set. |
| 3. Pre-processing | Calculating parameters and plotting the data to observe clusters in the data or detect important variables in the data [Dav02]. |
| 4. Feature selection | The most suitable variables and parameters are selected. |
| 5. Unsupervised pattern classification or clustering | The possibility of using clustering methods is studied. Otherwise, it can serve as a preprocessing stage for the supervised classification data. |
| 6. Discrimination | The classifier is designed using the training set from the stage 3. |
| 7. Assessment of the results | In this stage we test the classifier with the test set of data from stage 2. |
| 8. Interpretation | Finally, results are interpreted based on experience. |

### 4.2.1. Types of classification: supervised and unsupervised

**Supervised classification** is also known as **discrimination** and works with a set of familiar patterns (training set) to design a classifier. In this method, one case could be that the probability density function (PDF) of each class is known.

On the other hand, **unsupervised classification** looks for groups of information and features that distinguish each group from others (similarities between each group), **this method is known as clustering**. A combination of these types of supervised and unsupervised classification is also possible.

## 4.3. Approaches to Pattern Recognition

In this section, two methods of pattern classification are presented, namely *decision theory and discriminant functions*.

### 4.3.1. Decision theory

**Bayes Decision Rule for Minimum Error**

This rule is optimum if we seek to minimize the error. This rule assumes that the previous knowledge of the a priori distributions and the class-conditional distributions are known [Bis06].

The classification based on the Bayesian decision rule requires previous knowledge of the class-conditional density functions, $p(\boldsymbol{x}|\omega_i)$ (e.g. normal distributions whose parameters are estimated from the data) or nonparametric density estimated methods (such as kernel density estimator).

### 4.3.2. Discriminant functions

The essential difference between the Bayesian rule and the discriminant function approach is the form that the discriminant function is specified and not imposed by the underlying distribution.

A function of pattern $\boldsymbol{x}$ that leads to a classification rule is called a discrimination function. In a two-class problem, a discriminant function $h(\boldsymbol{x})$ must achieve the following

$$h(\boldsymbol{x}) > k \Longrightarrow \boldsymbol{x} \in \omega_1 \mid h(\boldsymbol{x}) < k \Longrightarrow \boldsymbol{x} \in \omega_2 \qquad (4.1)$$

with $k$ constant. If $h(\boldsymbol{x}) = k$, the pattern is assigned arbitrary to one of the two classes. For two class case, an optimal discriminant function is

$$h(\boldsymbol{x}) = \frac{p(\omega_1)}{p(\omega_2)}$$

with $k = \frac{p(\boldsymbol{x}|\omega_1)}{p(\boldsymbol{x}|\omega_2)}$ . Discriminant functions are not unique. If $f$ is a monotonic function then

$$g(\boldsymbol{x}) = f(h(\boldsymbol{x})) > k$$
$$g(\boldsymbol{x}) = f(h(\boldsymbol{x})) > \check{k}$$

where $\check{k} = f(k)$ leads to the same decision as equation 4.1.

In a $C$-group case, $C$ discriminant functions are defined

$$g_i(\boldsymbol{x}) > g_j(\boldsymbol{x}) \implies \boldsymbol{x} \in \omega_1$$

with $j = 1, ..., C$ ; $j \neq i$. And so, a pattern is assigned to a class with the largest discriminant function. For two classes, a single discriminant function is the following

$$h(\boldsymbol{x}) = g_1(\boldsymbol{x}) - g_2(\boldsymbol{x})$$
$$g_i(\boldsymbol{x}) = p(\boldsymbol{x}|\omega_i)p(\omega_i)$$

leading to the Bayes` decision rule [Bis06].

**Linear Discriminant Function**

The following equation represents the family of discriminant functions that are linear combination of the components of the pattern $\boldsymbol{x} = (x_1, ...., x_p)^T$

$$g(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{x} + w_0 = \sum_{i=1}^{p} w_i x_i + w_0 \tag{4.2}$$

This is a linear discriminant function, a complete specification of which is achieved by prescribing the weight vector $\boldsymbol{w}$ and the threshold weight $w_0$.

On the one hand, if normal distributions for the class densities are assumed, with equal covariance matrices, a linear discriminant function can be created. On the Contrary, without making distributions assumptions, a linear form of the discriminant can be formed and its parameters determined.

A pattern classifier that employs linear discriminant functions is termed a *linear machine*. The *minimum distance classifier* or *nearest neighbour rule*, is an important case of these types of classifier. Supposing a set of points $p_1, ..., p_C$, is considered, one for each of the $C$ classes $\omega_1, ..., \omega_C$. The *nearest neighbour rule* classifiers a pattern $\boldsymbol{x}$ to the class $\omega_i$ associated with the nearest point $p_i$, where for each point, the squared Euclidean distance is given by

$$|\boldsymbol{x} - p_i|^2 = \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^Tp_i + p_i^Tp_i$$

The minimum distance classification is achieved by comparing the expression $\boldsymbol{x}^Tp_i - \frac{1}{2}p_i^Tp_i$ and selecting the largest value. A linear discriminant function is given by

$$g(\boldsymbol{x}) = \mathbf{w}_i^T\boldsymbol{x} + w_{i0}$$

where $w_i = p_i$ and $w_{i0} = \frac{1}{2}|p_i|^2$.

If the prototype points, $p_i$, are the class means, then we have the nearest class mean classifier. In Figure 4.4, decisions regions for a minimum distance classifier can be seen, where each boundary is the perpendicular bisector of the lines joining the prototype points of contiguous regions. A characteristic of these types of decisions regions is that two arbitrary points lying
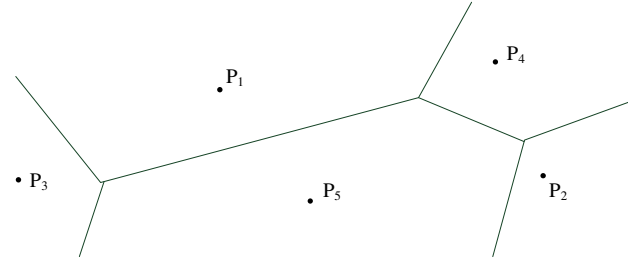


Figure 4.4.: Boundary decision for a minimum-distance classifier.

in the same region can be joined by a straight line that lies completely in the region (convex regions) [Bis06].

### 4.3.3. Pattern recognition cycle

A diagram describing the pattern recognition cycle is shown in figure 4.5. The different approaches of Pattern recognition can be separated first as *Unsupervised* or *Supervised Classification*.

In a Supervised method, the decision boundaries (classifier) can be obtained *directly* by means of discriminator functions (section 4.3.2), which form decision boundaries directly. Alternatively, it can obtain the decision boundaries *indirectly* by means of the *decision theory* (section 4.3.1), which forms decision boundaries by estimating density functions of the input data first and then, calculates the discriminant functions that form the decision boundaries [JDM00]. However, both supervised methods must be trained using a training set of data to construct the decision boundaries.

Unsupervised classification or clustering is used to explore the data and provide prototypes for a possible supervised classification. Clustering methods operate based on dissimilarity matrices and measurements on individuals.
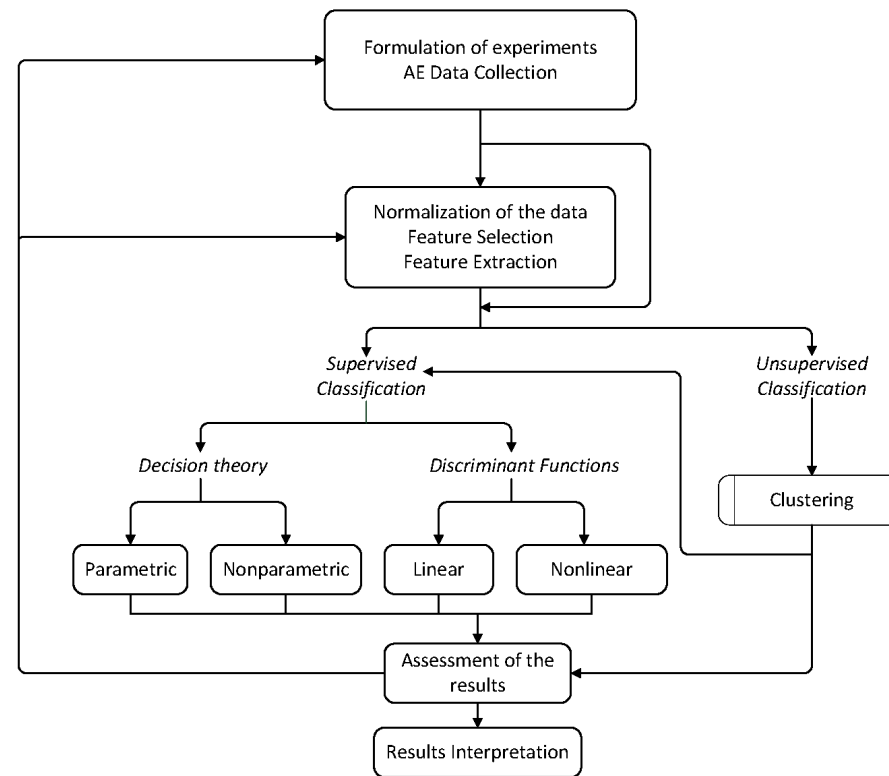
Figure 4.5.: Pattern Recognition Cycle (approaches in Pattern recognition).

## 4.4. Pre-processing of the Data

### 4.4.1. Introduction

After stage 2 in table 4.1, the data is examined to assess its quality and know more about its structure. Furthermore, the data is normalized to visualize and compare the input patterns in a suitable way. The objective is to familiarize with the data.

Features should be calculated for the whole raw data set; statistical features, waveform features, as well as, features from the frequency domain should be evaluated and can be displayed in a table to compare the values [CH08].

Scatter plots are useful to identify outliers. Different class of data can be plotted with different symbols to identify outliers. However, scatter plots can reflect only three-dimensional data, and therefore, data with a higher dimension can be difficult to evaluate. For these situations,

there are different data dimension reduction methods, such as *multidimensional Scaling* (MDS), *Linear Dimensional Analysis* (LDA) and *Principal Components Analysis* (PCA), which are useful to reduce the dimension of the data for exploratory data analysis.

How good is the performance of a classifier, depends on the number of the selected features, the inter-associations of the features and the type of classifier [Bis06]. Figure 4.6 illustrates the flow of the data in a Pattern Recognition process. The dashed square marks the stage of preprocessing of the data.
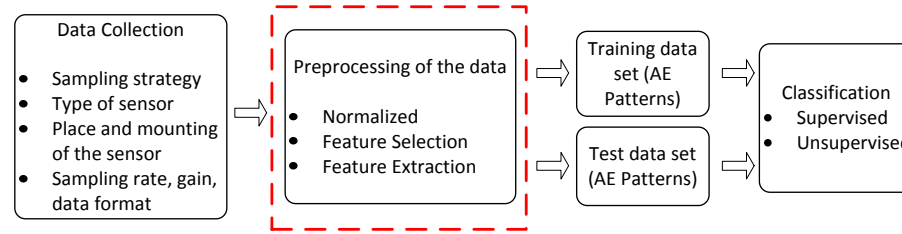


Figure 4.6.: Preprocessing of the data stage in the pattern recognition stage.

## 4.4.2. Data normalization

*Normalization of score* is useful to align different probability distributions to a normal distribution. The intention is to compare the normalized features from different data inputs.

Among the different types of normalization in statistics, *Standard Score Normalization* (SSN) is useful to normalize errors when the features are known. A standard score can present negative values if an observation is below the mean and positive if it is above the mean. Standard score can be seen as a dimensionless standardization, which is the product of the subtraction of the data mean ($\mu$) from an individual observation ($feature_{mn}$) and then dividing the difference by the standard deviation of the population $\sigma$ [LM12] (equation 4.3).

$$SC = \frac{Feature_{mn} - \mu}{\sigma} \tag{4.3}$$

SSN is defined only if the features or population parameters are known, rather than estimated.

Different scales can be standardized by means of SSN. This normalization technique is useful to examine the value of a feature relative to the rest of the features by using its distance from the mean and the size of scattering of the data.

### 4.4.3. Feature selection

After a measurement campaign, different features from the raw data set must be calculated. These features should represent the AE signals in all its dimensions; in other words, the signal must be represented in the time domain and frequency domain as well as in the time-frequency domain, if necessary. However, these features can be a huge amount. To solve this problem, an optimality criterion must be evaluated. The strategy adopted is independent of the optimal criterion, $J$, however the computational requirements do not depend on the optimality criterion.

### 4.4.4. Feature extraction

In feature extraction all features are used and transformed to a reduced feature space with the aim of replacing the original features by a reduced set of underlying features [CH08]. The extraction can be a linear or non-linear transformation process as well as supervised (using information about the class label of the features) or unsupervised process.

In this work, three features extraction techniques are review, namely *Principal Components Analysis* (PCA) (Linear and unsupervised), *Linear Discrimination analysis* (LDA) (Linear and supervised) and *Multidimensional Scaling* (MDS) (non-linear and unsupervised).

#### 4.4.4.1. Principal components analysis

Principal Component Analysis (PCA) was developed by Pearson [Pea01]. The principal objective of this technique is to derive new features in decreasing order of significance. This new features are uncorrelated and a linear combination of the original features. Also, this small group of underlying features describes better the data. From a geometric point of view, PCA is a rotation of the original feature space axes to new orthogonal axes. These new axes are ordered in terms of the amount of variation of original data.

PCA is a variable-directed or unsupervised technique, that is, it makes no assumptions about the pre-existence of clusters in the data. A geometrical description in two dimensions can be useful to understand this technique. Nonetheless, more than two dimensions can be selected in this technique.

Figure 4.7 illustrates different Pattern in the feature space. Supposing that this features represent; Energy, E, measured in $volt^2 \times sec.$ and the Amplitude, $A$, measured in $volt$, each from different AE signals. Now, the question is, what is the best straight line passing through these patterns? Considering the variable $x$ as the input variable and the variable $y$ as the dependent variable, the value of $y$ is expected as a function of $x$, $E[y|x]$, as a result, the best regression line of $y$ on $x$ (considering least squares).
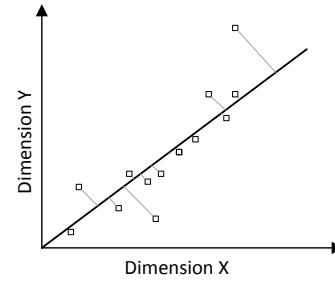
$$y = mx + c$$

Figure 4.7.: First principal component line [Pea01].

If $y$ were the regressor and $x$ the dependent variable, the solution would be the inverse situation.

So, there are two lines of best fit. It is important to note that, changing the scales of the variables does not alter the expected variables. PCA generates a single best line and the constraint that it satisfies is that minimizes the quadratic summation of perpendicular distance between points and the line. When the variables are measured in different units, the variance of each variable is made. Thus, the data is transformed to new axes, having as center the centroid of the data and coordinates in terms of standard deviation [Pea01].

The variable defined by the line of best fit is called the principal component. The second principal component is the variable defined by the line perpendicular to the first. In a two-dimension problem there is only one pair of orthogonal or principal components. In a higher dimension data, a plane of best fit is defined and successive principal components are formed in analogous way.

### 4.4.4.2. Linear discrimination analysis

The term *Linear Discrimination Analysis* (LDA) refers commonly to a technique that produces discriminant functions, which are linear with the input variables. LDA is also used as a technique that transforms the dimension space of the data to maximize the separability of classes and minimize the within-class variability.

Linear Discrimination Analysis for data reduction is based on the Fisher`s criterion. This is a linear supervised technique. The objective of the LDA is to transform the original data dimension space to a reduced dimension space.

The Fisher`s criterion is the ratio of between-class to within-class variances. Formally, a direction of $w$ is required, such that

$$J_F = \frac{|\mathbf{w}^T(\mathbf{m}_1 - \mathbf{m}_2)|^2}{\mathbf{w}^T S_w \mathbf{w}} \tag{4.4}$$

is a maximum. $\mathbf{m}_1$ and $\mathbf{m}_2$ are the group means, $S_W$ grouped within-class sample covariance matrix and can be obtained by

$$S_W = \tfrac{1}{n-2}(n_1\mathcal{P}_1 + (n_2\mathcal{P}_2)$$

where $\mathcal{P}_1$ and $\mathcal{P}_2$ are maximum probability estimates of the covariance matrices of classes $\omega_1$ and $\omega_2$. Also, there are $n_i$ samples in class $\omega_i$ [PFTV92].

### 4.4.4.3. Multidimensional scaling

Multidimensional Scaling (MDS) was introduced into the pattern recognition method by Sammon [Sam69]. It incorporates class information and provides nonlinear transformations for dimension reduction.

Multidimensional Scaling refers to the techniques that evaluate a matrix of distances or dissimilarities (proximity matrix). The aim of MDS is to produce a representation of the data in a reduced dimension space. All this techniques analyses a $n \times p$ data matrix $\mathbf{X}$, known as the sample covariance or correlation matrix. The difference is the manner these techniques work with the matrix. The first step is given a data matrix, a dissimilarity matrix is created and then MDS is applied. However, sometimes the data come already in the form of dissimilarities and so the first step is not needed.

MDS allows a nonlinear data-reducing transformation. Given a $n \times n$ matrix of dissimilarities and a distance measurement (normally Euclidean, see appendix A), a configuration of $n$ points $x_1, x_2, ..., x_n$ in $\mathbb{R}^e$ must be found, so that the distance between a pair of points is closer than the distance to the dissimilarity. All techniques have to find the coordination of the $n$ points and the dimension size of the space, $e$. Two basic types of MDS are *metric* and *non-metric MDS*. *Metric MDS* presumes quantitative data and a functional relationship between the inter-point distances and the dissimilarities [Wil02]. In case of *non-metric MDS*, the technique presumes that the data is qualitative and produces configurations that aims to keep the rank order of the dissimilarities [Kru64].

## 4.5. Supervised classification

### 4.5.1. Introduction

In a supervised classification, in advance a set of data sample (AE raw data in form of digital signals) is available. This raw data consist of measurements of AE events of interest, which are associated with labels or class types (fail condition, normal operation, etc.). This available data (training data or AE patterns) are used to train or design a computational classifier. The classifier assigns labels to input data. This assignment is made by means of a function, which represents the relationship between the input data and their labels.

There are two approaches for supervised classification namely via *Decision Theory* and *discriminant analysis* [McL92] (see figure 4.5).

The first approach (Bayes` rule) considers **normal based model (or parametric model)** and non-parametric **models** for density function estimation.

The second approach considers **linear and non-linear discrimination functions** for supervised classification.

Once the training classifier is created, it must be evaluated by means of an available test data set (AE Patterns, which have the same structure as the training set, but are separated) as shown in figure 4.8. Finally, the precision of the supervised classification method is evaluated and visualized.
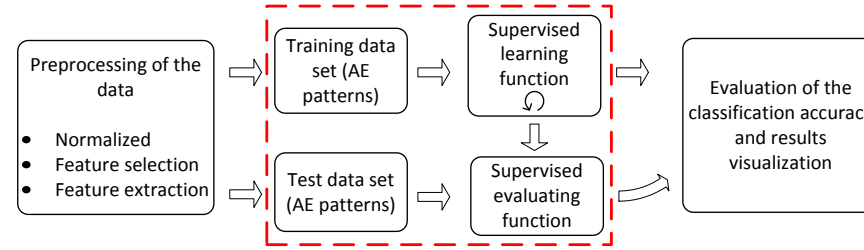


Figure 4.8.: Diagram of supervised classification process.

In this work, a discriminant analysis is employed for supervised classification, leaving out those approaches based on density function estimation.

## 4.5.2. Linear discriminant analysis

### 4.5.2.1. Overview

Linear Discriminant Analysis (LDA) assumes that the decision boundaries are linear. Also, linear discriminant functions are the basic for understanding nonlinear models, which are linear combinations of non-linear functions. There are basically two types of LDA, namely for two-class problems and multiclass problems. Although, the two-class methods are a special case of the multiclass methods. Nevertheless, only the theory for a two-class discriminant analysis is developed in this work, since it is used to form a Support Vector Machine (SVM) algorithm for classification. For further details about LDA, the following works are recommended [Fur97], [McL92], [Web02].

Let us consider a discriminant function with a set of training patterns $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, ..., \boldsymbol{x}_n$. Each pattern is assigned to one of the two classes $\omega_1$ or $\omega_2$. Having this in mind, a weight vector $\boldsymbol{w}$

and a threshold $\omega_0$, which fulfill the following is pursued

$$\boldsymbol{w}^T\boldsymbol{x} + \omega_0 > 0 \Longrightarrow \boldsymbol{x} \in \omega_1 \qquad and \qquad \boldsymbol{w}^T\boldsymbol{x} + \omega_0 < 0 \Longrightarrow \boldsymbol{x} \in \omega_2 \tag{4.5}$$

If we take $\mathbf{z} = (1, \boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, ..., \boldsymbol{x}_n)^T$ is the augmented pattern vector and $\boldsymbol{v}$ is a $(p+1)$-dimensional vector $(w_0, w_1, ..., w_p)^T$, e.g. equation 4.5 can be rewritten as

$$\boldsymbol{v}^T\mathbf{z} > 0 \Longrightarrow \boldsymbol{x} \in \omega_1 \qquad and \qquad \boldsymbol{v}^T\mathbf{z} < 0 \Longrightarrow \boldsymbol{x} \in \omega_2$$

Then, $\mathbf{z}$ could be $(1, \phi_1(\boldsymbol{x}), \phi_2(\boldsymbol{x}), ...., \phi_D(\boldsymbol{x}))^T$, where $\{\phi_i \quad i = 1, 2, ...., D\}$ are $D$ functions of the initial variables and $\boldsymbol{v}$ a $(D+1)$-dimensional vector of weights . The previous mathematical expression can be applied in a transformed feature space. Redefined all samples in class $\omega_2$ in the design set by their negative value (denoted by $y$), the value of $\boldsymbol{v}$ must fulfill

$$\boldsymbol{v}^T y > 0$$

For all $y_i$ corresponding to all $\boldsymbol{x}_i$ in the initial design set

$$y_i^T = (1, x_i^T), \ x_i \in \omega_1 \quad ; \quad y_i^T = (-1, -x_i^T), \ x_i \in \omega_2 \tag{4.6}$$

A solution of $\boldsymbol{v}$ where $\boldsymbol{v}^T y$ is positive for all samples in the design set minimizes the misclassification error. However, this situation is difficult to obtain and almost always there are misclassifications [Fur97]. If $\boldsymbol{v}^T y > 0$ is satisfied for all members of the design set then the data is linearly separable.

However, usually some other criterion is applied for discrimination between two classes. Some criteria are adequate for separable classes and other for overlapping classes. In the following section the Support Vector Machine (SVM) criterion is reviewed.

### 4.5.2.2. Support vector machine (two-class algorithm)

In the previous section, it was stated that linear discriminant functions may be applied to the original data variable or in a shifted feature space defined by a non-linear transformation of the original variables. SVM algorithm operates with the same concept. To achieve this, SVM maps patterns to a higher-dimensional feature space. Doing that, the best hyperplane, with *maximal margin hyperplane* is formed as is illustrated in figure 4.9.

### Linearly separable data

Linearly separable data must be capable of satisfying the following conditions. Having a set of training patterns $\{\boldsymbol{x}_i, \ i = 1, 2, ..., n\}$ assigned to one of the two classes $\omega_1$ or $\omega_2$, and each of them is labeled with $y_i = +1$ or $y_i = -1$ . Then, the linear discriminant function can be written as
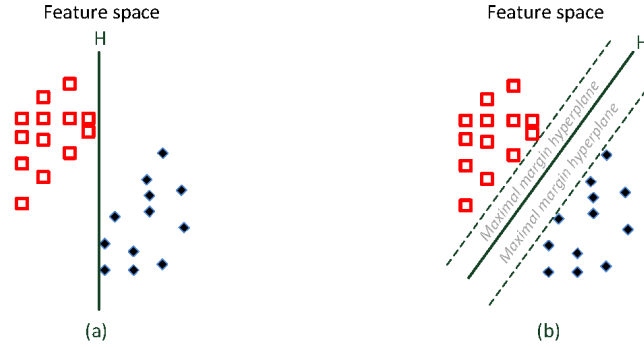
Figure 4.9.: Two linear separable set of data. (a) Separating hyperplane of the data with reduced margin and (b) Separating hyperplane with maximal margins.

$$g(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{x} + w_0$$

The decision rules derived from equation 4.5 becomes

$$\boldsymbol{w}^T \boldsymbol{x} + w_0 > 0 \Longrightarrow x \in \omega_1 \text{ with corresponding label } y = +1$$
$$\boldsymbol{w}^T \boldsymbol{x} + w_0 < 0 \Longrightarrow x \in \omega_2 \text{ with corresponding label } y = -1$$

and represent the same as equation 4.6. Figure 4.9(a) shows a set of separable patterns in the feature space with a separating hyperplane. Although, many possible separating hyperplanes exist, there is only one separating hyperplane with maximal margins (the distances to two parallel hyperplanes on each side of the hyperplane "H" that separate the data, as shown in figure 4.9(b)).

The larger the margin, the better is the generalization error of the linear classifier defined by the hyperplane.

Introducing a margin $b > 0$ and the following condition is obtained

$$y_i = \boldsymbol{w}^T \boldsymbol{x} + w_0 \geq b \qquad (4.7)$$

A solution for which all points $\boldsymbol{x}_i$ are at a distance greater than $b/\boldsymbol{w}$ from the separating hyperplane can be found. A scaling of the terms $b$, $w_0$, and $w$ keeps the distance the same and the equation 4.7 is still satisfied. Therefore, conveniently, $b = 1$ is defined to form the canonical hyperplanes; $H_1 : \boldsymbol{w}^T \boldsymbol{x} + w_0 = +1$, and $H_2 : \boldsymbol{w}^T \boldsymbol{x} + w_0 = -1$ , and thus equation 4.5 is obtained

$$\boldsymbol{w}^T \boldsymbol{x} + \omega_0 \geq +1 \quad for \ y_i = +1 \qquad and \qquad \boldsymbol{w}^T \boldsymbol{x} + \omega_0 \leq -1 \quad for \ y_i = -1 \qquad (4.8)$$

The distance between the *canonical hyperplanes* and the *separating hyperplane* ($g(x) = 0$), is equal to $1/\mathbf{w}$ and is known as the margin (Figure 4.9(b)). Figure 4.10 illustrates two canonical hyperplane and the separating hyperplane for two separable sets of data [Web02]. The patterns that lie on the *canonical hyperplanes* are called support vector.
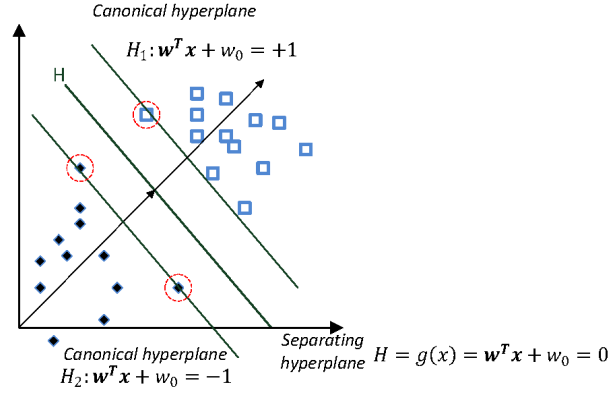


Figure 4.10.: The margin is the perpendicular distance between the separating hyperplane and a hyperplane through the closest patterns (marked by dashed circles). These patterns are called *support vectors* [Web02].

To maximize the margin, a solution that minimizes $|w|$ must be found, subjected to the constraints

$$\mathcal{C}_1 : \quad y_i(\mathbf{w}^T\mathbf{x}_i + \omega_0) \geq 1 \quad i = 1, ..., n \tag{4.9}$$

The Lagrange formalism [Fle88] is an approach for optimization problems with equality and inequality constraints and leads to the optimal form of the objective function $L_p$, formulated by

$$L_p = \frac{\mathbf{w}^T\mathbf{w}}{2} - \sum_{i=1}^{n} \alpha_i(y_i(\mathbf{w}^T\mathbf{x}_i + \omega_0) - 1) \tag{4.10}$$

where $\{\alpha_i = 1, 2, ..., n; \quad \alpha_i \geq 0\}$ are the Lagrange multipliers.

The primal parameters are $\mathbf{w}$ and $w_0$ and the number of parameters is $p+1$ where $p$ corresponds to the dimensionality of the feature space.

To find a solution to the problem of minimizing $\mathbf{w}^T\mathbf{w}$ to constraints equation 4.8 the singular point of the function $L_p$ must be determined, at with $L_p$ is minimized with respect to the weight vector $\mathbf{w}$, a threshold $w_0$ and at the same time maximized with respect to the Lagrange multiplier $\alpha_i$.

Now, differentiating $L_p$ with respect to $\boldsymbol{w}$ and $w_0$ and equaling to zero

$$\sum_{i=1}^{n} \alpha_i y_i = 0 \quad ; \quad \sum_{i=1}^{n} \alpha_i y_i \boldsymbol{x}_i \tag{4.11}$$

and substituting the equation above into equation 4.9, the dual form of the Lagrangian is obtained.

$$L_D = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j \boldsymbol{x}_i^T \boldsymbol{x}_j^T \tag{4.12}$$

This Lagrangian is maximized with respect to the $\alpha_i$ subjected to

$$\alpha_i \geq 0 \quad ; \quad \sum_{i=1}^{n} \alpha_i y_i = 0 \tag{4.13}$$

The equation 4.11 expresses the optimization criterion as inner products of patterns $\boldsymbol{x}_i$. This is an important issue and is used as a basis for the implementation of *non-linear vector machines*, which are discussed later in this chapter. The dual variables are Lagrange multipliers $\alpha_i$, and so the number of parameter, $n$, is equal to the number of patterns.

### 4.5.3. Nonlinear Discrimination Analysis

#### 4.5.3.1. Overview

This approach of supervised methods generalizes the discrimination model by supposing parametric forms for the discriminants functions $\phi$. A discriminant function with the following form is supposed

$$g_j(\boldsymbol{x}) = \sum_{i=1}^{m} w_{ji} \phi_i(\boldsymbol{x}; \mu_i) + w j_0, \qquad j = 1, ..., C \tag{4.14}$$

where, each basic functions, $\phi_i$, has $n_m$ parameters $\mu_i = \{\mu_{ik}, K = 1, ..., n_m\}$, and applying the discriminant rule

$$\text{assign } \boldsymbol{x} \text{ to class } \omega_i \text{ if } g_i(\boldsymbol{x}) = \max_i g_i(\boldsymbol{x})$$

so, $\boldsymbol{x}$ is allocated to the class, whose discriminant function is the largest [McL92]. In equation 4.14, the values $w_{ij}$ and $\mu_i$ are the parameters of the model and m the number of basic functions.

Equation 4.14 has the same form as the generalized equation for linear discriminant function, but more flexibility is allowed in the nonlinear function $\phi_i$. Equation $\phi_i$ can be written as

$$g(\boldsymbol{x}) = \boldsymbol{W}_\phi(\boldsymbol{x}) + w_0$$

where $\boldsymbol{W}$ is a matrix $C \times m$ and $(i,j)^{th}$ components $w_{ij}$, $\phi_{\boldsymbol{x}}$ is the $m$-dimensional vector with $i^{th}$ component $\phi_i\,(x, \mu_i)$ and $w_0$ is the vector $(w_{10}, w..., w_{C0})^T$.

### 4.5.3.2. Nonlinear support vector machine

In the previous section, support vector machine was reviewed as a tool for finding the best hyperplane for linear separable data. The support vector algorithm may be applied in a transformed feature space $\phi(x)$, for a nonlinear function $\phi$. many methods of pattern classification are based on this principle. The transformation of the input features nonlinearly to a space in which linear methods can be applied [McL92].

$$g(\boldsymbol{x}) = \boldsymbol{w}^T \phi(\boldsymbol{x}) + w_0$$

with the following decision rule

$$\boldsymbol{w}^T \boldsymbol{x} + \phi(\boldsymbol{x}) > 0 \implies x \in \omega_1 \; \textit{with corresponding label } y = +1$$
$$\boldsymbol{w}^T \boldsymbol{x} + \phi(\boldsymbol{x}) < 0 \implies x \in \omega_2 \; \textit{with corresponding label } y = -1$$

### 4.5.3.3. Back-propagation neural network

Nonlinear Discriminant functions are a sum of basis functions (nonlinear functions of linear projection of data) and have the form of equation 4.14.

$$g_j(\boldsymbol{x}) = \sum_{i=1}^{m} w_i \phi_i(\boldsymbol{x}; \mu_i) \tag{4.15}$$

where $w_i$ are the weights and $\mu_i$ the parameters of the nonlinear function $\phi_i$ [Web02]. As seen in the previous point (section 4.5.3.2) for SVM, nonlinear functions are not defined explicitly, but implicitly by the specification of a kernel defined in the data space.

**The multilayer perceptron**

Multilayer perceptron (MLP) was formulated by Rumelhart et al. (1986) [RHW86]. Since then is extensively employed in Pattern Recognition problems. Some generalizations of this technique are for example; MLP produces a transformation of a pattern $\boldsymbol{x} \in \mathbb{R}^P$ to an $\dot{n}$-dimensional space according to

$$g_j(\boldsymbol{x}) = \sum_{i=1}^{m} w_{ji} \phi_i(\alpha_i^T \boldsymbol{x} + \alpha_{i0}) + w j_0 \qquad j = 1, ..., n \tag{4.16}$$

The functions $\phi_i$ are fixed nonlinearities, usually identical and taken to be of the *logistic* form

representing the mean firing rate of a neuron as a function of the input current

$$\phi_i = \phi_z = \frac{1}{1 + exp(-z)} \tag{4.17}$$

Consequently, the transformation consists of $i$) projecting the data onto each of $m$ directions of the vectors $\alpha_i$. ii) Transforming the projected data by the nonlinear functions $\phi_z$. $iii$) Forming a linear combination using the weights $wji$.

Figure 4.11 illustrates the MLP process. The patterns are accepted by the input nodes. There are weights associated with the links between the input nodes and the hidden nodes that accept the weight combination $z = \alpha_i^T \boldsymbol{x} + \alpha_{i0}$ and perform the nonlinear transformation $\phi_z$. The output
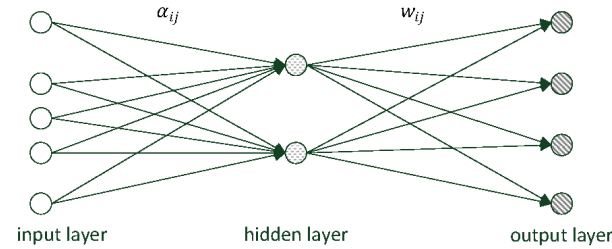


Figure 4.11.: Diagram of Multilayer Perceptron process.

nodes take a linear combination of the outputs of the hidden nodes and deliver these as outputs. MLP is a nonlinear model; the output is a nonlinear function of its parameters and the inputs. A nonlinear optimization scheme must be employed to minimize the selected criterion.

To specify the MLP structure, number of hidden layers, the number of nonlinear functions within each layer and the form of the nonlinear functions must be set. There are two stages of optimization, namely initialization of the parameter values and implementation of a nonlinear optimization scheme.

There are several ways to initialize the weights [FRHE01], for example at small random values with *Random Initialization*. Other two approaches are *Pattern Classifier Initialization* [Bed89] and *Independence model Initialization*.

## 4.5.4. Evaluation of the supervised classification

In a pattern classification the accuracy of the process is the degree of effectiveness of a classification function to assign labels to the input patterns when it is compared with the real labels.

Accuracy can be considered as a statistical measurement of how good a classification function identifies the labels of patterns.

In Supervised classification, the accuracy is the proportion of patterns labeled correctly with respect to the total input test patterns.

$$Classification\ accuracy = \frac{number\ of\ pattern\ classified\ correctly}{Number\ of\ Pattern\ to\ be\ classified} \times 100$$

where the number of pattern classified correctly is obtained comparing the real class labels with the predicted class label.

## 4.6. Unsupervised Classification or Clustering

### 4.6.1. Introduction

Clustering is used to explore the data in order to find groups of pattern and use them in a supervised method. Existing methods operate based on dissimilarity matrices or measurements on individuals. Each of them models the data structure. The aim of clustering is to discover a structure in the data. Ideally, patterns of the data should be close or similar in each group and dissimilar from other patterns in other groups. Clustering explores the data and searches for groups to label them. After clustering, an identifiable structure of the data is available, which can be used to hypothesize about the data, and then these hypotheses should be tested with new data (test data) [Bis06].

Figure 4.12 illustrates the clustering process. In clustering the AE Patterns are classified without an external reference. As in supervised techniques, the accuracy of the unsupervised classification method is evaluated and visualized.
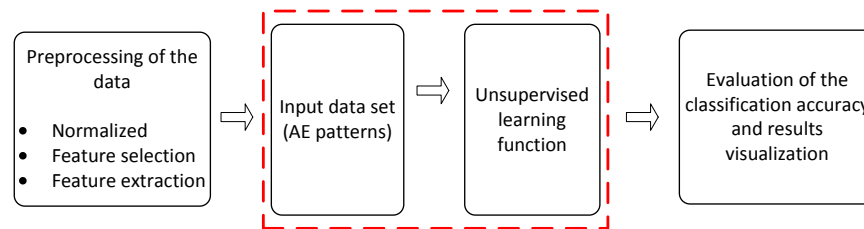


Figure 4.12.: Diagram of Unsupervised classification process or Clustering.

### 4.6.2. Sum of square methods

The Clustering methods reviewed in this chapter pertain to the sum of square methods. These methods search a partition of the data that maximizes a pre-established clustering method

[JD88]. This clustering criterion is based on the within-class and between-class scatter matrices. All these methods differ in the way the clustering criterion is optimized. Nevertheless, all the clustering methods seek to solve the same, that is, given a set of $n$ data samples, the algorithm must divide the data into $g$ groups or clusters so that the criterion is optimized.

**Clustering criteria**

Being $x_{1,\ldots,x_n}$, the $n$ data samples. The sample covariance matrix, $\hat{\sum}$, given by

$$\hat{\sum} = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i - \boldsymbol{m}) - (\boldsymbol{x}_i - \boldsymbol{m})^T$$

with $\boldsymbol{m} = \frac{1}{n}\sum \boldsymbol{x}_i$ , the sample mean. Assuming $g$ clusters, *within-class scatter matrix (cross-products matrix)* is given by

$$S_W = \frac{1}{n}\sum_{j=1}^{g}\sum_{i=1}^{n} z_{ji}(\boldsymbol{x}_i - \boldsymbol{m}) - (\boldsymbol{x}_i - \boldsymbol{m})^T$$

the sum of sum of square and cross-products matrices over the $g$ clusters [JD88]. $z_{ji=1}$ if $\boldsymbol{x}_i \in$ group $(j,0)$ if not, $\boldsymbol{m}_j = \frac{1}{n_j}\sum_i^n z_{ji}\boldsymbol{x}_i$ is the mean of the cluster $j$ and $n_j = \sum_{i=1}^{n} z_{ji}$, the number in cluster $j$. The *between-class scatter matrix* is given by

$$S_B = \hat{\sum} - S_W = \sum_{j=1}^{g} \frac{n_j}{n}(\boldsymbol{x}_i - \boldsymbol{m}) - (\boldsymbol{x}_i - \boldsymbol{m})^T$$

The three clustering methods described in this chapter are based on the univariate functions of the matrices described above.

### 4.6.3. K-means clustering

The objective of this algorithm is to divide the AE patterns in $k$ clusters in order to minimize the within- group sum of squares ($S_W$ criterion). The basic form of the K-means clustering is based on alternating two procedures. In the first one, the patterns are allocated to the groups, usually to the group whose mean is closest (in a Euclidian logic). In the second one, the calculation of new group of means based on the allocations takes place. The iterations finish, when no more allocations of patterns to another group will reduce the with-group of sum of square.

Assignations are done pattern by pattern, rather than pass the total number of patterns. A pattern $\boldsymbol{x}_i$ in a group $l$ is allocated to a group $r$ if

$$\frac{n_l}{n_l - 1}d_{il}^2 > \frac{n_r}{n_r + 1}d_{ir}^2$$

with $d_{il}$ is the distance to the $l^{th}$ centroid and $n_l$ is the number in group $l$. The highest reduction in the sum-squared error is obtained by selecting the group for which $n_r d_{il}^2/(n_r+1)$ is a minimum [JD88]. Now, developing a nonlinear optimization, the within-groups sum of squares criterion can be expressed by the following equation.

$$T_r(S_W) = \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{g} z_{ki} \sum_{j=1}^{p} (\boldsymbol{x}_{ij} - \boldsymbol{m}_{kj})^2 \tag{4.18}$$

where $\boldsymbol{x}_{ij}$ represents the $j^{th}$ coordinate of the $i^{th}$ pattern $(i = 1, ..., n)$; $j = 1, ..., p$, $\boldsymbol{m}_{kj}$ represents the $j^{th}$ coordinate of the mean of the $k^{th}$ cluster and $z_{ki} = 1$ if the $i$th pattern pertains to the $k^{th}$ cluster and 0 if not.

### 4.6.4. Fuzzy C-means clustering

The partitioning method of the K-means presented in the section 4.6.3 has the property that each pattern pertains to one cluster only. Though, a mixture model can be considered as providing degrees of cluster association. Actually, the first works on fuzzy clustering was strongly associated to multivariable mixture models. Fuzzy means clustering methods are based on the idea that pattern are permitted to pertain to all clusters with different degree of association. Dunn presented the first work on K-means [Dun73]. The fuzzy C-means algorithm tries to find a solution for parameters $y_{ji}(i = 1, ..., n; \quad j = 1, ..., g)$ for which

$$J_r = \sum_{i=1}^{n} \sum_{j=1}^{g} y_{ji}^r |\boldsymbol{x}_i - \boldsymbol{m}_j|^2 \tag{4.19}$$

is minimized subjected to the following restrictions

$$\sum_{j=1}^{g} y_{ji} = 1 \qquad 1 \leq i \leq n$$

$$y_{ji} \geq 0 \qquad i = 1, , , .n; \quad j = 1, ..., g$$

The parameter $y_{ji}$ is the degree of *association or membership* function of the $i^{th}$ pattern with $j^{th}$ cluster. In equation 4.19, $r$ represents a scalar called the *weighting exponent* which controls the *fuzziness* of the resulting clusters $(r \geq 1)$ and $\boldsymbol{m}_j$ represents the centroid of the $j^{th}$ cluster

$$\boldsymbol{m}_j = \frac{\sum_{I=1}^{n} y_{ji}^r \boldsymbol{x}_i}{\sum_{i=1}^{n} y_{ji}^r} \tag{4.20}$$

If $r = 1$, it results in the same problem as the nonlinear optimization scheme. In this event, it is known that a minimum of equation 4.19 provides values of the term $y_{ji}$ that are either 0 or 1.

### 4.6.5. Vector quantization clustering

Vector Quantization (VQ) is an application of the two clustering methods presented before. Different clustering techniques are reformulated in VQ literature. During clustering a distortion measure is optimized which is often based on the Euclidian distance. For further fundamentals about this clustering technique, it is recommendable to read the work of Gray [Gra84].

VQ encodes a $p$-dimensional pattern or vector $\boldsymbol{x}$, as one from a *codebook* of $g$ vectors, $z_1, z_2, ..., z_g$, called the *code vectors*. The objective of VQ is basically to perform data compression. Conceptually a vector quantiser is composed of two components, namely encoder and decoder as is illustrated in figure 4.13. The first step is to encoder a input vector (pattern), $\boldsymbol{x}$, to a scalar
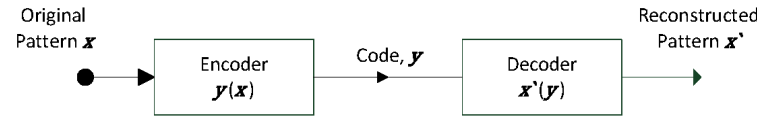


Figure 4.13.: Process of encoding-decoding in vector quantization technique [Gra84].

variable, $y$, taking discrete values $1, ..., g$. Then the index is transmitted and finally the reverse operation is done, that is an approximation $\boldsymbol{x}$` of the original vector is reproduced. The decoding process is a mapping from the index set $\mathbb{I} = 1, ..., g$ to the codebook $\mathbb{C} = z^1, ..., z_g$. Given the set of training samples, the problem of the codebook design is to determine the codebook entries. From a clustering point of view, the problem of codebook design can be associated to clustering the data and then choose a representative vector (or pattern) for each cluster. This pattern could be the means for example and they constitute the entries in the codebook and are indexed with integer values. Therefore, the code vector for a given input vector $\boldsymbol{x}$ is the representative vector, $\boldsymbol{z}$, of the cluster to which $\boldsymbol{x}$ pertain. The degree of association of pattern with a cluster may be determined on a nearest-to-cluster basis. The error can be calculated as the distance between the representative vector $\boldsymbol{z}$ and the vector to be associated to a group or cluster.

The problem in VQ is to determine the set of representative vectors or codebook that characterize a data set. This can be achieved by selecting the set of vector for which a distortion measurement (error) between an input vector or pattern, $\boldsymbol{x}$, and its reconstructed or quantized vector, $\boldsymbol{x}$`, is minimized. Various ways of calculating the distortion measures have been proposed, but the most common is based on the squared error measure, producing average distortion, $D_2$,

$$D_2 = \int p(\boldsymbol{x})d(\boldsymbol{x},`)dx = \int p(\boldsymbol{x})\|\dot{\boldsymbol{x}}(y(x)) - \boldsymbol{x}\|^2 \qquad (4.21)$$

where $p(x)$ is the PDF over samples $\boldsymbol{x}$ for training of the vector quantiser and $\|\cdot\|$ represents the norm of a vector. Other types of distortion measurement are given in table 4.2

Table 4.2.: Distortion measurements used in VQ [Gra84].

| Type of norm | $d(\boldsymbol{x},\boldsymbol{x}^{\grave{}})$ |
|---|---|
| Euclidian, $L_2$ | $|\grave{\boldsymbol{x}} - \boldsymbol{x}|$ |
| Minkowski | $\max_{1 \leq i \leq p} |\grave{\boldsymbol{x}}_i - \boldsymbol{x}|$ |
| Quadratic | $(\grave{\boldsymbol{x}} - \boldsymbol{x})^T \mathbf{B}(\grave{\boldsymbol{x}} - \boldsymbol{x})$ |

### 4.6.6. Evaluation of the unsupervised classification

All clustering techniques will divide the data set of patterns even if no natural clusters are present in the data. Also different techniques produce different classifications. There are three main classes of identifying real clusters [Gor96], namely Poisson Model [Boc85], Unimodel model [Boc85] and Random dissimilarity matrix [Lin73].

Also a measurement of the similarity between two clusters can be obtained using the Rand Index (RI) [Ran71]. RI is commonly used to validate cluster performance; it basically compares two partitions, one from the clustering technique and the other from one specific (external) criterion. The internal structure of the Rand Index can be modified as a function of the specific data of set of data to be classified; this is called the *adjusted Rand Index*. Mathematically, RI is associated with accuracy.

Having a set of input Patterns $\mathcal{S} = (\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n)$. Two partitions of $\mathcal{S}$ are given by $\mathcal{X} = \{X_1, X_2, ..., X_3\}$ and $\mathcal{Y} = (Y_1, Y_2, ..., Y_m)$, containing each partition, $j$ and $m$, clusters of the input data $\mathcal{S}$ respectively. Defining the relationships

1. $a$, is the number of patterns in $\mathcal{S}$ that are present in same clusters in $\mathcal{X}$ and in the same clusters in $\mathcal{Y}$.

2. $b$, is the number of patterns in $\mathcal{S}$ that are in different clusters in $\mathcal{X}$ and in different clusters in $\mathcal{Y}$.

3. $c$, is the number of patterns in $\mathcal{S}$ that are present in same clusters in $\mathcal{X}$ and in different clusters in $\mathcal{Y}$.

4. $d$, is the number of patterns in $\mathcal{S}$ that are in different clusters in $\mathcal{X}$ and in the same clusters in $\mathcal{Y}$.

The RI can be mathematically defined by

$$RI = \frac{a + b}{a + b + c + d} = \frac{a + b}{\binom{n}{2}} \tag{4.22}$$

where $(a + b)$ express the degree of correspondence between $\mathcal{X}$ and $\mathcal{Y}$ and $(c + d)$ represent the degree of divergence between $\mathcal{X}$ and $\mathcal{Y}$.

## 4.7. Summary

In this chapter the main topic of Pattern Recognition Technique is reviewed. The main applications of PRT in the mining industry are presented, which are principally in the automation field. Also, the association between PRT and AET is explained. The process of Acoustic Emission Testing is assimilated in the process or cycle of Pattern Recognition (Table 4.1). The concept of pattern is explained from an AET point of view and the relation between PRT and Neural Network is presented.

For the preprocessing of the data; normalized, feature selection and feature extraction methods are presented.

**Principal Components Analysis** (PCA) forms a *linear* transformation based on matrices of first and second order statistics. This technique is based on correlation or covariance matrix.

**Linear Discriminant Analysis** (LDA) is a supervised linear supervised technique for data dimension reduction. LDA is based on the Fisher`s Criterion.

**Multidimensional Scaling** (MDS) refers to a range of technique, which analyses dissimilarities matrices and create coordinates of points in a reduced dimension.

Three approaches for supervised classification methods (linear and nonlinear) are presented. One technique for linear discrimination is reviewed, namely **support vector machine** algorithm for two classes (quadratic algorithm).

A **nonlinear two-class** classification approach for **Support Vector Machine** is also reviewed. Nonlinear SVM is reliable in many real applications. Once the kernel is fixed, nonlinear SVMs have only free parameter.

**Back-propagation neural networks** technique takes the sum of univariate nonlinear functions $\phi$ of linear projections of the data, $\boldsymbol{x}$, onto a weight vector $\alpha$, and use this information for classification. The non-linear functions are an imposed form and a weighted sum is form (equation 4.16).

The Evaluation of the classification accuracy of a supervised classification method can be evaluated by the **classification accuracy** (section 4.5.4).

Also, three approaches for unsupervised classification methods are presented.

**K-means clustering** is extensively used in pattern recognition and forms the basis for other clustering techniques. It codes the data in a pattern pertaining to one cluster only.

**Fuzzy C-means** permits a pattern pertaining to more than one cluster, depending on an association function, the clusters resulting from this algorithm are faintly overlapped. The degree of overlapping is controlled by a parameter specified by the user.

**Vector quantization clustering** technique produces tree-structured algorithms. The advantage of this type of clustering is its fast coding.

The Evaluation of clustering performance in an unsupervised classification method can be evaluated by the **Rand Index**, (section 4.6.6).

# 5. Experimental cases of study

## 5.1. Introduction

Appearing in the middle of previous century, Acoustic Emission Technique (AET) is a Non-Destructive Test (NDT) developed from the seismology. Different research works have been carried out in a variety of fields using this technique. AET involves the capture of microsismic signals, which travel through a material using contact AE sensors.

As it was reviewed in section 2.6, nowadays AET can be applied to a variety of process in the mining industry. The most common application of this technology takes place in the condition monitoring of machine components. Nevertheless, since a decade ago the use of AET has been extended to different areas in mining.

In this work, two experimental applications are reviewed, namely AET in column flotation cells and AET in rock cutting.

In case of column flotation, AET is used for monitoring the bubble activity and bubble size, as a means of improving the efficiency of the column flotation process. Using AET in combination with PRT, small changes in air flow can be detected, making it possible to monitor the bubble activity inside the column flotation.

In rock cutting, AET is used to identify the rocks being cut as means of automation in the operation of underground mining and tunneling. AET together with PRT is used to classify AE signal coming from different cutting process of different rocks.

The experiments were carried out in a laboratory scale, using AE measurements to analyze the processes. Pattern Recognition Technique is used in combination with classic and advanced signal processing to characterize the collected AE signals.

The AE signals were collected by a DAS. Then a signal preprocessing of the raw data was carried out, namely normalization, feature selection and feature extraction.

Unsupervised and supervised classification algorithms are employed to classify the AE data. The accuracy of these algorithms are evaluated and compared.

## 5.2. AET in column flotation cell

### 5.2.1. Introduction

Flotation is a physico-chemical separation process that utilizes the difference in surface properties of valuable minerals and the unwanted gangue mineral [WNM06]. This is the most commonly used method for separating valuable minerals from non-valuable materials (gangue). In this process, slurry is fed into a tank (flotation cell) and a flow of air (bubbles) is introduced at the bottom with the aim of collecting and lifting particles of valuable minerals to the surface of the tank. Consequently, a froth layer is created with the help of chemical reagents at the top of the tank. This froth layer is recovered from the tank, forming concentrate, and the gangue at the bottom of the tank is expelled.

The selective attachment of the valuable mineral particles to air bubbles is possible due to the differences in physico-chemical surface properties of the particles. This process can be applied only to fine particles, otherwise the adhesion between the bubble and the particle will not be strong enough and the particle will drop from the bubble. Moreover, industrial flotation cells are designed for a specific particle size range. Consequently, it can be stated that for each flotation system of a certain size, there is a distribution of bubbles at which the performance is close to the ideal [FD91].

Two main types of industrial flotation cells can be distinguished, mechanical and column cells. Mechanical cells are agitated by an impeller located at the bottom of the cell, whereas column cells work without an impeller. Air is injected into both types and the air flow can be controlled to regulate their performance.

The performance of the flotation column process is mainly determined by the primary variables, namely concentrate grade and recovery. The bubble characteristics in flotation cells are at the core of the problem to increase productivity in the flotation process [FD91]. This is a secondary variable that can be controlled to optimize the process. The ideal bubble size depends on the size of the particles in the slurry fed into the flotation cell.

In column cells, Acoustic Emissions (AE) can be detected and analyzed to understand the gas dispersion properties. First of all, AE are present during the bubble formation close to the porous sparger [MFD⁺01]. Next, when the bubbles rise, they collide with one another. There are also signals from bubbles distorted by hydrodynamic forces (caused by ascending of other bubbles in the vicinity of the studied bubble)[CL92]. Finally, AE is emitted during the bursting of the bubbles on the surface and due to coalescence of the bubbles in the froth layer [Lub89].

In flotation columns, the optimization of the flotation process is not easy to achieve because its performance depends on the main variables, namely froth properties, gas rate, bias rate, gas hold up and bubble distribution. However, secondary variables, which can easily be regulated and measured in real-time (e.g. air flow) can be used to optimize the process [Sat98].

The aim of this study is to recognize differences or characteristics in the acoustic emission

signals emitted by bubbles with different sizes when they form and rise from the bottom to the surface of a flotation column in a laboratory-scale apparatus. For this purpose, a broadband Acoustic Emission sensor and an accelerometer sensor were used. Different stainless steel rods were placed along the column cell, acting as waveguides, which transmit the signals from the inside of flotation column to the sensors. This study considers only two phases, namely water and air. For analysis of a system with three phases, (water-air-solid) more variables should be taken into account [ZSC12].

Also, the problem of turbulence is addressed and analyzed by means of pattern recognition technique. In column cells, the increase of the air flow produces turbulence. Initially, with a moderate turbulence, an increment in the flotation rate is produced due to an increase in the collision probabilities between the bubbles and the mineral particles. Then, with further increase in turbulence, the flotation rate decreases due to detachment of the mineral particles from the bubble [PDFR02].

A specific air flow cannot be recommended because the turbulence degree depends on many factors such as the altitude at which the flotation cell is located, its height, percentage of solids in the slurry and other factors related to the kinetics of the flotation. Pattern recognition is applied on different signals from different air flows to identify the bubble activity changes to monitor the presence of excessive turbulence.

As unsupervised classification *K-means clustering* is used in combination with PCA and LDA data dimensionality reduction techniques.

For supervised classification; Back-propagation neural network algorithm is used to classify the AE patterns from different air flow (bubble activity).

The use of these contact sensors was motivated by the impracticality of other methods for the characterization of bubbles outside of the laboratory, including techniques like high speed photography or active ultrasonic sensors [Nis72].

### 5.2.2. Theory

Acoustic waves are energy in the form of vibrational waves and are related to the traditional field of sound and vibration. Sound waves are longitudinal when they travel across gases, liquids and solids [Kin82]. In solids, acoustic waves can also travel in a transverse manner, that is to say, perpendicular to the direction of the waves, e.g. in stainless rods. Due to the low viscosity of liquids and gases, transversal waves do not occur in these phases.

In 1933 Minnaert [Min93] stated that the frequency of the sound emitted by a bubble when it is formed by a nozzle is related to its size by

$$f = \frac{1}{2\pi r}\sqrt{\frac{3\kappa P}{\rho}} \qquad (5.1)$$

where, $f$, is the frequency in Hz, $r$, the radius of the bubble (assuming a spherical bubble), $\kappa$, the ratio of specific heats for the gas, $P$, the absolute liquid pressure and, $\rho$, the density of the liquid.

When a bubble rises to the surface of the column flotation, it displaces the water in front of it. Consequently a wake is formed during its way to the surface, producing a pressure wave. In addition, it experiences a volumetric oscillation caused by shock excitation due to the ascension of other bubbles (turbulence pressure fluctuations). As a result, the bubble oscillates producing an acoustic pressure wave which can be measured. The resonance frequency for a given bubble does not experience a dramatic fluctuation with depth [SCWL99] and this parameter can be considered a second order effect in tank with depths below 10 m.

This acoustic pressure wave can be attenuated by other bubbles. The wave can be scattered or transformed into heat; hence the distance at which the wave can be detected by a sensor is reduced [FD91]. Moreover, a high air flow causes bubbles to come into contact, producing coalescence of the bubbles [KMP07].

**Turbulence in flotation Cells**

It is well known that turbulence is an important factor in the performance of the column flotation. Turbulence has an important effect on the recovery and grade of the concentrate in column flotation. The hydrodynamics theory is used to understand the fluctuation of the turbulence [Fal87].

Turbulence kinetic energy, $q$, is formed by means of the fluctuations of the flow in the three dimension of the space [WP89]

$$q = \frac{1}{2}(u_i^2 + u_y^2 + u_z^2)$$

In practical applications this is difficult to measure, because many other factors take part in the flotation process. Also the air flow that must be introduced into a column flotation cell is conditioned by the absolute pressure at the overflow and the absolute pressure at the bottom of the column cell, percentage of solids in the slurry and kinetics factor. In this context, a method to measure qualitatively the air flow and its changes may be provided by the AE signals emitted by the air flow. Using Pattern Recognition technique, the features of each flow can be classified and the air flow changes detected.

To identify each air flow, three features are extracted from the AE signals, namely *Root Mean Square* (RMS), *Variance* ($\sigma^2$) and *Entropy* ($E$) (see chapter 3).

### 5.2.3. Experimental methods

A laboratory flotation column was used for the study of AE emitted by bubbles. An acrylglass column with a diameter of d = 200 mm and a height of H = 2000 mm was fitted with stainless

rods ($\varnothing = 22$ mm and l $= 250$ mm) every 30 cm and used as test bench, as it is shown in figure 5.1.

Seven stainless rods were placed along the column flotation, crossing it (figure 5.1) to investigate whether there are changes in the signals emitted by bubbles during their ascent. These rods act as waveguides capturing the signals from the inside of column flotation in an attempt to reduce the attenuation of the signals. Moreover, two short rods (l $= 25$ mm) were placed with the aim of analyzing the advantage or disadvantage of using long rods as waveguides (figure 5.2 and figure 5.3). These short rods are labeled rod 10 and rod 11 in figure 5.1. It must be noted that waveguides could introduce some further complexity for the frequency analysis of AE waves [GO10]. Air was introduced at the bottom of the flotation column and bubbles were produced
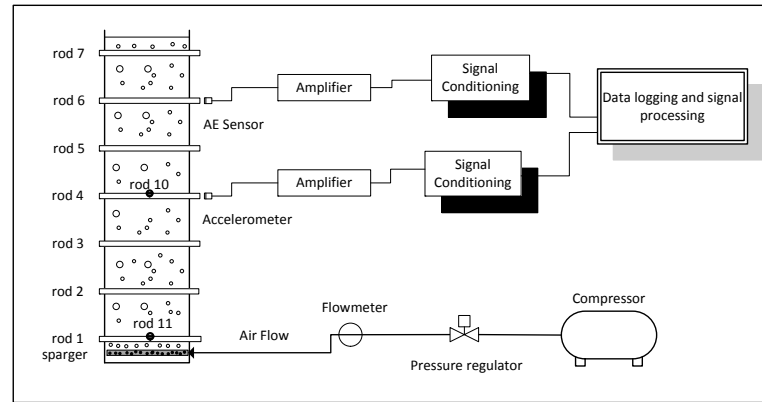


Figure 5.1.: Schematic diagram of the test-bench and the equipment.

with a porous tube capable of sparging bubble diameters from 1 mm to 9 mm approximately depending on the air flow. The air flow was regulated under controlled pressure to produce different bubble sizes [Oco90] and bubble flows. Tests were conducted with five different flows to analyze the signals emitted by the ascending bubbles of different sizes (Table 5.1). These air flows are relatively low in order to avoid turbulence.

Table 5.1.: Air Flow Rates.

| Air Flow | Air flow volume [l/min] |
|---|---|
| A | 0,4 |
| B | 0,8 |
| C | 1,5 |
| D | 3 |
| E | 5 |

The acquisition system is shown in figure 5.1; it consists of two contact sensors, which are described in table 5.2. The AE sensor was used to analyze the signals in the ultrasound domain with a high sampling rate. In addition, an accelerometer was used to analyze the signals in the sound domain. Every sensor picks up signals from two different frequency bands to improve the analysis.

Table 5.2.: Description of the used sensors.

| Type of Sensor | Model | Freq. Range | Freq. Peak |
|----------------|----------|-------------|------------|
| AE Sensor | VS160-NS | 100-450 kHz | 150 kHz |
| Accelerometer | 622A01 | 3-9000 Hz | 20 Hz |

After the AE signals and accelerometer signals were obtained, they were amplified. Next, AE signals were recorded at a sampling rate of 1MHz and prepared for processing. In the same way accelerometer signals were recorded at a sampling rate of 50 kHz. The signals were recorded and processed with the software LabVIEW.
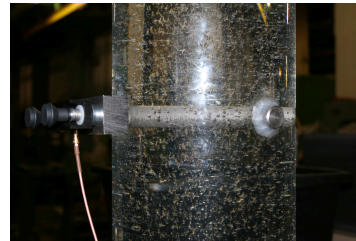


Figure 5.2.: AE Sensor placed in position 4 and short rod 10.



Figure 5.3.: Accelerometer sensor placed in position 4 and short rod 10.

### 5.2.4. Results

#### 5.2.4.1. Time domain analysis

**Attenuation of signals**

After analyzing the data from the AE sensor and the Accelerometer, it is clear that AE sensors are more suitable for a wave-form analysis in time domain in comparison with the accelerometer. Figure 5.4 shows the AE signal emitted by bubbles produced by introducing an air flow C and picked up from position 4 during 3 s. Figure 5.5 shows a signal obtained using the accelerometer with the same flow and in the same position.

The AE signal in figure 5.4 shows the peaks produced by the bubbles when they rise at the top of flotation column, passing by rod 4. This signal was picked up at position 4, with help of a waveguide which enable to diminish the attenuation suffered by the signal, when they travel across the water.
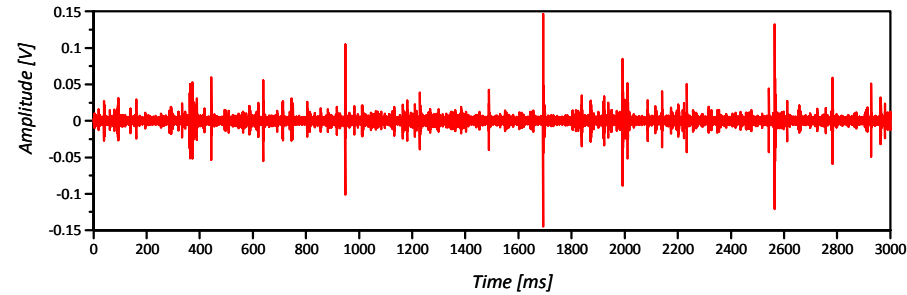


Figure 5.4.: AE signal at position 4 and flow C.
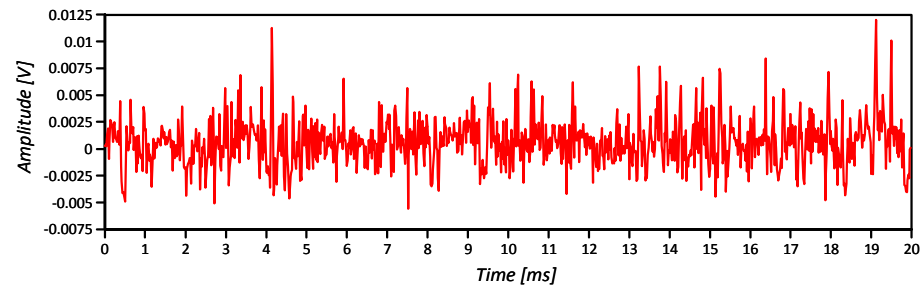


Figure 5.5.: Accelerometer signal emitted by bubbles at position 4 and flow C.

On one hand, the accelerometer signal in figure 5.5 seems to show in general more activity than the AE signal, but on the other hand it does not show information about bubbles at first

sight. Despite the fact that the accelerometer signal in Figure 5.5 has a duration of just 20 ms, it represents the signals that could be found in this research using the accelerometer and the waveguides.

Additionally, Figure 5.6 shows an AE signal and figure 5.7 shows an accelerometer signal, both of them picked up at position 10 (without stainless rod as a waveguide), using the same air flow, that is to say, air flow C.
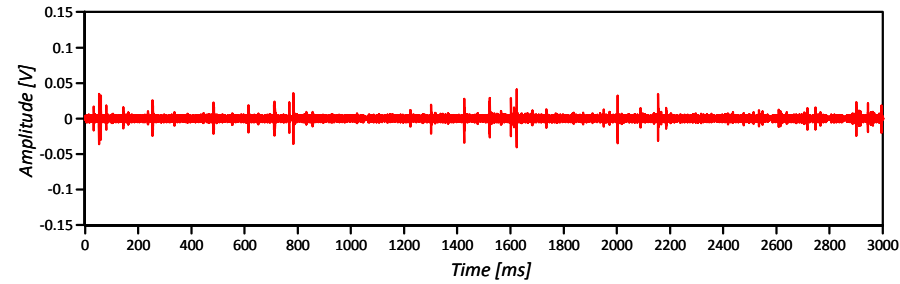


Figure 5.6.: AE signal emitted by one bubble at position 10 (without waveguide) and flow C.

From figure 5.4 and figure 5.6 it can be seen that there are differences between an AE signal picked up with waveguide and without waveguide. Without the help of a waveguide, the AE signal is attenuated before it reaches the AE sensor.
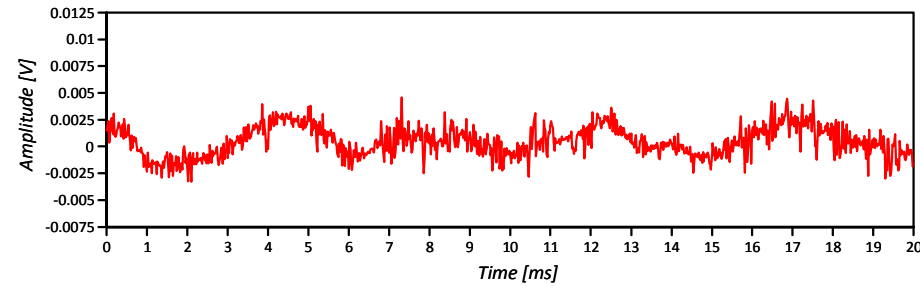


Figure 5.7.: Accelerometer signal emitted by bubbles at position 10 (without waveguide) and flow C.

Figure 5.7 shows that the attenuation suffered by the accelerometer signal picked up without waveguide is high in comparison to an accelerometer signal picked up with waveguide (figure 5.5). This suggests that the signal is too weak to reach the accelerometer across the water due to the scattering and attenuation produced by the bubbles in the vicinity of the signal source.

**Signal analysis from different positions**

By changing the position of the sensor it is possible to obtain signals from various AE sources corresponding to the process. First of all, signals obtained at the bottom of the column, that is to say, position 1 contain information about the formation of the bubbles that take place in the porous tube. The signals emitted by the bubbles during their formation relate to equation 5.1 and can be used to estimate the bubble size. At this stage, the sound signal is caused when air bubbles detach from the porous sparger. At this moment the bubbles oscillate at their natural frequencies until they reach equilibrium. Moreover, at high flow rates turbulence noise and collisions between bubbles are present, and equation 5.1 is no longer valid. Figure 5.8 shows an AE signal from a bubble captured in position 1 and flow B.
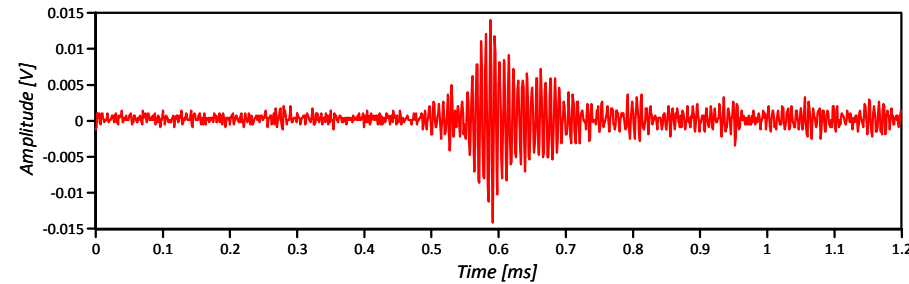


Figure 5.8.: AE signal emitted by one bubble produced using flow B in position 1.

Secondly, when the bubbles float to the top of the flotation column (from position 2 to position 6) they displace the volume of water in front of them, generating a wake that produces hydro-dynamic forces, which distort others bubbles in the vicinity. Consequently, the bubbles start to oscillate again generating a sound wave. However, these waves are weak and most of them are dissipated by other bubbles, hence the elastic pressure wave cannot reach the AE sensor. Accordingly, only some bubbles can be detected by the AE sensor (figure 5.4 and figure 5.6).

Finally, at the top of the flotation column, the bubbles reach the surface of the water and bubble bursting is produced (position 7). Bubble bursting is supposed to emit sound signals. However it was not possible to recognize any waveform in the time domain with either the AE sensor, or the accelerometer. The frequency analysis using the accelerometer shows that the bursting of bubbles emits signals with frequencies between 50-300 Hz.

### 5.2.4.2. Frequency domain analysis

Theoretically, acoustic waves emitted by small bubbles as in the flotation process contain frequencies in the sound frequency range [Min93]. This is corroborated by the results of the present research. Accordingly, with a frequency analysis of AE signals (with frequencies above

20 kHz it was not possible to recognize patterns in the spectrum in order to characterize the AE signals. However, the spectrum analysis of the accelerometer signals is able to identify the bubble flow of the flotation column.

Figure 5.2.8 shows the spectrum emitted by bubbles introduced at air flow D at position 1 and the spectrum obtained when no air flow is introduced (thermal noise). This graph should help to convey an idea of the spectrum generated by bubbles and shows the weakness of the emitted signals.

Using equation 5.1, it is possible to calculate the diameter of a bubble as a function of its natural oscillation frequency, emitted when it is forming. To take an example, a $r = 0.5$ mm bubble, $\kappa = 1.2$ considering the average of the possible values for air, $P = (101325 + 17658)$ Pa considering that the bubble is formed under 1,8 m of water and water density $\rho = 1000$ kg/m3, the frequency emitted by the bubble is 6588 Hz. Now considering a $r = 4.5$ mm bubble, under the same conditions, the frequency emitted by it is 732 Hz.

These frequencies are considered as lower and upper limits of the range of frequencies emitted by the bubbles in this study [732 - 6588] Hz (in accordance with the sparger utilized). Different bubble sizes are formed during the bubble sparging with the aim of producing a wide frequency range. The frequencies outside of this range can be considered as signals from turbulence produced by the bubbles, collisions between bubbles as well as due to the sensitivity of the sensor at 20 kHz (frequency peak). Analyzing the same flow and conditions as in figure 5.9,
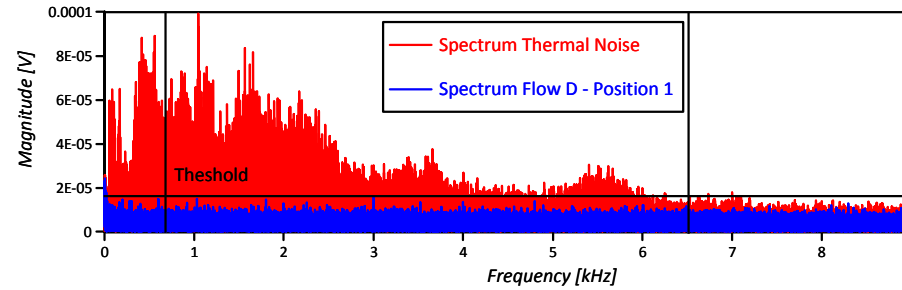


Figure 5.9.: Spectrum using accelerometer sensor in position 1 without bubbles (thermal noise) and with an air flow D.

but this time, extracting the power spectral density, the difference is even clearer. Figure 5.10 shows the power spectral density (PSD) from an accelerometer sensor placed in position 1 and measuring an air flow D. From figure 5.10, it is clear that it is possible to use a threshold to identify the bubbles. Since the signals are in the sound domain (frequency domain: 0-20 kHz), the background noise depends strongly on where the column cell is located. However, using a contact sensor, it is much less likely for background noise to contaminate the signals from the bubbles.
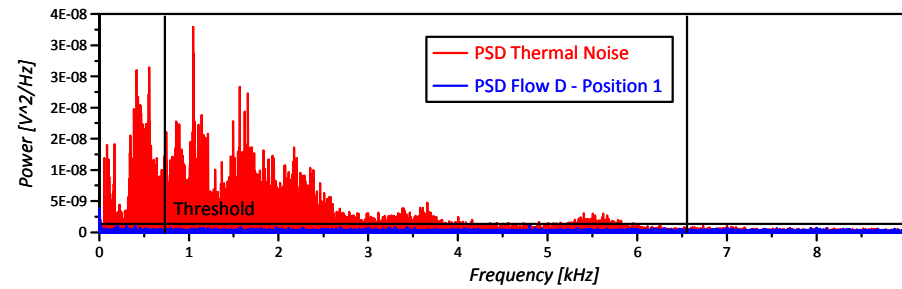
Figure 5.10.: Power Spectral Density using accelerometer in position 1 without bubbles (thermal noise) and with an air flow D.

The frequency emitted by the bubbles is delimited by the vertical lines in figure 5.9 and figure 5.10. Using a band-pass filter on the signals, it is possible to isolate the bubble frequencies and use them as way to identify the bubble size.

A band-pass filter between 732 - 6588 Hz is used to filter out the frequencies from other phenomena such as turbulence (see figure 5.9). In this case, the power signal amplitude of electrical noise floor or thermal noise is less than 1x10-9 V2/Hz and must be applied as a threshold for all Power Spectral Density calculations (see figure 5.10). Accordingly, figure 5.11, figure 5.12 and figure 5.16 shows the PSD analysis corresponding to air flows B, C, and E respectively, all of them taken with the accelerometer placed in position 1 and filtered. Figure
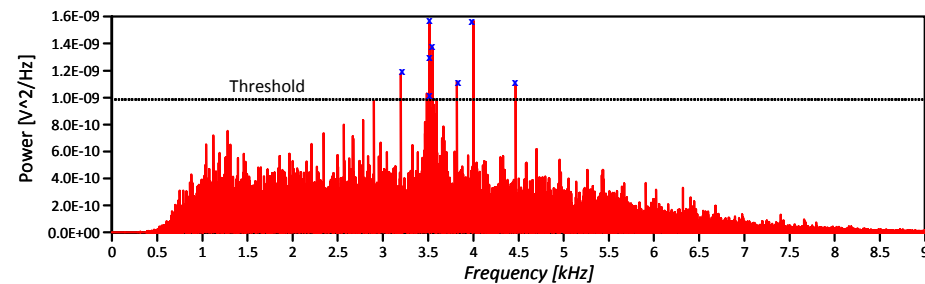


Figure 5.11.: Power Spectral Density using accelerometer in position 1 and air flow B.

5.11 shows 8 peaks corresponding to bubble frequencies, where the power is related to the vicinity to the accelerometer sensor at which the signal is emitted from the bubble.

Table 5.3.: Bubble diameters as a function of the bubble frequencies (accelerometer in position 1 and flow B).

| Frequency [Hz] | Power $[V^2/\mathbf{Hz}]$ | Bubble diameters [mm] |
|---|---|---|
| 3200 | $1{,}17 \times 10^{-9}$ | 2.06 |
| 3484 | $1{,}03 \times 10^{-9}$ | 1.89 |
| 3515 | $1{,}55 \times 10^{-9}$ | 1.88 |
| 3518 | $1{,}29 \times 10^{-9}$ | 1.87 |
| 3549 | $1{,}37 \times 10^{-9}$ | 1.86 |
| 3818 | $1{,}12 \times 10^{-9}$ | 1.73 |
| 4000 | $1{,}57 \times 10^{-9}$ | 1.65 |
| 4471 | $1{,}09 \times 10^{-9}$ | 1.47 |

Analyzing table 5.3 it can be observed that for the flow B the predominant bubble diameters are in the range of 1.47 - 2.06 mm. It must be clear that using PSD for this analysis instead of the magnitude of the peak spectrum makes the results independent of the signal duration and number of samples used. However, the frequency at which the bubble sizes are calculated should correspond with how often the bubbles are formed in the porous tube.

Figure 5.12 shows the PSD for signals emitted having an air flow C picked up in position 1. In this case, many more signals frequencies are exceeding the background noise threshold. This indicates that using a higher air flow; other bubble sizes are formed in the porous tube. Figure
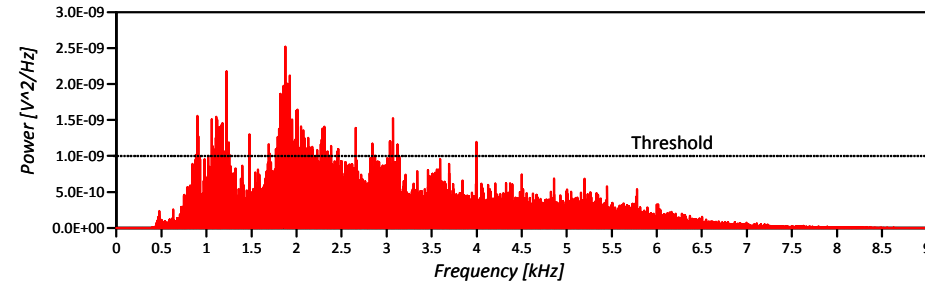


Figure 5.12.: Power Spectral Density using accelerometer in position 1 and air flow C.

5.13 shows the histogram of frequencies for the signal in figure 5.12, it can be observed that using the flow C, the sparger produces bubbles, which emit signals with frequencies mainly around 2 kHz. With this data and using equation 5.1, it is possible to calculate the bubble size distribution present in the flotation column see (figure 5.14). It should be noted that bubble diameters are rounded to two significant digits.
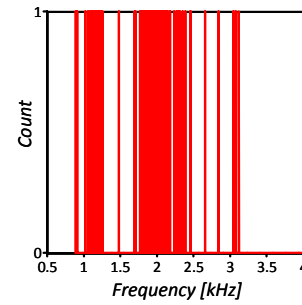
Figure 5.13.: Histogram of frequencies for accelerometer signal in position 1 and flow C.
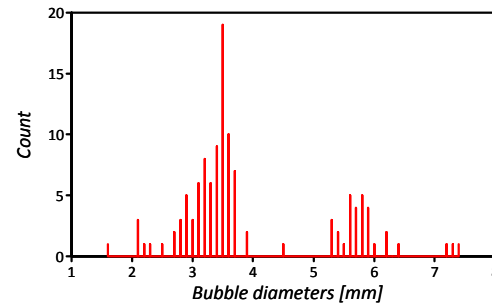


Figure 5.14.: Histogram of bubble diameters for accelerometer signal in position 1 and flow C.

Having the histogram of bubble diameters, it is easy to calculate the total volume of air contained by the bubbles and total superficial area of the bubbles for the analyzed signal, which corresponds to 5,1 $cm^3$ and 63,9 $cm^2$ respectively, in this case.

Figure 5.15 shows the normalized intensity graph of the spectrogram for the signal in figure 5.12. During five seconds, the frequencies emitted by the bubbles are represented by different colors in the spectrogram. Through the period, different frequencies reach a peak in every second, with peaks between 1 kHz and 3 kHz occurring most often. Finally, figure 5.16 shows the PSD for signals emitted using an air flow E and picked up in position 1. Due to the high air flow used (flow E), the number of signals exceeding the threshold is significant. From figure 5.16, it can be stated that the porous tube used for bubble sparging tends to produce much more bubbles with frequencies between determined ranges, for example 1-2 kHz and produces less bubbles with frequencies in the range 4-5 kHz, corresponding to bubble`s diameters in the range of 3.3 - 6.6 mm.
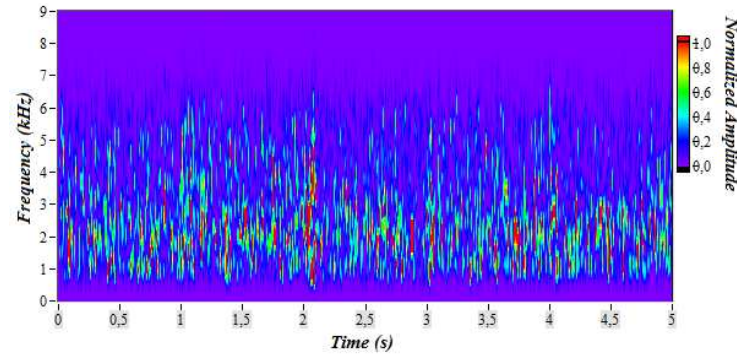
Figure 5.15.: Time-frequency representation for the signal in Figure 5.12
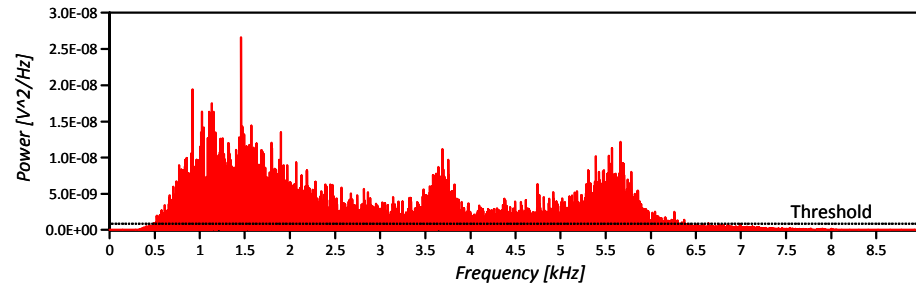


Figure 5.16.: Power Spectral Density using accelerometer in position 1 an air flow E.

Figure 5.17 shows a 3D representation of the Power spectral density, frequency and time for the signal of figure 5.16, during 33 seconds. Like figure 5.15, this representation helps to visualize the behavior of the signal under determined parameters of operation, and therefore understand the bubble size distribution during the analyzed period.

From figure 5.16 and figure 5.17, it can be observed that for air flow E, the sparger that was used, produces bubbles with three defined ranges of size corresponding to the frequencies centered around 1.5 kHz, 3.7 kHz and 5.6 kHz. This frequency-power spectral density distribution is relatively constant through the time, suggesting that the sparger emits bubbles with a constant diameter under a specific air flow.

Unfortunately, the signals emitted using air flow A could not be identified, since they were too faint to distinctly exceed the threshold even using PSD analysis.
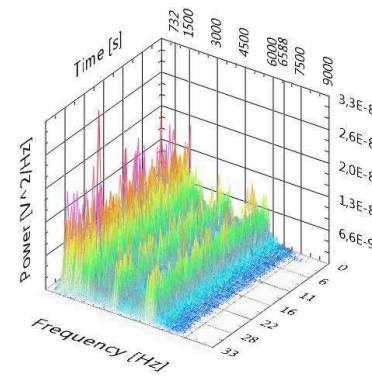
Figure 5.17.: Waterfall plot of the signal in figure 5.16, during 33 seconds.

### 5.2.4.3. Pattern recognition analysis

**Pre-processing and feature selection**

In this section, Pattern recognition is used to classify AE signals from different air flows. The aim is to identify and classify changes in the air flow. For this purpose, three air flows were used, namely flow A, flow B and flow C (see table 5.1). Monitoring the bubble activity independently of an air flow meter permits identifying change in the bubble activity affected by factors such as; changes in the slurry composition, changes in the used reagents and the correct operation of the sparger.

Following the steps in table 4.1; after collection of the data, it is divided into training data set and test data set (step 2), for each of three air flows. For Pattern Recognition analysis, only the signals from the AE sensor in the measurement place road 11 are used.

Three features are selected to characterize each AE signal, namely RMS, Variance and Entropy as figure 5.18 shows. For further information about these features, please see chapter 3. These features were selected after examination of the data by means of plotting and comparison of the features obtained (pre-processing, step 3 and step 4). As a result, these features are selected since they are capable of representing each flow in a distinctive way.

Each pattern is formed by these three features and represents a point in the feature space (figure 4.2). These patterns act as the fingerprint of the signals and are the raw material for the classification process. Then a pattern array is formed with the different patterns (figure 5.19). This array is the input data to be used in a supervised or unsupervised method.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} Root\ Mean\ Square \\ Variance \\ Entropy \end{bmatrix} = \begin{bmatrix} RMS = \sqrt{\dfrac{1}{n}\sum_{i=0}^{n}|x_i|^2} \\ Std.\,Dev = \sqrt{\sum_{i=1}^{n} p_i(x_i - \mu)^2} \\ WPE = \int_{-\infty}^{\infty} f(t)^2 dt = \sum_{j=1}^{n} E_j \end{bmatrix}$$

Figure 5.18.: Pattern, $\mathbf{x}$, of three features used to characterize the flows (feature extraction).
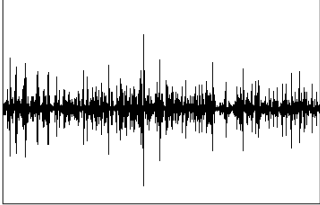


Figure 5.19.: Pattern array from air flow AE signals.

**Normalization and data dimension reduction**

After the pattern arrays are formed, the classification algorithm should normalize the data input. This normalization permits a visualization and comparison of the patterns in the feature space. Normalization shifts and scales the axis in the feature space. This permits working with different features containing different units. *Standard score normalization* (section 4.4.2) is used to normalize the signals. This type of normalization can present negative values and can be seen as a dimensionless standardization.

Figure 5.20 shows the AE patterns formed by the normalized features of figure 5.18. This process permits visualizing and comparing features with different ranges and units, such as *Entropy* and *Variance*. For reducing the data dimension from a three to two-feature space, *Principal Component Analysis* (PCA) and *Linear Discrimination Analysis* (LDA) are employed.

PCA technique permits deriving new features in decreasing order of significance, being this new feature a linear combination of the original feature. As a result, a reduced group of features is obtained, which describes better the data (see section 4.4.4.1).

The main characteristic of this data reduction technique is that it makes no assumptions about
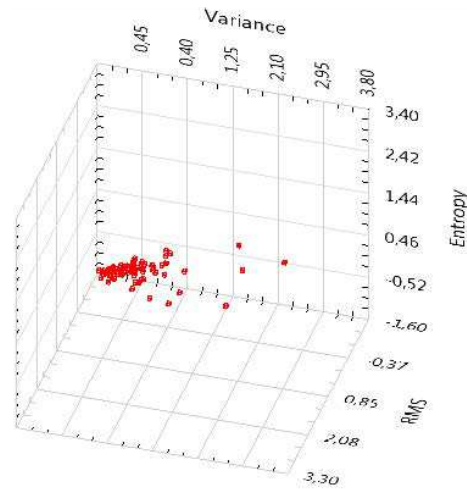
Figure 5.20.: 3-D plot of the normalized AE patterns containing three features each one.

the natural clusters in the data (unsupervised technique).

This process is called feature extraction (section 4.4.4), which takes these three features and transforms them into two features. The new features form a reduced pattern, which have a lower dimension with minimum loss of information. This transformation permits visualizing the features in a suitable form and a simplified structure of data is achieved.

Figure 5.21 shows the feature extraction of the data set illustrated in figure 5.20 in a reduced feature space.

This reduction of the data dimensionality using PCA permits a classifier algorithm to have a better performance. Also it increases the speed of the algorithm and reduces redundancy. *Linear Discrimination Analysis* (LDA) is a data reduction technique based on the Fisher`s criterion. The principle of LDA is the maximization of the separability of classes, together with the minimization of the within-class variability (section 4.4.4.2).

These two data extraction techniques are used to prepare the input data before classification. Also, these techniques are compared and their influence on the general clustering performance is analyzed.

After preprocessing of the data, it is prepared to be classified. If previous knowledge about the data is available, is it possible to label the input data and used it as a training set for a supervised classification. If not, unsupervised classification or clustering can be performed to classify the data.

In this work both methods are evaluated and the classification methods compared each other.
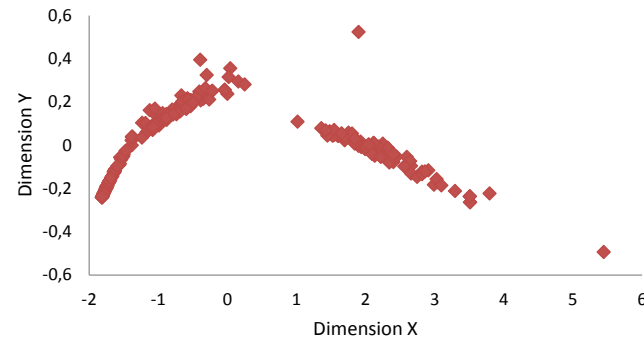
Figure 5.21.: Dimension reduction of the input data using PCA.

**Unsupervised classification**

Sometimes, no information about the parameters of the process is available or the flotation process inside of the column is uncertain. In these situations an unsupervised classification of the different states in the process can be performed. The aim is to detect changes in the kinetics of the process to avoid turbulence inside of the column flotation.

In this work, *K-means clustering* is employed to classify the AE signals emitted by the bubbles as a product of different air flows. *K-means clustering* divides the input pattern in $k$ clusters based on the minimization of the sum of square error of each cluster (Section 4.6.3).

The clustering accuracy is evaluated by the Rand Index (RI). RI is a measurement of the similarity between two clusters based in a criterion (section 4.6.6).

Although, the number of types of flows is known, namely flow A, flow C and flow E (see table 5.1), the aim is to evaluate the capacity of the k-means method to cluster these three types of flows and identify them.

Figure 5.22 shows the 2-D projection of the input data using PCA as data extraction method. The three-feature patterns were first normalized. As can be observed in the feature space, some flow A patterns are in the same area or close to some flow C patterns. However, as the air flow is incremented, its AE patterns move to the right side of the feature space and they are distinct separated from the patterns of other air flows. It important to mention that this is the real distribution of AE patterns from different flows. It can be illustrated with the help of the available pattern labels of each AE pattern.

However, sometimes no information is available about the AE signal sources. Thus the labels of each patterns does not exist. In these situations, a clustering method can be applied to identify and classify the signals (patterns).

Figure 5.23 illustrates the clustering of the input data using PCA as data extraction methods and *K-means* as clustering method. In this figure the symbols assigned to each cluster are
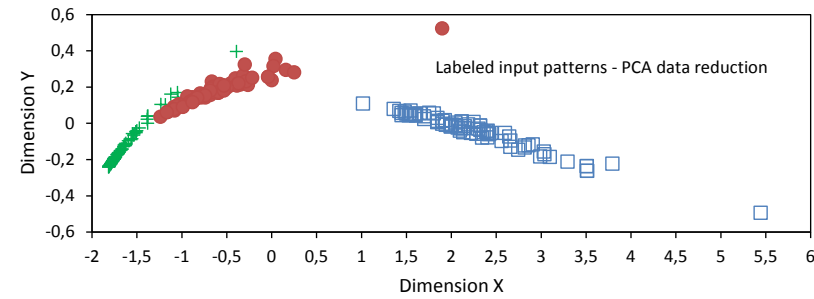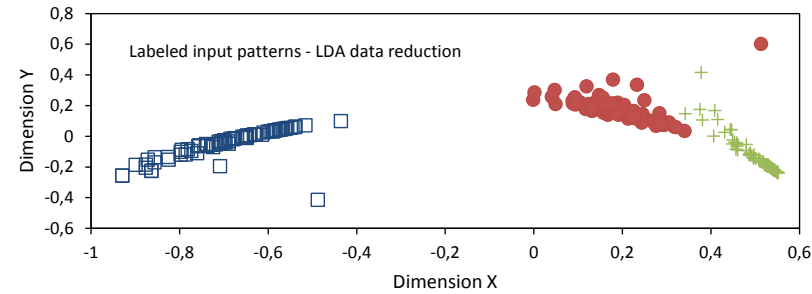
Figure 5.22.: Labeled input patterns using PCA; + air flow A patterns, ● air flow C patterns, □ air flow E patterns.

not relevant, but the pattern classification should be as similar as possible with respect to the labeled classification in figure 5.22.

The clustering accuracy obtained is RI=0,96 which is pretty high considering that a value of RI = 1 represents a classification without errors. The loss of accuracy is caused by the outliers in flow A and flow C, which are misclassified. The patterns from flow A and flow C are overlapped and therefore some of them are confound. However, the patterns of flow E are quite separated from the other and can be classified in an acceptable manner. Now, LDA is used as a data
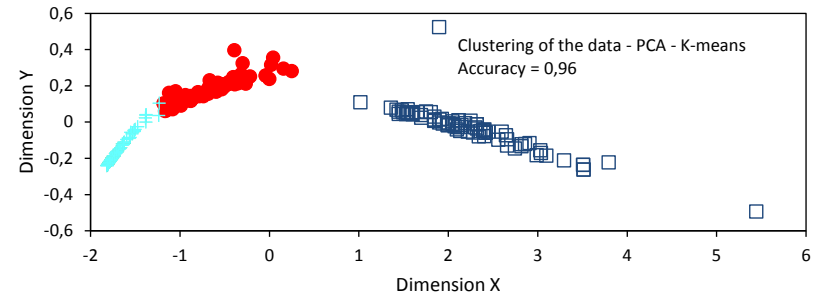


Figure 5.23.: Clustering of the feature patterns (PCA - K-means); + 1° patterns cluster ● 2° pattern cluster, □ 3° pattern cluster.

extraction method in combination with the same clustering method (*K-means*) to evaluate and compare the global clustering performance by means of the *Rand Index*.

Figure 5.24 illustrates the labeled input patterns; this time using LDA technique to reduce the dimensionality of the data. Using LDA, the patterns corresponding to the flow E are more isolated from the patterns of others flows (in comparison with the results using PCA).
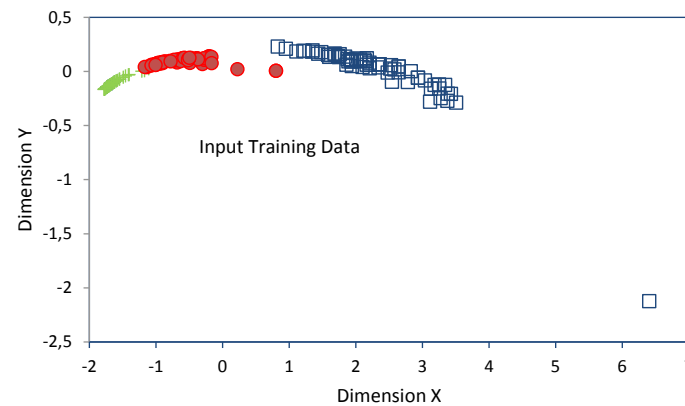
Figure 5.24.: Labeled input patterns using LDA; + air flow A patterns, ● air flow C patterns, □ air flow E patterns.

Figure 5.25 illustrates the clustering of the same input patterns, this time using LDA as a data reduction method. The clustering accuracy for this combination is RI = 0,97 being slightly superior with respect to the previous one. It is important to note that the flow C patterns are all correctly classified and isolated in a region of the feature space.
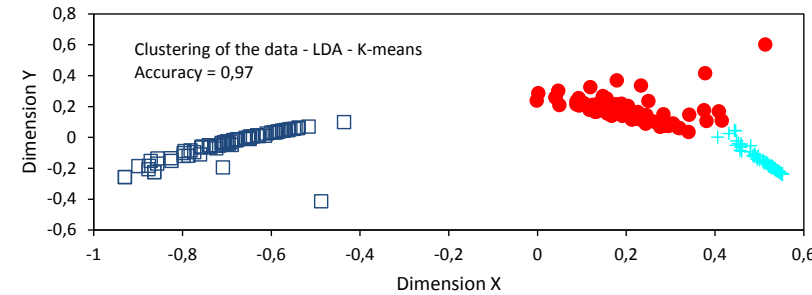


Figure 5.25.: Clustering of the feature patterns (LDA - K-means); + 1° patterns cluster ● 2° pattern cluster, □ 3° pattern cluster.

As a result, it can be concluded that there is no relevant difference in the clustering performance using PCA or LDA as a data extraction method. Also, both methods offer good results with *Rand index* close to the unit. This clustering data can be used as the input data to employ a supervised method.

**Supervised classification**

In supervised classification, the labels of each AE event of interest must be known and available. In this case the AE events are the AE signals coming from different air flows (From the air bubbles inside the column flotation to the AE sensor in the measurement place "road 11").

This information is used to train a classifier. The Classifier algorithm assigns labels to the new input data or test data (see section 4.5).

In this section a nonlinear discrimination function is used. These functions are a sum of basis functions. Multilayer perceptron (MLP) is a technique used in PRT, which produces a transformation of the input data (patterns) into a different dimensionality space by means of nonlinear functions and then it forms a linear combination of the data in the output nodes (please see section 4.5.3.3).

*Back-Propagation Neural Network* is a nonlinear optimization scheme for MLP, which evaluates an error function and its derivatives with respect to the MLP weights.

Coming back to the problem of recognizing the different air flows present in the column flotation, the same three flows used in unsupervised classification are used, namely flow A, flow C and flow E (see table 5.1).

Also, three-feature AE patterns are used as in the previous section, but this time with the aim to evaluate the classification accuracy of the supervised method; Back-propagation neural network.

For data extraction, PCA (section 4.4.4.1) is used to reduce the data dimensionality from a three to a two-feature space.

The effectiveness of this supervised method is evaluated by the classification accuracy (section 4.5.4).

Figure 5.26 illustrates the labeled pattern distribution in the feature space from different air flows. This set of patterns is used to train the supervised classification (*Back-propagation neural network* algorithm).



Figure 5.26.: Labeled input patterns using PCA; + air flow A patterns, ● air flow C patterns, □ air flow E patterns.

As can be seen, the training data set presents an outlier within the flow E patterns. However, in this configuration, this outlier does not affect the performance of the supervised classification because it is placed alone and close to patterns of its group.

Figure 5.27 illustrates the supervised classification of the AE patterns using *Back-propagation neural network*. The AE patterns are identified with class labels predicted or calculated by the BP neural network algorithm.

The number of hidden layers used in the algorithm was five and the classification accuracy obtained was 97,5 porcent. These accuracy is very high and it is a proof that the features that form the patterns were correctly selected. This results are encouraging because they show it is possible to identify and classify small changes in air flow in order of a few liters per minutes.

These results support the use of this type of analysis in flotation columns to monitor the AE activity and the air flow changes. This is especially desired in processes where a simple air flow meter cannot always indicate the real bubble activity of the air bubble due to the numerous factors that affect the process.
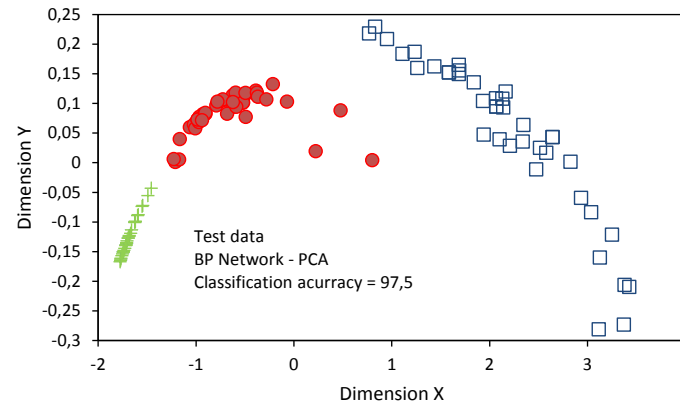


Figure 5.27.: Supervised classification of the feature patterns (PCA - BP neural network); + air flow A patterns, ● air flow C patterns, □ air flow E patterns.

Although there are techniques that measure the bubble size and bubble activity by means of image processing, these techniques are difficult to employ, further the images require signal processing. On the other hand, the AE sensors are simple to mount (on the external shell of the column flotation) and the measurement system can be set in automatic mode with the help of PRT.

Other features can be employed to form new patterns in accordance with different process conditions. Also, other supervised algorithms can be evaluated to be used in column flotation.

### 5.2.5. Conclusions

Two different sensors are used to characterize the bubbles produced by a porous sparger in a column flotation. The use of the stainless steel rods as waveguides is beneficial for accelerometer and AE signal collection. Signals emitted by rising bubbles are picked up for analysis and they show the presence of turbulence and collisions between bubbles.

Despite the fact that AE signal frequency content did not yield information relevant to bubble size, the shape of the bubble signals in the time domain obtained with AE signals proved to be useful to obtain more parameters in order to characterize the process. Accordingly, exploration of the use of other types of AE sensor and methods to get more information about the bubbles in the time domain analysis is recommended. However, AE signals were employed successfully in the identification and classification of different air flows (bubble activity). This capacity is especially desired to identify or prevent turbulence inside the column flotation.

The signals obtained from the accelerometer contain information in the frequency domain and can be correlated with the Minnaert`s equation to determinate the size of the bubbles immediately after they are formed. Valuable information could not be observed from the spectrum of AE signals, which are in a higher frequency range. These results reaffirm the existing theory about frequency contents in signals from bubbles as well as makes use of this monitoring method sensitive to the background noise in an industrial plant. However, this may not play a large role because sources must be close to the contact sensor, and frequency filters could be used to separate the bubble signals.

For each particular case, a suitable threshold in the frequency domain must be set to isolate the bubble information from the background noise. Showing the spectral analysis in terms of Power Spectral Density resulted to be an efficient tool to differentiate between the floor noise and the frequencies emitted by bubbles. Moreover, the frequencies that exceed the threshold are more important for the analysis than the magnitude of these frequencies, making use of PSD suitable for the calculations.

Further analysis is recommended to completely understand the sources of the acoustic emission. To obtain bubble size distribution information as accurately as possible, the frequency at which bubbles are formed in the sparger needs to be correlated with the air flow introduced and the spectral analysis. Another independent method such as visual measurement is suggested to confirm the physical size of the bubbles assessed by the contact sensors.

The use of contact sensors to monitor a column flotation process is a promising tool, particularly the use of accelerometers. This information can be used for online-monitoring, automation of the process and improvement of process production efficiency. The use of AET can be complemented with pattern recognition technique to identify and classify different air flows to prevent turbulence.

*K-means clustering* is used in combination with two data extraction methods, namely PCA and LDA to evaluate the feasibility of clustering AE patterns coming from different air flows.

Both combinations provided high degree of clustering accuracy, confirming that unsupervised technique can be used to identify and classify different air flows.

For supervised classification; *Back-propagation neural network* algorithm is performed in combination with PCA technique for data extraction. Using this combination high classification accuracy is obtained and small changes in air flow were capable of being identified and classified.

## 5.3. AET in Rock Cutting

### 5.3.1. Introduction

AE waves are generated due to dislocation motions (discontinuity of displacements as cracking) in a solid. These waves spread inside the solid and can be detected by an AE sensor on its surface. Since, there are different sources of AE signals, the characteristics of AE signals can widely vary depending on every configuration. Two important characteristics of AE technique are the facts that first, only the formation and expansion of cracks into the rocks are able to be detected with this technique and secondly, every measurement is not perfectly reproducible under the same conditions owing to the random path formation of crack and the stochastic nature of the AE.

In rocks, the rapid release of elastic energy by the process of crack growth or deformation within the rock generates transient pulses of elastic wave energy in form of AE events. AE measurements in rocks cover the frequency range from 1 kHz to 1 MHz [GO10]. Nevertheless, the frequencies of interest in our case cover the range between 60 kHz to 200 kHz. Although rocks are not homogeneous, their properties depend on their dimensions.

AE signals are composed of a deterministic part and a stochastic part. However, the deterministic part is very small [Lyo04], that is, the deterministic part of an AE signal is present in the low frequencies only, in the same way, the influence of the deterministic part is negligible in high frequency. Consequently an AE signal can be considered to be stochastic signals.

Signals that cannot be described with a time-domain equation are called *noise signals*. However, this kind of signals can be quantified in a valuable way using the statistical measures of random sequences [Lyo04]. These statistical indicators can be used as signal analysis tools, such as the *Standard Deviation*, of discrete random samples. Also other parameters or features of the discrete signals can be used to identify and classify the signals.

The study of AE signals emitted by rocks as means of material identification in rock cutting machinery is relatively new. Thus, this kind of research has not been broadly addressed. Shen, et al., [SHK97], founded that AE technique is suitable for the recognition of high dust generation occurring during rotary coal cutting operations. However, material recognition is not studied in depth. More recently, Crosland, et al.,[CMH09] founded a set of major dominant frequencies using fast Fourier transforms. They concluded that the frequency content in AE

signals can be used to characterize the acoustic emissions emitted by rocks during the cutting. Nevertheless, the measurements are made with a resonant AE sensor at a sampling rate of 200 kHz. Consequently, only the frequency spectrum up to 100 kHz could be analyzed.

Nienhaus, et al.,[NKBB12], found that AE frequency range in rock cutting takes place from 60 kHz to 200 kHz approximately. This study considers this investigation and focuses on the signal processing of the signals obtained previously in combination with Pattern Recognition Technique.

Different parameters are used to form fingerprints of the each signal. These fingerprints are called patterns and are the raw data to identify and classify different AE signals coming from the cutting of rocks.

Different techniques for Data Extraction, such as *Principal Component Analysis* (PCA), *Multidimensional Scaling* (MDS), and *Linear Discriminant Analysis* (LDA) are employed and evaluated.

Also, different Unsupervised Classification techniques such as *K-means clustering*, *Fuzzy C-means clustering*, *Vector quantization* (VQ) clustering are employed to evaluate the feasibility of clustering of the AE patterns.

Finally, different Supervised Classification methods such *Back-propagation neural network*, *Linear Support Vector Machine* and *Non-linear Support Vector Machine* are employed and their accuracies are evaluated.

## 5.3.2. Experimental Methods

The test bench and the data acquisition system used to pick up the AE signals are described in this section. A linear rock cutting machine was used at constant speed for the tests. The cutting speed was measured by an inductive sensor at a sampling rate of 1 kHz [NKBB12]. The AE sensor was placed on one side of the cutting tool. The cutting configuration was the following:

- Cutting width of tool: 10 mm
- Cutting speed: 14 mm/s
- Cutting depth: 10 mm
- Cutting angle: 45°

For the AE measurements a broadband sensor model VS160-NS was used, frequency range [100-450] kHz and the peak frequency response is 150 kHz. For data acquisition, a PXI National Instrument system was used with sampling rate of 1MHz [NKBB12]. LabView 2011 and Matlab 8.2 code was used for acquisition and processing of the AE signals (Figure 5.28).
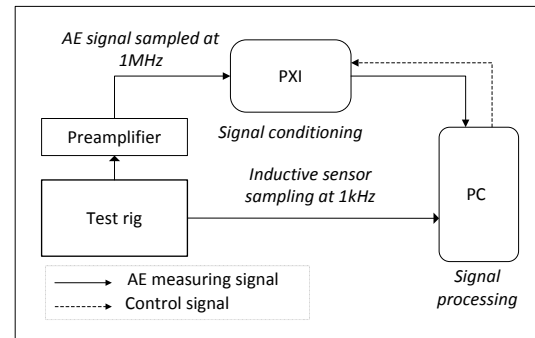
Figure 5.28.: Diagram of the data acquisition system [NKBB12].

**Types of Rocks Tested**

Three types of rocks and gypsum were used in this study; coal, dead rock (clayey sediments) and oil shale [NKBB12]. To allow mechanical stability, a sample box was used as shown in figure 5.29. Later, to maintain the rocks fixed in the box, gypsum was used as supporting material. The coal was mined in Auguste Viktoria (AV) mine, located in North Rhine-Westphalia, Germany. In the same way, the dead rock was obtained from AV mine and is made of clayey sediments with pieces of coal and mudstone. In addition, oil shale was taken from a bore sample.



Figure 5.29.: Sample box containing two rows; coal (upper row) and dead rock (lower row) [NKBB12].

### 5.3.3. Results and discussion

#### 5.3.3.1. AE cutting signals

Time domain representation is the first step of analysis. In these representations of the signals it is easy to appreciate that every rock emits very different AE signals during the cutting

process as illustrated in figure 5.30, figure 5.31, figure 5.32 and figure 5.33, where the signals are normalized. The data necessary to generate these graphs can be considered to be the raw material and the initial point of the signal processing. At this point, the obvious question is; how long should be an AE signal to represent itself in an adequate manner? The answer could be the duration enough to contain all the details of the AE signal, namely bursts and the continuous part. Apart from AE signal of coal, the others signals show fluctuations in an erratic way and it seems that at first they need more than one second to be represented. However, just a fraction of one second AE signal is enough to identify every AE signal.

The sample rate used to record the AE signal is 1MHz. In other words, there are one million points per every second of AE signals. These points are treated as if they were a random sequence to extract statistical parameters as well as others features.
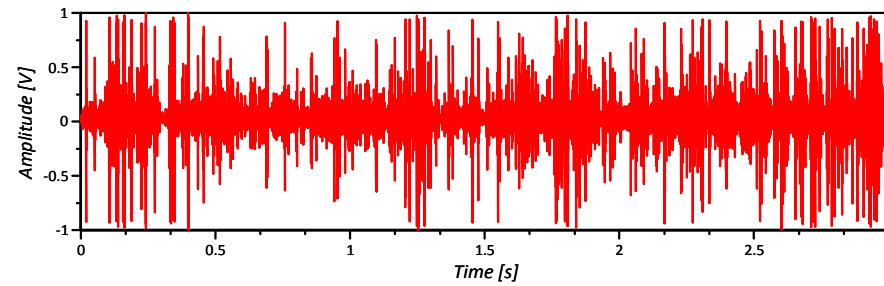


Figure 5.30.: Normalized AE signal of coal cutting.
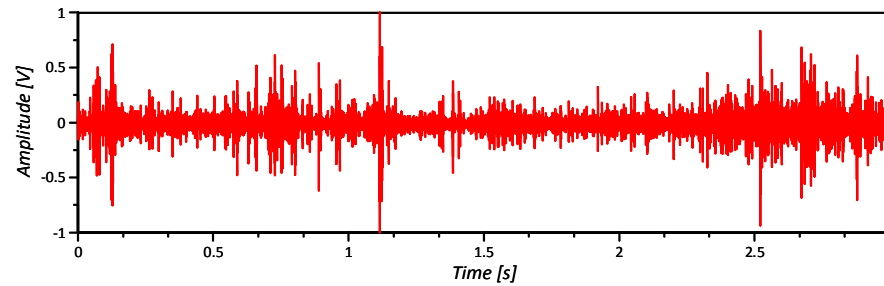


Figure 5.31.: Normalized AE signal of dead rock cutting.
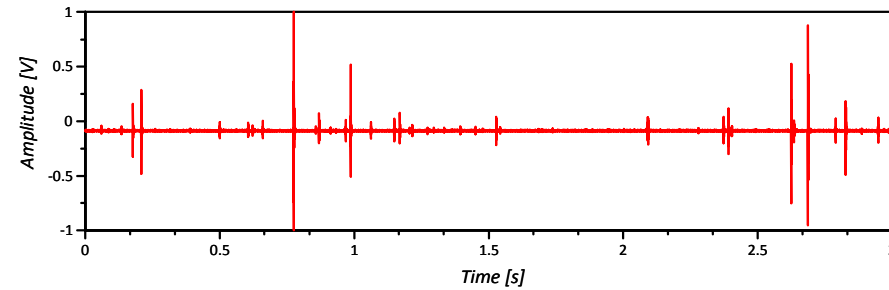
Figure 5.32.: Normalized AE signal of oil shale cutting.
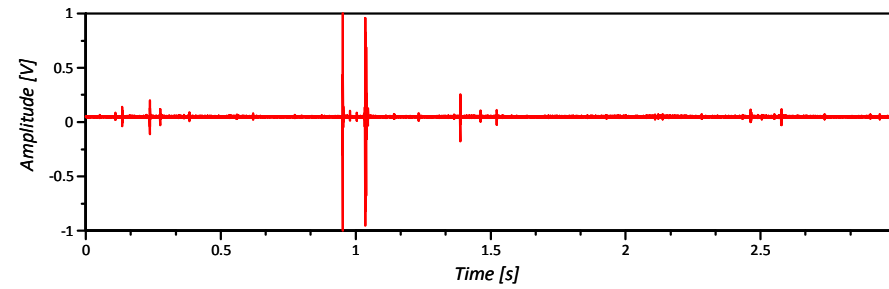


Figure 5.33.: Normalized AE signal of gypsum cutting.

The figures above are the AE signals used to obtain different features to form AE patterns of each rock cutting. Clearly, in the time domain, these signals have different waveforms. The challenges are to identify and classify these AE signals in an automatic manner.

It must be kept in mind that the coal is the most important and valuable material among the considered materials. Thus, the principal aim is to identify the coal and separate the coal signal from the other ones. To achieve this task, signal processing is employed together with Pattern Recognition Technique.

In the following sections, unsupervised and supervised classification methods are employed to evaluate the feasibility of identifying and classifying AE signals coming from different rock cuttings.

### 5.3.3.2. Pre-processing and feature selection

The most important and fundamental part for any pattern recognition method is the collection of appropriate data. For signals emitted in rock cutting, these must have a good quality to

represent and describe each rock. Taking the appropriate number of features for each signal, the rocks may be identifiable using signal processing and pattern recognition.

After recording the AE data from different tests, AE data is divided into training set and test set of data. Also different rocks of the same type were cut to have a better representation of the rock. The examination of the training set permits visualizing the best parameters to be used and not to consider indistinct parameters.

The number of **features selected** to describe each cutting are eight, which characterize the signals not only in time domain, but also in frequency domain (figure 5.34). Features; *NP*, Kurt, $\sigma^2$, *RMS*, and $\sigma$ are related to the time domain. While the features *Freq. Peak* and *Pow. Peak* are related to the frequency domain. Furthermore, the feature *WPE* is related to the time-frequency domain. These features are selected based on their capacity of representation, because each of them has a range of characteristic values for each rock. These features are used for building the patterns $\boldsymbol{x}_i$ representing each rock cutting in the feature space. From figure 4.2 each pattern represents a point in the feature space and also this point represents an AE event (a specific rock cut).

$$
\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} = \begin{bmatrix} \text{Number of peaks} \\ \text{Kurtosis value} \\ \text{Variance} \\ \text{Root Mean Square} \\ \text{Standard Deviation} \\ \text{Frequency Peak} \\ \text{Power Peak} \\ \text{Wavelet Packet energy} \end{bmatrix} = \begin{bmatrix} NP = number\ of\ peak\ found\ in\ the\ data \\ Kurt = \dfrac{\frac{1}{n}\sum_{i=0}^{n-1}(x_t(i)-\mu)^4}{\sigma^4} \\ \sigma^2 = \sum_{i=1}^{n}\dfrac{(x_i-\mu)^2}{n} \\ RMS = \sqrt{\dfrac{1}{n}\sum_{i=0}^{n}|x_i|^2} \\ \sigma = \sqrt{\sum_{i=1}^{n}\dfrac{(x_i-\mu)^2}{n}} \\ Freq.\,Peak = \dfrac{\sum Power\ spectrum_{(j)} * (j*df)}{\sum(Power\ sprectrum_{(j)})} \\ Pow.\,Peak = \dfrac{\sum(Power\ spectrum_{(j)})}{ENBW} \\ WPE = \int_{-\infty}^{\infty} f(t)^2 dt = \sum_{j=1}^{n} E_j \end{bmatrix}
$$

Figure 5.34.: AE signal feature extraction. The vector $\boldsymbol{x}$ represents a pattern containing eight features.

Once each group of features for each signal is defined, the next step is to collect the vectors and form the Pattern Arrays, which are our input to the different supervised and unsupervised methods (Figure 5.35). The Pattern array is the input data set for the pattern recognition algorithms and each of their rows represent a pattern or feature vector of the signal.

Sometimes, it is convenient in the pre-processing stage (table 4.1) to rearrange the features as a combination of them; e.g. *NP/Kurt*. This combination of features permits a reduction in the dimension of each pattern and consequently the feature space.
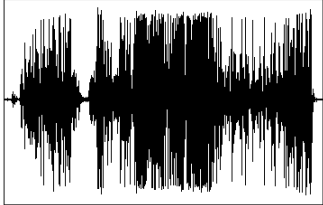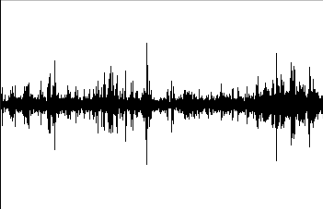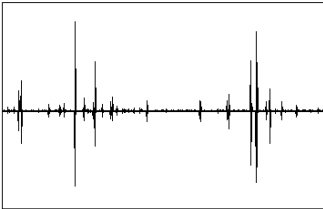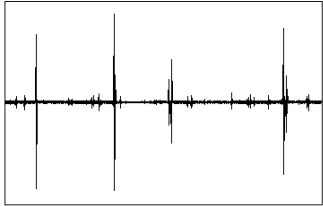
| Signals | Rock Cutting Waveform  Signals | Pattern Arrays |
|---|---|---|
| Coal Signal |  | *Coal* *Pattern Array* $= \begin{bmatrix} x_{1,1} & \cdots & x_{1,8} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m8} \end{bmatrix}$ |
| Dead Rock Signal |  | *Dead rock* *Pattern Array* $= \begin{bmatrix} x_{1,1} & \cdots & x_{1,8} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m8} \end{bmatrix}$ |
| Oil Shale Signal |  | *Oil Shale* *Pattern Array* $= \begin{bmatrix} x_{1,1} & \cdots & x_{1,8} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m8} \end{bmatrix}$ |
| Gypsum |  | *Gypsum* *Pattern Array* $= \begin{bmatrix} x_{1,1} & \cdots & x_{1,8} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m8} \end{bmatrix}$ |

Figure 5.35.: Pattern arrays from rock signals.

**Normalization and data dimension reduction**

After calculation of the pattern arrays, all these features should be normalized to be compared

in a suitable way in the feature space. The normalization or scaling permit stretching the axis in the feature space from zero to one. Thus the number of count or peak amplitude can be seen in an easier way. However, in statistics and its application such as Pattern Recognition, normalization can have a different range of values. For different data set, normalized values allow their comparison by means of shifted and scaled versions. When the population parameters are known, *standard score normalization* can be used. It can present negative values, if an observation is below the mean and positive, if it is above the mean. Standard score can be seen as dimensionless standardization, which is the product of the subtraction of the data mean ($\mu$) from an individual observation ($Feature_{mn}$) and then dividing the difference by the standard deviation of the population $\sigma$ [LM12]. The mathematical formulation is,

$$SC = \frac{Feature_{mn} - \mu}{\sigma} \tag{5.2}$$

Before clustering the data, it is convenient to reduce the dimensions of it. Projecting the data (feature vectors or pattern) into a reduced feature space, helps to analyze data. Clustering can be easier and accurate in the reduced space than in the original space. The data in the reduced space must preserve the most important characteristics. The objective is to process the data in a more manageable low-dimension feature space before classification. There are many dimension reduction algorithms. All of them have the objective of making the pattern recognition analysis more effective.

To illustrate the concept of data reduction, an example in rock cutting patterns is given, using the *isometric Feature Mapping* (Isomap) algorithm. It is shown in figure 5.36. Isomap algorithm uses geodesic distance induced by a neighborhood graph (the high space graph) embedded in the classical scaling. This geodesic distance is defined as the sum of edge weights along the shortest path between two feature vectors. The new n-space is represented by the n-eigenvectors of the geodesic distance matrix.

This data dimension reduction technique is employed just for illustrating the process, but later on this technique is not employed due to the computational resources required to work with it.
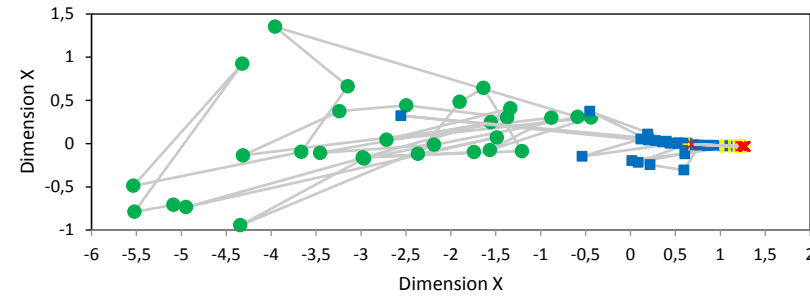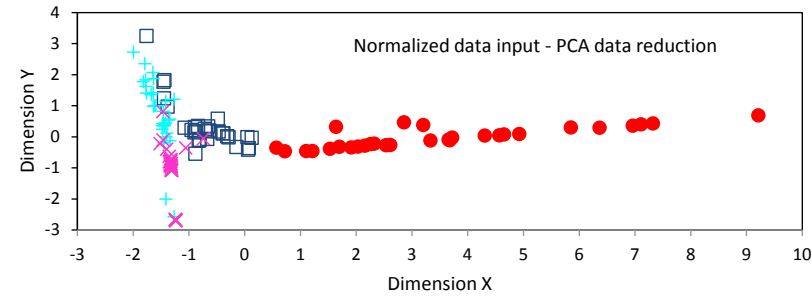


Figure 5.36.: 2-D projection of feature patterns; ● coal vector feature patterns, ■ dead rock features patterns, + Oil shale features, **x** Gypsum features patterns.

Also other traditional algorithms could be used to reduce the dimension of the data such as *Principal Component analysis* (PCA) or *Multidimensional Scaling* (MDS).

### 5.3.3.3. Unsupervised classification

**Introduction**

In this work, a number of clusters to be in the classification are known, that is to say, one for each rock. Also, data representing each rock is available, thus a supervised classification is possible to be carried out. Sometimes, data classification with the help of previous data cannot be counted. In this case, an unsupervised analysis should be carried out. In this work unsupervised classification is evaluated to corroborate its feasibility to be used in rock cutting identification.

Unsupervised pattern recognition uses similarities between patterns to classify them into different groups or clusters. Unsupervised techniques do not require previous knowledge or labeled data. Without previous information, the number of classes or labels must be estimated and evaluated. After clustering the cluster data can be used as the training set for supervised classification.

Clustering is an important approach in unsupervised learning and can be used to identify different groups of patterns in AE data. For example, in an AE testing it can be used to group different signals from different sources.

There is no single best technique for clustering, thus different clustering algorithms produce different results and some methods are not adequate for detecting simple clusters. The reason is that each cluster algorithm forces the data to a specific data structure. Also, the number of clusters must be decided, which can be a problem if the process being measured is not completely understood. It must be considered that clustering algorithms tend to produce clusters even if the input data is random.

To overcome partly the clustering problems, different clustering methods are tested in this work to compare their results and decide which method is most suitable for AE rock cutting patterns. First, two cluster methods are going to be tested and evaluated namely.

- *K-means clustering*

- *Fuzzy C-means clustering*

These approaches need initial partition of the data. Groups means are calculated by means of the nearest class mean rule.

Also, other clustering method which works completely different is tested with rock cut AE features and evaluated

- *Vector quantization (VQ) clustering*

The principal aim of *VQ* is to compress the data. This is made by encoding the feature vector $\boldsymbol{x}$ as one from a codebook of $g$ vectors and then the code vectors are termed (section 4.6.5).

In rock cutting process, the AE signals picked up by an AE sensor are product of non-stationary irreversible phenomena and depend on the section of the mine being cut; differences in the AE signals for the same type of rock could be observed. All that made clustering methods recommendable to be employed.

With each clustering analysis, the Rand index (section 4.6.6) is used to compare the quality and validate the results of each clustering algorithm. In data clustering, this index is a measurement of the similarity between two data clustering.

In this work the input data or features vectors are normalized with *standard score normalization* (equation 5.2) to compare them in a suitable form.

Also three types of data reduction algorithms are tested and their results are compared. All of them are set to reduce the dimensionality of the data from eight to three dimensions in the feature space.

- *Principal Component Analysis* (PCA)

- *Linear Discriminant Analysis* (LDA)

- *Multi-dimensional Scaling* (MDS)

This analysis is made to find the best clustering algorithm for clustering AE Patterns from rock cutting, as well as the best data reduction algorithm for this purpose.

**K- means clustering**

This clustering method corresponds to the sum-of-squares methods (section 4.6.3). These kinds of methods find a partition of the input data that maximizes a predefined criterion. Particularly, *K-means* divides the input data ($n$ AE pattern) into $k$ clusters, so that the sum of squares of each group is minimized (Each pattern is cluster with the nearest mean). This algorithm works alternating two procedures;

1. Allocation of patterns to a group based on the minimization of the total within-group sum of squares about the centroids (Section 4.6.3).

2. Calculation of a new group mean based on the previous allocation.

As seen in figure 5.34, the input data correspond to feature vectors containing eight features. For each AE signal rock there are different patterns as shown in figure 5.37.

In figure 5.37 the normalized input data is showed to illustrate what the real pattern distribution of each rock patterns look like.

Normally, when clustering is applied, the information presented in figure 5.37 is not available, that is, the label of each pattern. From this figure, it can be observed that the coal patterns are separated from the others rocks patterns and occupy a large horizontal band in the future space.
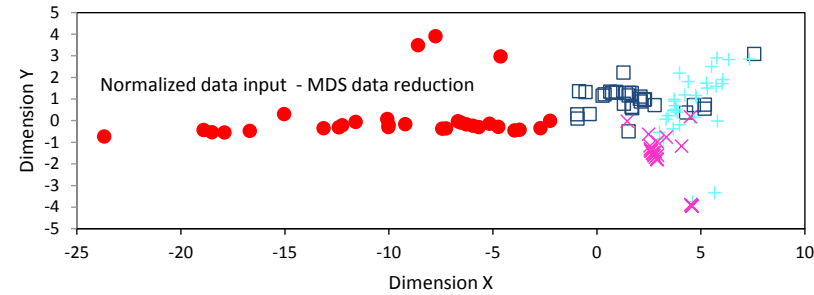
Figure 5.37.: 2-D projection normalized input feature patterns(PCA); ● coal vector feature patterns, □ dead rock feature patterns, + Oil shale feature patterns, × Gypsum feature patterns.
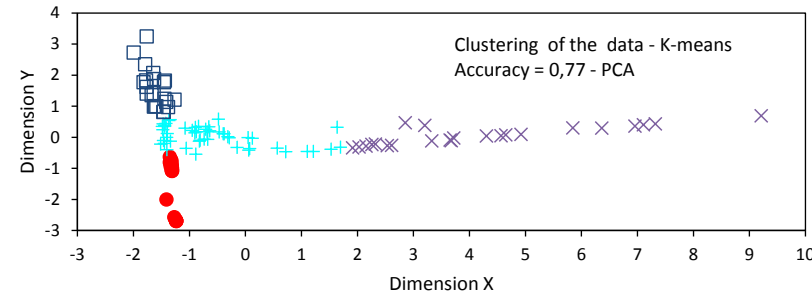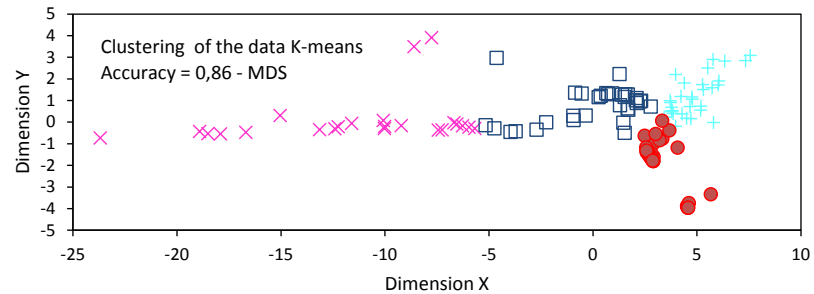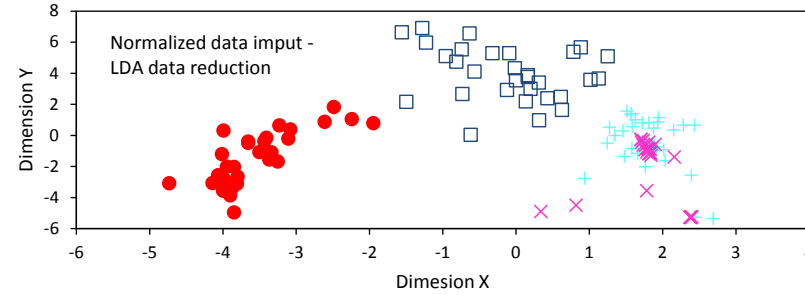


Figure 5.38.: 2-D projection of clustering feature patterns (PCA - K-means); ● 1°pattern cluster, □ 2°pattern cluster, + 3°patterns cluster, × 4°patterns cluster.

Also, the other patterns are not completely segregated each other, existing some overlapped clusters in the feature space. This situation is punctual for this input data and it would change with other input data set.

Visualization of the data is useful to provide an insight into the nature of the multi-feature data. These figures can be useful to detect important (or unimportant) features, natural clusters in the data and detect outlier patterns.

In figure 5.38 the clustering of the data input set is presented. It is important to note that the symbol assigned to each cluster group is not important, but the distribution of the patterns of each cluster on the feature space should be as similar as possible to the input data (Figure 5.37). The accuracy (Rand Index) of the K-means algorithm, using PCA data reduction algorithm is of RI= 0,77 (the rand index ranges from 0 to 1, being assigned 0 to the worst performance of the clustering algorithm and 1 assigned to the best, (see section 4.6.6)). This result is not satisfactory to classify the rock because all the rocks pattern clusters are overlapped in the feature space.

Now, Multi-dimensional scaling (MDS) data reduction algorithm is used instead of Principal Components Analysis (PCA) using the same data input set. MDS is a class of technique that analyses a matrix of distances or dissimilarities (section 4.4.4.3) with the aim to produce a representation of data points in a reduced dimension space.



Figure 5.39.: 2-D projection normalized input feature patterns (MDS); ● coal vector feature patterns, □ dead rock feature patterns, + Oil shale feature patterns, × Gypsum feature patterns.

As shown in figure 5.39, using an MDS data reduction algorithm, the data is rearranged in a different manner than with PCA data reduction algorithm. Nevertheless, also in this case, the coal features are isolated from the other pattern, while the remainder of pattern groups are overlapped each other. Other interesting characteristic of the MDS is that it produces outlier patterns in the coal cluster, situation that is not observed with PCA data reduction algorithm.
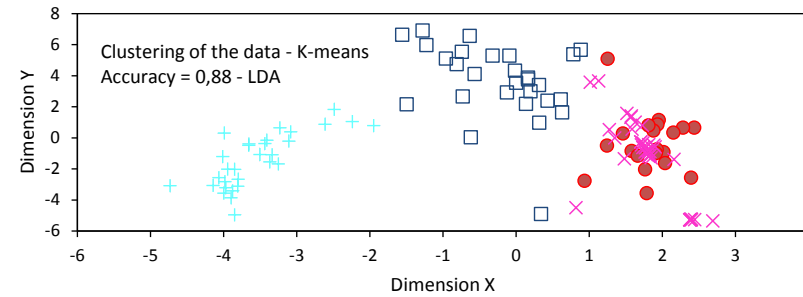


Figure 5.40.: 2-D projection of clustering feature patterns (MDS - K-means); ● 1°pattern cluster, □ 2°pattern cluster, + 3°patterns cluster, × 4°patterns cluster.

Figure 5.40 shows the clustering of the data in the feature space. As explained before, the symbols of the patterns are not important but their locations define the accuracy of the *K-means* method using MDS. MDS together with $K - means$ methods give better result than PCA algorithm. The accuracy obtained using MDS was of RI= 0,86 which can be considered

a good result. The similarity between the labeled input data and the clustering of the data is satisfactory. However, some coal patterns are clustered as members of other cluster as can be observed comparing the figure 5.39 and figure 5.40.

Finally, *Linear Discriminant analysis* (LDA) is employed to reduce the data dimension. LDA finds a linear combination of the AE features with the aim of separating two or more classes of patterns. This linear combination of AE features is used to reduce the dimension of the data input (section 4.4.4.2).

LDA shares some characteristics with PCA as they both search linear combinations of variables which best represent the data. However, LDA searches for those patterns in the input data that best discriminate among classes instead of those that best describe the data. On the contrary, PCA does not take into consideration any difference in classes.

Figure 5.41 shows the same input data as in figure 5.37 and figure 5.39. In comparison with the two previous data reduction methods, LDA gives clusters of the coal and dead rock isolated in the feature space. These results are particularly good, since the principal aim is to isolate the coal patterns from other patterns. In this case, only the patterns belonging to dead rock and gypsum clusters are overlapped each other.



Figure 5.41.: 2-D projection normalized input feature patterns (LDA); • coal vector feature patterns, □ dead rock feature patterns, + Oil shale feature patterns, × Gypsum feature patterns.

Figure 5.42, shows the clustering data using LDA as data reduction methods. The accuracy obtained is of RI=0,88 which can be considered acceptable. Also, this data reduction method permits differentiation between the coal patterns and other patterns distinctly and the cluster of coal patterns is not overlapped with other clusters.

Up to now, the number of clusters used has been four because the clustering tests were performed to evaluate the capacity of a specific clustering method to differentiate four classes of patterns from four AE signals. The purpose is to identify the optimal clustering method as well as the optimal data reduction method and not the optimal number of clusters because this variable is known in advance.
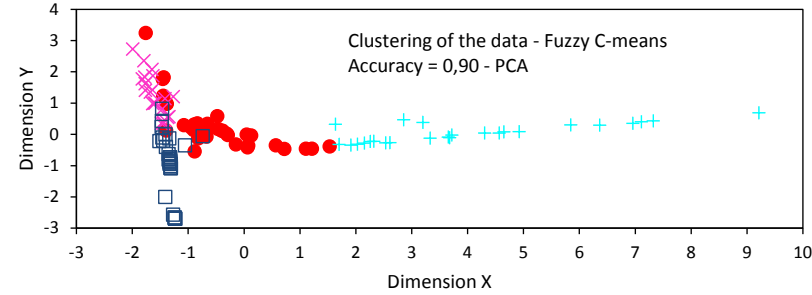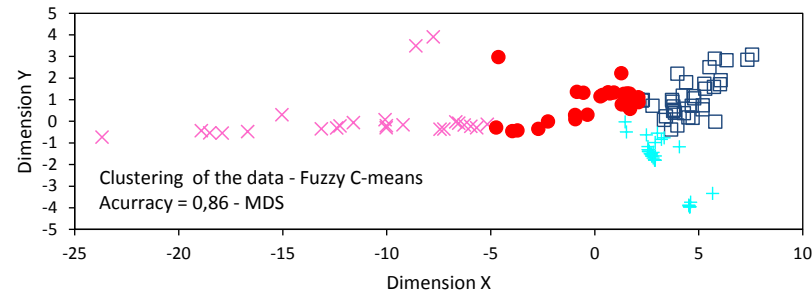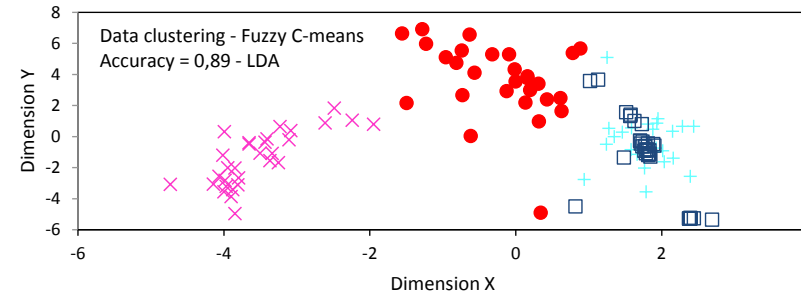
Figure 5.42.: 2-D projection of clustering feature patterns (LDA - K-means); ● 1°pattern cluster, □ 2°pattern cluster, + 3°patterns cluster, × 4°patterns cluster.

**Fuzzy C-means clustering**

The concept of the Fuzzy C-means (FCM) algorithm is that the AE patterns are permitted to belong to all clusters with different degree of involvement. This is the main difference of this clustering method with K-means, which assign one AE pattern to one group only. However, Fuzzy C-means provides degree of cluster involvement. This manner of assignment is called *Fuzzy*, where each AE pattern has a degree of belonging to different clusters.

FCM is one of the commonly used fuzzy clustering algorithms and was first developed by Bezdek [Bez81]. The basic algorithm is iterative and can be described as follows. For further details, please refer section 4.6.4.

1. Set the number of clusters, exponential weight, initial partition matrix, and the finalization criterion.

2. Calculate the fuzzy cluster centers using the initial partition matrix.

3. Calculate the new partition matrix using the fuzzy cluster centers

4. Calculate the new partition matrix (as the difference of the actual minus the previous one) and go to the step 2. This is iterative till the finalization criterion is fulfilled.

It is important to note that the algorithm applied on this calculation was run without information about the degree of membership matrix.

The analysis is applied as in the previous section, using the same input data to compare the results. Since the same data input is analyzed, but this time using FCM algorithm, the figures of data input using different data reduction algorithms are the same (Figure 5.37, Figure 5.39, Figure 5.41).

Again, the three data reduction algorithms are tested, this time with FCM. The *Principal component analysis* (PCA) algorithm is tested first. The normalized data input corresponding to the figure 5.37 is the same in this case.

The figure 5.43 shows the clustering of the AE patterns using FCM. The accuracy obtained is RI = 0,90 corresponding to a clustering of very good performance. However, if the clustering of the data is compared with the labeled input data in figure 5.37, it can be observed that some coal patterns are confounded with patterns of other cluster by the FMC. Nevertheless, FMC together with PCA are effective to cluster the other groups of patterns.

This result indicates that the combination of PCA for data reduction with FCM, has more effectivity in comparison to the combination of K-means and PCA.



Figure 5.43.: 2-D projection of clustering feature patterns (PCA - FCM); • 1°pattern cluster, □ 2°pattern cluster, + 3°patterns cluster, × 4°patterns cluster.

Now, MDS is used as data reduction method in combination with FCM. Again the data input set is the same illustrated in figure 5.39). The clustering of the input patterns using FCM is showed in figure 5.44. The clustering accuracy RI = 0,86 is the same as the obtained using K-means as clustering method (Figure 5.40).
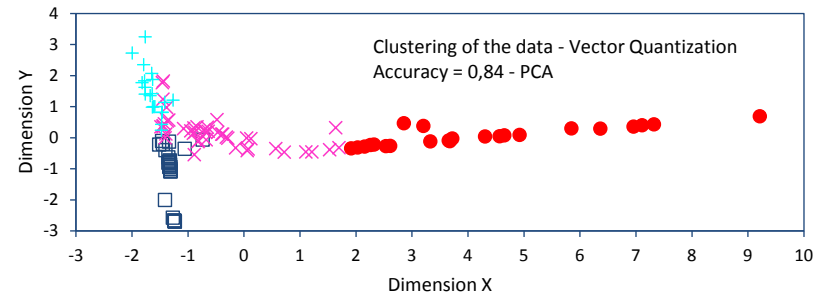


Figure 5.44.: 2-D projection of clustering feature patterns (MDS - FCM); • 1°pattern cluster, □ 2°pattern cluster, + 3°patterns cluster, × 4°patterns cluster.

The clustering results using MDS are same for both; K-means and FCM, thus the figure 5.40 and Figure 5.44 are practically the same with some differences between the members of each cluster.

In this case, some coal patterns are assigned wrongly to different clusters and confounded with patterns belonging to dead rock cluster. Also, all cluster are overlapped each other and outliers from all groups are obtained using this combination.

Finally, the LDA data reduction method is tested with FMC. Figure 5.45 shows the clustering data input which has an accuracy RI = 0,89. The input data can be seen in figure 5.41. LDA permits a complete separation of the coal patterns from the other clusters and in combination with FMC is effective to cluster coal pattern.
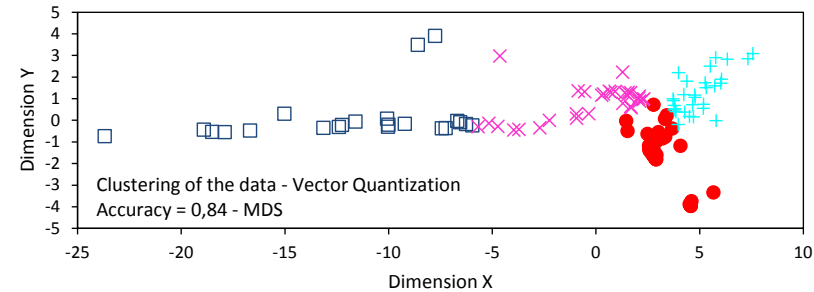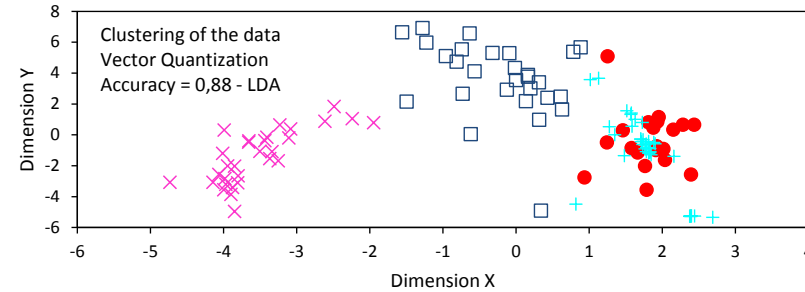


Figure 5.45.: 2-D projection of clustering feature patterns (LDA - FCM); • 1°pattern cluster, □ 2°pattern cluster, + 3°patterns cluster, × 4°patterns cluster.

However, figure 5.45 shows that except from coal patterns, some other patterns belonging to other clusters are confounded and misclassified.

**Vector quantization clustering**

Vector Quantization (VQ) is a method which models the probability density function by the distribution of prototype vectors (Section 4.6.5). VQ divides the input AE patterns into clusters having approximately the same number of AE patterns closest to them. VQ splits the data into small blocks or vectors, which are then encoded. The objective is to define a set of vectors or codebooks, which are representative of the input information. VQ encodes pairs up each source vector with the closest matching vector from the codebook, so quantizing it. Then each group is represented by its centroid point.

Figure 5.46 shows the clustering of the data using VQ and PCA for data dimension reduction. The input patterns correspond to the figure 5.37. The accuracy obtained with this clustering methods is RI = 0,84. Comparing the figure 5.37 with figure 5.46, it can be seen that some coal patterns are misclassified and confounded with dead rock patterns. All clusters of pattern are overlapped, although the accuracy obtained by PCA together with VQ can be considered as satisfactory.

Figure 5.47 shows the clustering of data, this time using MDS as data dimension reduction algorithm. The input patterns correspond to those illustrated in figure 5.39. Using MDS in

Figure 5.46.: 2-D projection of clustering feature patterns (PCA - VQ); ● 1°pattern cluster, □ 2°pattern cluster, + 3°patterns cluster, × 4°patterns cluster.

combination with VQ can be obtained a clustering accuracy RI = 0,84. MDS produces clusters with outliers of each cluster, making difficult to cluster the patterns. From figure 5.47, it can be seen that the outlier patterns are misclassified in other clusters and all clusters are overlapped each other.

It can be presumed that the outlier patterns presented in the input data are product of a wrong selection of AE features, but certainly this pattern distribution in the feature space is product of using MDS as data dimension reduction algorithm.

Considering that MDS is a technique, which analyzes a matrix of distances or dissimilarities and then produces a representation of the data points, it can be deduced that some features, which define the input patterns are not aligned with the features of their same type.



Figure 5.47.: 2-D projection of clustering feature patterns (MDS - VQ); ● 1°pattern cluster, □ 2°pattern cluster, + 3°patterns cluster, × 4°patterns cluster.

Finally, Figure 5.48 shows the clustering of the data input using LDA as data dimension reduction method and VQ as clustering method. The input data corresponds to the data illustrated in figure 5.41. The clustering accuracy is RI = 0,88. LDA together with VQ are effective in clustering coal patterns. However, this combination is not so precise when clustering other rock input

patterns.



Figure 5.48.: 2-D projection of clustering feature patterns (LDA - VQ); • 1°pattern cluster, □ 2°pattern
cluster, + 3°patterns cluster, × 4°patterns cluster.

The results presented till now are made using a set of input data (or input patterns) called
data input 1 and are summarized in table 5.4.

Table 5.4.: Data Input 1. Summary of the clustering accuracies (Rand Index), using different algorithms (8
features).

| Type of clustering method | Accuracy using PCA data reduction method | Accuracy using MDS data reduction method | Accuracy using LDA data reduction method | Features in each pattern |
|---|---|---|---|---|
| K-means | 0,77 | 0,86 | 0,88 | 8 |
| Fuzzy C-means | 0,90 | 0,86 | 0,89 | 8 |
| Vector Quantization | 0,84 | 0,84 | 0,88 | 8 |

From table 5.4, it can be observed that the best data dimension reduction is *Linear Discriminant
Analysis* (LDA) and the clustering method with the best accuracy is *Fuzzy C-means* (FCM).
It is important to remark that these results and conclusions are only applicable to the set of
input data 1. For other data set using the same feature vectors from figure 5.34, these results
should be similar.

To assure repeatability, two more data set are tested with the same data dimension reduction
and clustering algorithms. The results are presented in table 5.5 (for data input set 2) and
table 5.6 (for data input set 3). As can be seen, the results for others set of data are not exactly
the same as in data input set 1 in table 5.4. However, the main conclusion remains identical,
that is to say, the best clustering accuracy is obtained using FCM as clustering method and
the best data reduction method is LDA.

Table 5.5.: Data Input 2. Summary of the clustering accuracies (Rand Index), using different algorithms (8 features).

| Type of clustering method | Accuracy using PCA data reduction method | Accuracy using MDS data reduction method | Accuracy using LDA data reduction method | Features in each pattern |
|---|---|---|---|---|
| K-means | 0,80 | 0,80 | 0,86 | 8 |
| Fuzzy C-means | 0,85 | 0,85 | 0,89 | 8 |
| Vector Quantization | 0,80 | 0,80 | 0,86 | 8 |

Table 5.6.: Data Input 3. Summary of the clustering accuracies (Rand Index), using different algorithms (8 features).

| Type of clustering method | Accuracy using PCA data reduction method | Accuracy using MDS data reduction method | Accuracy using LDA data reduction method | Features in each pattern |
|---|---|---|---|---|
| K-means | 0,82 | 0,82 | 0,83 | 8 |
| Fuzzy C-means | 0,82 | 0,81 | 0,83 | 8 |
| Vector Quantization | 0,82 | 0,82 | 0,83 | 8 |

Table 5.7 summarizes the results obtained for the three data sets. The average accuracies of the three data set input are shown. Again the results and conclusions made previously are confirmed, that is FCM is the most suitable clustering method for rock cutting patterns and using LDA as data dimension reduction before clustering, the best clustering results are achieved.

Although, FCM is mentioned as the best clustering method, good results are obtained using K-means and VQ in combination with LDA too. Using PCA as data dimension reduction in the pre-processing of the data, satisfactory results are obtained depending of the data input set.

Even though the average clustering accuracies do not go beyond 90 porcent, these results can be attributed to the lack of capacity of clustering methods to classify gypsum and oil shale patterns.

To obtain better results, new features for oil shale and gypsum patterns should be examined. However, the clustering methods tested and in particular the FCM method, are efficient in clustering coal and dead rock patterns.

Table 5.7.: Average clustering accuracies (Rand Index), of Data input set 1, 2 and 3 (8 features).

| Type of clustering method | Accuracy using PCA data reduction method | Accuracy using MDS data reduction method | Accuracy using LDA data reduction method | Features in each pattern |
|---|---|---|---|---|
| K-means | 0,80 | 0,83 | 0,86 | 8 |
| Fuzzy C-means | 0,86 | 0,84 | 0,87 | 8 |
| Vector Quantization | 0,82 | 0,82 | 0,86 | 8 |

**Three-feature patterns**

Now, to evaluate the importance of the number of features that form every pattern, AE rock cutting patterns of just three features are tested. The patterns are formed as figure 5.49 illustrates. These patterns are a reduced version of the patterns illustrated in figure 5.34. This reduction in features is made in an attempt to shrink the computational resources needed to process the input data.

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} Number\ of\ peaks \\ Standard\ Deviation \\ Wavelet\ Packet\ energy \end{bmatrix} = \begin{bmatrix} NP = number\ of\ peak\ found\ in\ the\ data \\ \sigma = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \mu)^2}{n}} \\ WPE = \int_{-\infty}^{\infty} f(t)^2 dt = \sum_{j=1}^{n} E_j \end{bmatrix}$$

Figure 5.49.: Feature extraction from an AE signal. The vector $\boldsymbol{x}$ represents a pattern containing three features.

Basically, a subset of the original feature is selected. This is performed for the following reasons

1. Allow an easier subsequent analysis (improvements in speed and reduction of data requirements)

2. Provide a relevant subset of AE features

3. Increase classification performance and

4. Remove irrelevant or redundant features.

This feature selection is based on a quantitative analysis of the data. Furthermore, these three-feature patterns are subjected to feature extraction (data dimension reduction analysis) by means as PCA, MDS and LDA. The three-dimension data is reduced to a two-dimension data input. Feature extraction methods transform the original data (using all the original features) to a subset having a reduced number of features.

Having each pattern only with three features, these can be represented in a 3-D plot as can be seen in Figure 5.50. Each square represent a pattern and their tone (color) is just to visualize the space better, that is, the tone of each pattern does not signify that the pattern belongs to a group. As before, the input data was normalized and then illustrated.

The same analysis performed previously with patterns containing eight features is performed for these patterns. However, only the figures for the best clustering method are illustrated.

These three features in each pattern, describe the signals in the time domain (number of peak and standard deviation) and in the time frequency domain (wavelet packet energy). All data input set are normalized before their dimensions are reduced by means of PCA, MDS and LDA algorithms.
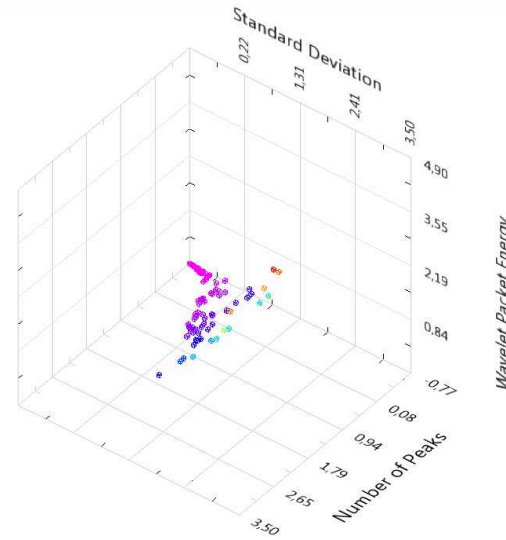


Figure 5.50.: 3-D plot of the AE input patterns (with three features).

Table 5.8 shows a summary of the clustering accuracies obtained using different clustering data reduction methods and data input set 1. As the results show, using AE patterns containing three features, there is a reduction in the clustering accuracies of all clustering methods. It is interesting to note that the effect of the data dimension reduction algorithm is less relevant using three-feature patterns. Also, the clustering accuracy using different clustering methods is slightly different.

The clustering accuracy for this input data set can be considered as unsatisfactory, since is about just 75 percent.

Table 5.9 and table 5.10 show the a summary of the clustering results using input data set 2

Table 5.8.: Data Input 1. Summary of the clustering accuracies (Rand Index), using different algorithms (3 features).

| Type of clustering method | Accuracy using PCA data reduction method | Accuracy using MDS data reduction method | Accuracy using LDA data reduction method | Features in each pattern |
|---|---|---|---|---|
| K-means | 0,74 | 0,74 | 0,76 | 3 |
| Fuzzy C-means | 0,77 | 0,77 | 0,77 | 3 |
| Vector Quantization | 0,72 | 0,74 | 0,75 | 3 |

and input data set 3 respectively. The same conclusions made before applies for these results and in general a reduction in the clustering accuracy is the most noticeable.

Table 5.9.: Data Input 2. Summary of the clustering accuracies (Rand Index), using different algorithms (3 features).

| Type of clustering method | Accuracy using PCA data reduction method | Accuracy using MDS data reduction method | Accuracy using LDA data reduction method | Features in each pattern |
|---|---|---|---|---|
| K-means | 0,73 | 0,73 | 0,79 | 3 |
| Fuzzy C-means | 0,76 | 0,76 | 0,78 | 3 |
| Vector Quantization | 0,73 | 0,73 | 0,77 | 3 |

Table 5.10.: Data Input 3. Summary of the clustering accuracies (Rand Index), using different algorithms (3 features).

| Type of clustering method | Accuracy using PCA data reduction method | Accuracy using MDS data reduction method | Accuracy using LDA data reduction method | Features in each pattern |
|---|---|---|---|---|
| K-means | 0,73 | 0,74 | 0,77 | 3 |
| Fuzzy C-means | 0,76 | 0,76 | 0,77 | 3 |
| Vector Quantization | 0,73 | 0,73 | 0,76 | 3 |

Table 5.11 shows a summary of the average clustering accuracies obtained from the three data input sets tested. These results corroborate the conclusions made previously and show that again the best clustering method to be used in AE patterns of rock cutting signals is Fuzzy C-means. Also LDA data dimension reduction method in combination with FCM gives the

best accuracy, that is RI = 0,77. Nonetheless, PCA and MDS methods have a similar accuracy as LDA.

The reduction in clustering accuracy indicates that the five features left out are relevant for clustering AE patterns and is convenient to work with them in order to achieve an acceptable classification performance.

Table 5.11.: Average clustering accuracies (Rand Index), of Data input set 1, 2 and 3 (3 features).

| Type of clustering method | Accuracy using PCA data reduction method | Accuracy using MDS data reduction method | Accuracy using LDA data reduction method | Features in each pattern |
|---|---|---|---|---|
| K-means | 0,74 | 0,74 | 0,77 | 3 |
| Fuzzy C-means | 0,76 | 0,76 | 0,77 | 3 |
| Vector Quantization | 0,73 | 0,73 | 0,76 | 3 |

Figure 5.51 illustrates the labeled data input set 2 in the feature space. The patterns are normalized and their dimension is reduced from three to two dimensions, using LDA algorithm.
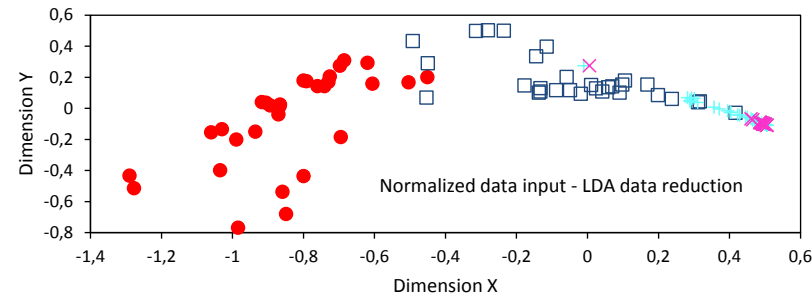


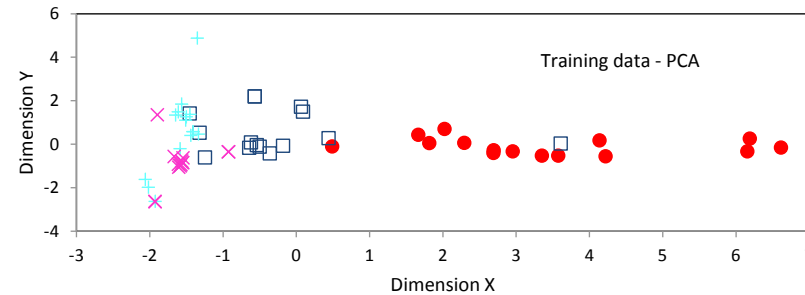Figure 5.51.: 2-D projection of clustering feature patterns (three features)(LDA); ● coal vector feature patterns, □ dead rock feature patterns, + Oil shale feature patterns, × Gypsum feature patterns.

As it can be seen in Figure 5.51, the groups of pattern are not clearly separated in the feature space. All the groups of patterns are overlapped and close to each other. This denotes a shortage of representation of each group using three-feature patterns.

Figure 5.52 illustrates the clustering of the data using LDA as data dimension reduction method and Fuzzy C-means as clustering method. The clustering accuracy RI = 0,78 is the highest obtained using three-feature patterns. However, this degree of accuracy is unsatisfactory to recognize pattern from AE signals in rock cutting.
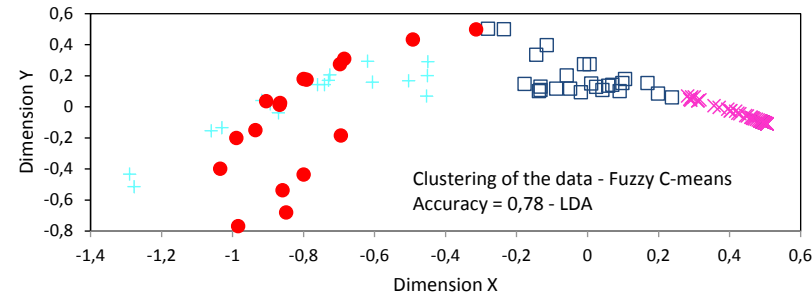
Figure 5.52.: 2-D projection of clustering feature patterns with three features (LDA - FCM); • 1°pattern cluster, □ 2°pattern cluster, + 3°patterns cluster, × 4°patterns cluster.

The clusters of the AE pattern groups are overlapped and all rock cutting patterns are confounded or incorrectly labeled. These results make AE patterns with these three specific features not recommendable for clustering.

### 5.3.3.4. Supervised classification

**Introduction**

To perform supervised classification, a labeled training data set must have available (section 4.5). In this case, the labels correspond to the cutting AE signal of each rock, that is to say, four labels. Once the input data is separated into training data and test data, a supervised algorithm is trained with the training data set and then tested with the test data set. The evaluation of the accuracy of this supervised algorithm is carried out by the classification accuracy (section 4.5.4).

As in the previous section (unsupervised classification), patterns containing eight features are employed as illustrated in figure 5.34. The data input data is normalized using the standard score normalization (equation 5.2) and then two feature extraction methods are used, namely

1. Principal Component Analysis (PCA) (section 4.4.4.1)

2. Multi-dimensional Scaling (MDS) (Section 4.4.4.3)

The algorithms are used to reduce the dimensionality of the data from eight to three features in the feature space.

This feature extraction using PCA is performed in an attempt to increase the computational speed and reduce redundancy.

In this work, three discriminant algorithms are used to perform supervised classification, namely

1. Back-propagation Neural Network algorithm

2. Linear Support Vector Machine algorithm (two-class)

3. Non-linear Support Vector Machine algorithm (two-class)

After these algorithms are employed, its performance is evaluated and compared.

As explained before, the most important material to be recognized and classified is the coal. Consequently, when using the two-class algorithms, the coal represents a class and the rest (other rocks) represent the second class.

**Back-propagation Neural Network**

This supervised classification algorithm is a nonlinear model, that is the output is a nonlinear function of its parameters and the inputs (section 4.5.3.3). It can be considered as an optimization of the Multilayer Perceptron model based on the calculation of an error function in the multilayer network.

Figure 5.53 illustrates the training data or labeled patterns used to train the supervised classification. In this case, the data is first normalized and then PCA is performed as data extraction method. As it can be seen, an outlier pattern pertaining to the dead rock group is located in the coal pattern zone in the feature space. Although, this situation is not ideal, it is a common situation when large amount of data is handled.

Unfortunately, the clusters are not good separated and the four clusters overlap each other especially the Oil Shale and Gypsum clusters.
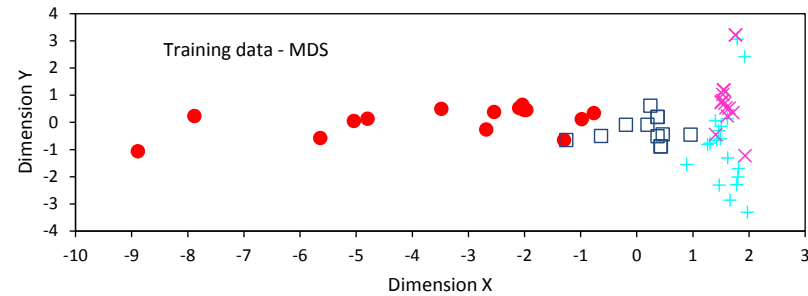


Figure 5.53.: Training data using PCA; • coal vector feature patterns, □ dead rock feature patterns, + Oil shale feature patterns, × Gypsum feature patterns.

Figure 5.54 illustrates the supervised classification of the test data or unlabeled patterns, using PCA as data extraction method in combination with BP Neural Network as supervised classification algorithm. The PCA algorithm reduces the data dimensionality to three dimensions. The number of hidden layers used in the BP Neural Network algorithm is five.

The classification accuracy obtained is 90 percent which is high enough to be considered as satisfactory. This accuracy is affected by outliers patterns from the Oil Shale and Gypsum clusters.

The data coal patterns from this test data are isolated from the rest, making easier the work to the BP Neural Network algorithm. Despite the classification accuracy being 90 percent, it can be concluded that only the coal patterns are classified satisfactorily and the other ones are confounded by this combination of algorithms.
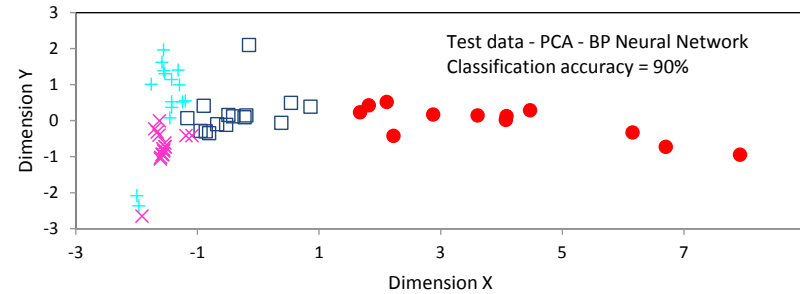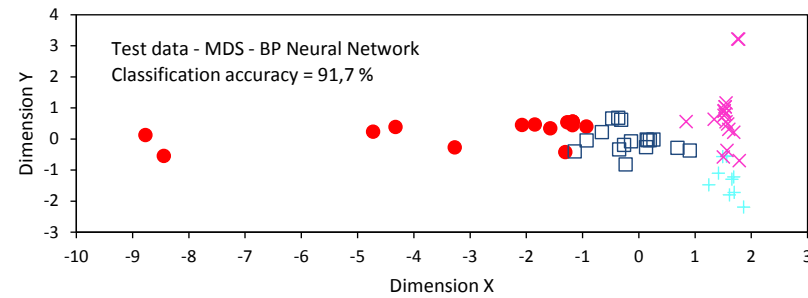


Figure 5.54.: Supervised classification of the test data using PCA - BP Neural Network; ● coal vector feature patterns, □ dead rock feature patterns, + Oil shale feature patterns, × Gypsum feature patterns.

Table 5.12 shows the classification accuracies of three set of data, namely data set 1 (analyzed above) and two data set more, namely data set 2 and data set 3, which are not illustrated. Also the average accuracy using this combination of algorithm (PCA and BP Neural Network) is calculated (92,5 percent).

Now *Linear Multi-dimensional Scaling* (MDS) algorithm is employed as data extraction method in combination with BP Neural Network to evaluate its influence on the general classification performance. MDS is set to reduce the data dimensionality from eight to three in the feature space. Again the number of hidden layer employed to perform the BP Neural Network algorithm is five.

Figure 5.55 illustrates the training data or labeled patterns, this time using MDS to reduce the dimensionality of the data. As it can be seen, again outlier patterns are obtained using MDS. All clusters are overlapped, even the coal cluster.

Figure 5.56 illustrates the supervised classification of the test data or unlabeled patterns using BP Neural Network as classification algorithm and MDS to reduce the data dimensionality. The classification accuracy obtained using this combination of algorithms is 91,7 percent, which is similar to the obtained previously using PCA.

Table 5.12 shows the classification accuracies of three set of data using MDS, namely data set 1 (analyzed above) and data set 2 and data set 3, which are not illustrated. Also the average accuracy using this combination of algorithm (MDS and BP Neural Network) is showed (90,2 percent).

Again it can be concluded that BP neural Network in combination with MDS algorithm, are effective only to identify and classify coal patterns reliably. Patterns from other clusters are

Figure 5.55.: Training data using MDS; • coal vector feature patterns, □ dead rock features patterns, + Oil shale features, × Gypsum features patterns.

confounded and are difficult to classify due to their proximity in the feature space.



Figure 5.56.: Supervised Classification of the test data using MDS - BP Neural Network; • coal vector feature patterns, □ dead rock features patterns, + Oil shale features, × Gypsum features patterns.

Analyzing the results from table 5.12, it can be concluded that using PCA as data extraction method before classifying with BP Neural Network gives slightly better classification accuracies than using MDS. Furthermore, PCA is easier to implement and it does not consumes much computational resources. MDS can be quite laborious to implement and takes more time to perform a calculation which is not desirable in continuous process.

Table 5.12.: Classification accuracies of data input set 1, 2 and 3. Using BP Neural Network algorithm.

| Clustering method | Data set 1 | Data set 2 | Data set 3 | Average accuracy |
|---|---|---|---|---|
| PCA | 90% | 93,3% | 91,2% | **92,5%** |
| MDS | 91,7% | 88,3% | 90,6% | **90,6%** |

**Linear support vector machine algorithm (two-class)**

In this subsection, linear SVM for two classes is evaluated. The first class corresponds to the coal patterns and the second class corresponds to the other patterns from the rest of rocks.

Before classification, the input data is normalized and its dimensionality reduced from eight to three, using *Principal Component Analysis* (PCA) algorithm.

Linear SVM algorithm projects the pattern into a higher dimension space in order to scatter them, and so makes it easier to find a linear classifier (see section 4.5.2.2). This algorithm produces linear decision boundaries, thus the input data must be linearly separable. Linear separable data must fulfill the equation 4.7.

The aim of the linear SVM algorithm is to find the best separation line or *hyperplane*, which separates the input data in two classes with maximal margins.

Figure 5.57 illustrates the training data or training patterns used to train the linear SVM algorithm.
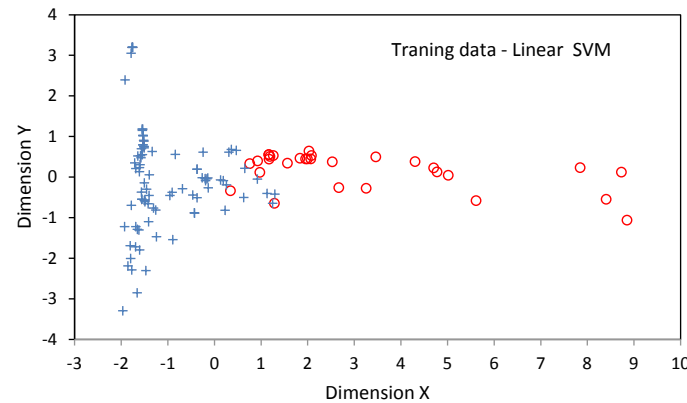


Figure 5.57.: Training data for Linear SVM. ● coal vector feature patterns, + Others rock patterns.

From this example, it can be observed that the data is not linearly separable, in other words there is not a possible straight line, which could separate the data or patterns into two groups in the feature space. However, SVM transforms the original space to other, where the data is separable.

In this case, after application of PCA algorithm, the data has three-dimensions and linear SVM must find a hyperplane to separate the two-class data. Normally, it is difficult to find or observe the hyperplane with maximal margins in the original feature space and even more, considering that the figure 5.57 is a two dimension representation. In the original feature space, the patterns are tight. Hence, linear SVM takes these patterns into a higher dimension space

and sparse them in this new dimension space. This transformation increases the possibility to find a hyperplane for the test data.

Linear SVM gives the support vectors and the maximal margins to separate the data. There are a lot of algorithms to perform this task. In this work, the *Sequential Minimal Optimization method* [Pla98] is employed.

The most important parameter and condition used in this algorithm are the marging failures $C$=0,1 and Tolerance of inaccuracies of the KKT conditions $Tol = 0,001$ respectively.

Figure 5.58 illustrates the application of the linear SVM over the test data or test pattern. The straight line represents the hyperplane in the new higher dimension space.
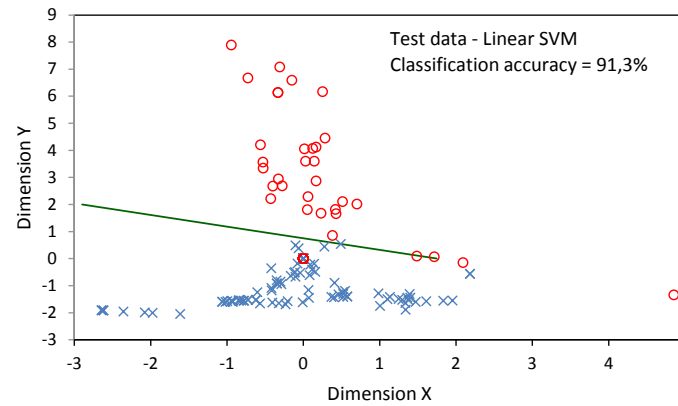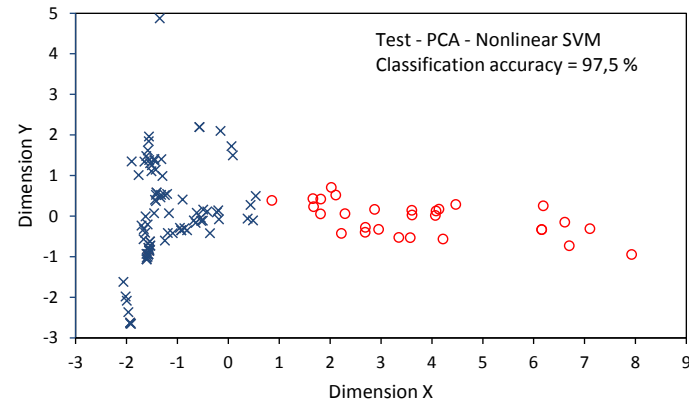


Figure 5.58.: Classification of the test data using linear SVM and PCA. ∘ coal vector feature patterns, ×
                 others rock patterns.

As can be seen, the hyperplane obtained is located over same patterns, indicating that no margins exists. Also a coal pattern is located at the other side of the hyperplane, apart from the other coal patterns.

These results indicate that the input data (training data and test data) is not appropriate to be classified using this algorithm. However, the classification accuracy obtained is superior i.e. 90%.

Table 5.13 shows the classification accuracies for two data sets (data set 2 and data set 3) apart from the data set presented and evaluated above (data set 1). The evaluation and illustration of the data set 2 and data set 3, gives the same problem, that is, no hyperplane with margins is found. This situation indicates that a non-linear classification method should be used for classification of this AE patterns.

In the following subsection, the same data sets are used, but this time employing nonlinear SVM algorithm.

Table 5.13.: Classification accuracies of data input set 1, 2 and 3. Using Linear SVM algorithm.

| Type of clustering method | Data set 1 | Data set 2 | Data set 3 | Average accuracy |
|---|---|---|---|---|
| PCA | 91,3% | 89,8% | 90,1% | **90,4%** |

**Non-linear support vector machine algorithm (two-class)**

Non-linear SVM is capable of classifying non-linear separable data. It transforms the feature space using non-linear functions (see section 4.5.3) and so the data can be separated in the new feature space.

The use of a technique known as the *kernel trick* permits employing SVM to nonlinear separable data. The algorithm is similar to the linear SVM algorithm except for the nonlinear kernel function included instead of computing the space transformation.

Figure 5.59 illustrates the training data used in the nonlinear algorithm. This corresponds to the same input data used in the previous section (figure 5.57), but used in other algorithm.
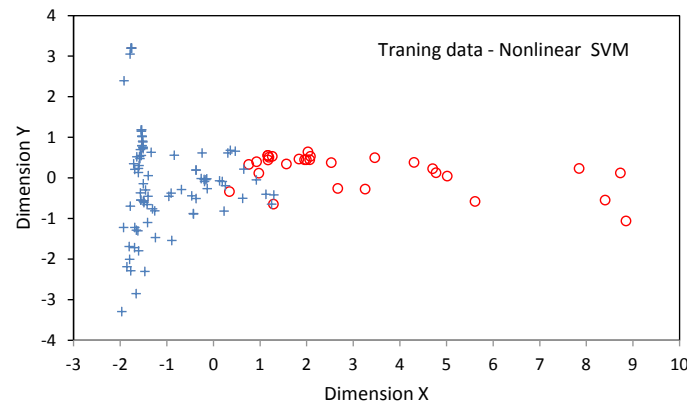


Figure 5.59.: Training data for nonlinear SVM. ● coal vector feature patterns, × Others rock patterns.

There are numerous kernel types that can be employed in the algorithm. In this work the Gaussian Kernel is employed due to the best results that are obtained in comparison with other kernels (Polynomial kernel, Sigmoid kernel).

The parameter of the kernel is 3, the regularization parameter for the training error, $C = 0,001$ and the soft margin parameter is also 3. These correspond to the most important parameters of the algorithm.

Again the dimension of the input data is reduced by means of the PCA algorithm after normalization. The dimensionality of the data is reduced from eight to three-dimension in the feature space before classification.

Figure 5.60 illustrates the classification of the AE patterns (test data), where the patterns are labeled with predicted class labels by the nonlinear SVM algorithm.



Figure 5.60.: Classification of the test data using non-linear SVM and PCA. ∘ coal vector feature patterns, × others rock patterns.

The classification accuracy is 97,5 percent, which is superior to the accuracies obtained using linear SVM. This degree of accuracy can be considered as satisfactory and could be used to recognize and classify AE patterns coming from rock cutting, specifically to separate and classify AE rock signals.

Table 5.14 shows the results of other two data sets, namely data set 2 and data set 3. The data set 1 corresponds to the results analyzed above.

Table 5.14.: Classification accuracies of data input set 1, 2 and 3 using non-linear SVM algorithm.

| Type of clustering method | Data set 1 | Data set 2 | Data set 3 | Average accuracy |
|---|---|---|---|---|
| PCA | 97,5% | 96,8% | 97,9% | **97,4%** |

### 5.3.4. Conclusions

The data input is normalized with standard score normalization. Six data sets are evaluated, the first three containing AE patterns of eight features each one. The other three containing

pattern of three features each one.

Using patterns containing eight features resulted in better clustering accuracies than using pattern of three features (in this specific case). The highest clustering accuracies achieved are about 86% (eight-feature patterns) and about 76% (three-feature patterns). These results suggest that the features selected are relevant and necessary for data clustering of rock cutting AE patterns.

Although the clustering accuracies do not surpass the 90%, the clustering of coal pattern, which are the patterns of interest in this study, could be clustered and isolated in a satisfactory manner, using LDA algorithm in combination with FCM algorithm.

In this work, the number of clusters is a known variable, that is why no cluster validity is performed. Patterns from four rock cutting signals are tested.

Based on the results of this work, some recommendations can be given for the improvement of process of clustering of AE patterns;

1. Plot the input in the feature space to understand its structure. The data should be normalized and the use of PCA could be useful.

2. Select the right feature from the raw data and try different patterns containing different amount of features.

3. Once the data is plotted, if possible, detect and remove outliers. This should raise the clustering performance.

4. Try different clustering methods and choose the appropriate. The accuracy of each method vary depending on each specific data input

For supervised classification three methods are evaluated, namely Back-propagation *Neural Network* algorithm, *Linear Support Vector Machine* algorithm (for two classes) and *Nonlinear Support Vector Machine* algorithm (for two classes).

Using BP Neural Network in combination with MDS as data extraction method gives clustering accuracies superior than 90%. However, this supervised classification algorithm is only effective in classifying coal patterns.

Linear SVM method is employed for classification. Specifically, the *Sequential Minimal Optimization method* algorithm is used. The results of three data sets show classifications accuracies superior to 90%. However, no separate line or hyperplane could be found between the coal patterns and the patterns of the other rocks. This result indicates that the AE patterns are non-linear separable and must be classified by a non-linear classification method.

Non-linear SVM algorithm is employed in combination with PCA to reduce the dimensionality of the data from eight to three-dimension. The classification accuracies obtained from three test data sets are superior than using a linear SVM classifier. Nonlinear SVM obtained very high classification performances, which makes it suitable for the classification of AE coal cutting patterns.

## 5.4. Summary

In this chapter, two experimental applications of AET and PRT are used in combination for resolving problems in mining process. Problems related to flotation process and rock cutting.

Unsupervised and supervised methods from PRT are employed. As unsupervised algorithms; *K-means clustering*, *Fuzzy C-means clustering* and *Vector Quantization clustering* are used. As supervised algorithm; *Back-propagation Neural Network*, *Linear Support Vector Machine* and *Nonlinear Support Vector Machine* are used.

In Unsupervised classification; *Rand Index* is used to evaluate the *clustering accuracy*, while in supervised classification; classification accuracy is used.

For data reduction dimensionality, three algorithms are used, namely Principal Component Analysis (PCA), *Multi-dimensional Scaling* (MDS) and *Linear Discriminant Analysis* (LDA).

**In flotation column** an accelerometer and an AE sensor are used to characterize the bubble activity. The signals obtained from the accelerometer are used to obtain information in the frequency domain and are correlated with the Minnaert`s equation to determinate the size of the bubbles immediately after they are formed. AE signals are employed successfully in the identification and classification of different air flows (bubble activity).

K-means clustering is used in combination with PCA and LDA to evaluate the feasibility of clustering AE patterns coming from different air flows. Both combinations provided high degree of clustering accuracy.

For supervised classification; Back-propagation neural network algorithm is used in combination with PCA. Using this combination high classification accuracy has been obtained and small changes in air flow are capable of being identified and classified.

**In Rock cutting**, unsupervised classification algorithms are evaluated, namely *K-means*, *Fuzzy C-means* and *Vector Quantization* in combination with different data extraction algorithms, namely PCA, MDS and LDA.

The best combination of clustering and data extraction algorithms is obtained using LDA algorithm for data dimension reduction and Fuzzy C-means algorithm for clustering of the data sets.

AE features patterns containing eight and three features are evaluated to compare their influence on the general classification accuracy of different data sets. Results show that best clustering accuracies are obtained employing eight-feature patterns. The clustering accuracy obtained with three-feature patterns can be considered as unsatisfactory.

In general the clustering accuracies obtained are not superior to 90%. However, the clustering algorithms evaluated showed good results for clustering coal patterns, which is the most important rock to be classified.

Also supervised method are evaluated, namely Back-propagation *Neural Network*, *Linear SVM* and *Nonlinear SVM*.

BP Neural Network is evaluated in combination with two data extraction algorithms; PCA and MDS Algorithms. The classification accuracies of these combinations are similar, being the classification accuracy of PCA slighter superior than 90%.

Linear and nonlinear SVM algorithm (two-class) are tested, both of them in combination with PCA algorithm for data dimension reduction. Since the aim of the rock cutting pattern classification is to recognize, classify and isolate the coal signal, the first class is assigned to the coal cutting signals and the second class is assigned to the rest of rock cutting signals.

Using linear SVM algorithm, the results were not satisfactory due to the nonlinearity of the AE data. On the contrary, using nonlinear SVM algorithm the results were satisfactory showing that this type of classifier are suitable for AE signal of coal cutting.

Pattern recognition is proven to be an effective method to work in combination with AET. Data analysis and signal processing can be improved with this method for automatic pattern recognition of multiple AE signals picked up from mining machinery.

# 6. Conclusions and outlook

## 6.1. Summary and general conclusions

In this work, different AE patterns are developed and proposed as a means of recognizing and classifying different states and process in mining machinery. Acoustic Emission Technique (AET) and Pattern Recognition Technique (PRT) are described and analyzed to be used in combination. Signal processing is used to process and form parameters in order to identify signal features. These features are used as raw data for the PRT. Features in time domain, frequency domain as well as time-frequency domain are selected. Statistical parameters, waveform parameters and frequency parameters are evaluated as features of signals from AE events of interest.

The implementation of AET in PRT requires the correct choice of AE parameters. For this task, AE emission source waves are measured directly by an AE sensor using a data acquisition system. All conditions and mechanisms in the process or object measured are taken into account. This combination of techniques can be used to achieve an assessment and prediction tool for mining machinery. Basically, the objective is to identify and classify similar AE signal from the operation of mining machinery.

The importance of doing these types of experimental works resides in the fact that PRT cannot interpret itself the results and it needs the proper interpretation of the results and examples of use to correlate the AE patterns with processes of interest in the machine. In this way, the results presented in this work can be used as a guide to obtain a satisfactory classification accuracy of AE signals in a specific mining process being monitored.

As explained in section 4 the first stage in the pattern recognition process is to select the correct and suitable AE features or parameters to characterize the process. This is actually, the most important step in the process and it is critical to achieve good results. This feature selection stage requires time to find the right combination that in further stages of the process is used to form the AE patterns or fingerprints of the process. Normally, this task can be supported by using graphical representations of the features, mathematical analysis of the data as well as the operational experience to correlate the AE features with the AE events being observed. In this work, graphical representation and analysis of the data are used to select the AE features.

After data collection, the normalization of the data is an important stage, which serves to compare AE patterns in the feature space. Normalization permits comparing AE features with different ranges and different units of measurement. The method used in this work for

normalization is Standard Score Normalization. This step is fundamental to achieve a correct selection of features from the signals.

Feature extraction is used to transform the complete set of selected features into a reduced set of features. This helps to have a low-dimension representation with minimum loss of information, reduce redundancy, enhance discrimination, increase the computational speed and obtain significant underlying features. In this work, three algorithms for data reduction are used, namely *Principal Components Analysis* (PCA), *Linear Discrimination Analysis* (LDA), and *Multidimensional Scaling* (MDS). Between these data extraction methods, PCA resulted to be the easiest to apply and its performance is comparable with the others methods.

When a set of data samples is available in advance, a supervised classification method can be used to classify AE patterns. This set of data must be associated with labels or class types. This supervised classifying algorithm assigns labels to the input data. Two approaches exist to supervised classification namely via *Decision Theory* and *discriminant analysis*. In this work, only a discriminant analysis is employed for supervised classification, namely Support Vector Machine (two-class), Nonlinear Support Vector Machine and Back-propagation Neural Network. All these methods are evaluated using classification accuracy. Their performance depended on the type of data being evaluated as well as the type of input patterns (linear or nonlinear).

On the contrary, clustering is used to explore the data in order to find groups of patterns and use them in a supervised method. Clustering can be based on dissimilarities matrices or measurements on individuals. Both of them explore the data to find groups and label them. After clustering a defined structure of data is achieved. Each clustering method imposes a specific structure to the data. A disadvantage is that clustering methods produces groups from the data even if there are not natural groups in the data, thus it is recommendable to have an idea of the possible AE clusters that could be found in the data. Three clustering methods were used and evaluated in this work, namely K-means clustering, Fuzzy C-Means Clustering and Vector Quantization Clustering. For the evaluation of the clustering performance a measurement of the similarity between two clusters was used, that is to say the Rand Index (RI).

In this work two experimental applications were performed namely AET in column flotation cells and AET in rock cutting. Both of them were analyzed using a combination of AET and PRT.

In the case of column flotation, AET and accelerometers sensor are used to monitor the bubble activity and bubble size, as a means of improving the efficiency of the column flotation process. The signals obtained from the accelerometer contained information in the frequency domain and could be correlated with the Minnaert`s equation to determine the size of the bubbles immediately after they are formed.

Also, changes in air flow could be detected, making possible the bubble activity monitoring inside of the column flotation. For this purpose three features are selected to characterize each AE signal, namely *RMS*, *Variance* and *Entropy*. For reducing of the data dimension from

three to two-feature space, *Principal Component Analysis* (PCA) and *Linear Discrimination Analysis* (LDA) are evaluated.

For unsupervised classification, K-means clustering was employed to classify the AE signals emitted by bubbles as a product of different air flows. The aim is to detect changes in the kinetics of the process to avoid turbulence inside of the column flotation. For this purpose, three air flows are used, namely flow A, flow B and Flow C. High clustering accuracies could be obtained, suggesting that clustering is a suitable way to identify different flows and bubble activity monitoring in column flotation cells.

For supervised classification *Back-Propagation Neural Network* algorithm is evaluated in combination with PCA as data extraction method. Using this combination high classification accuracy has been obtained and small changes in air flow are possible to be identified and classified.

Employing, this procedure changes in the air flows could be detected and classified, in form of AE activity from the air bubbles. However, it is advisable to a have a reference to interpret the results.

In rock cutting, AET is used to identify the rocks being cut as a means of automation in the operation of underground mining and tunneling. AET together with PRT are used to classify AE signal coming from different cutting process of different rocks, namely coal, dead rock, oil shale and gypsum. First, different clustering methods are evaluated; K-means clustering, Fuzzy C-means clustering and Vector Quantization clustering. Also for data dimension reduction three algorithms are evaluated, namely Principal Component Analysis (PCA), Multi-dimensional Scaling (MDS) and Linear Discriminant Analysis (LDA). For this AE patterns Fuzzy C-means clustering in combination with LDA had the best performance to classify AE signals of rock cutting.

Patterns containing eight and three features are evaluated to test their capacity of identifying the AE signal. Findings show that the highest classification accuracies are obtained using patterns of eight features. This indicates that the selected eight features are not redundant. Thus, the search for new parameters that identify the rock cutting not only from the presented rock but from others is recommendable. The selection of these parameters is fundamental for the classification results.

Furthermore, for supervised classification three methods are evaluated, namely *Back-propagation Neural Network algorithm*, *Linear Support Vector Machine algorithm* (for two classes) and *Nonlinear Support Vector Machine algorithm* (for two classes). This supervised classification methods are evaluated in combination with the same data reduction algorithms used in unsupervised analysis. The performance of the *Back-propagation Neural Network algorithm* is very high, thus it is recommendable to use as well as *Nonlinear Support Vector Machine algorithm* (for two classes). However, since these AE patterns are not linear, the performance of *Linear Support Vector Machine algorithm* (for two classes) is not satisfactory.

## 6.2. Outlook

In Acoustic Emission Technique new types of AE sensors can be used to explore other frequency range and try to obtain further information about the processes. Also, the influence of the sampling rate in the wave form could be tested. Different AE sensors, especially manufactured for the process being measured should be tested.

In Signal Processing Technique new parameters should be found in time and frequency domain as well as in time-frequency domain. These features must be representative of each process or AE event of interest.

In Pattern Recognition Technique new classifying algorithms should be developed for unsupervised classification as well as supervised classification. These algorithms must enable monitoring in real time and should not demand huge computational resources to be run.

PRT limitations should be taken into account in order to use this technique in a proper manner with AET. Limitations caused by a wrong selection of signal features in the pre-processing of the data is the most important issue to be considered.

For each process of interest, a specific classifier should be developed, because each system possesses its own characteristics as background noise and different AE sources.

Using supervised classification, a condition record or data base can be formed for specific machines. This allows automatic signal classification for all possible condition of the machine being monitored. Also, using unsupervised classification methods can be useful to classify unfamiliar AE signals and form label data for further supervised analysis.

It is recommendable to carry out new tests in a real mining process context to verify if these techniques are suitable to monitor real-life operations. However, the experimental tests presented in this work in a laboratory scale show promising results for material identification in rock cutting and signal clustering in column flotation.

# A. Appendix: measurements of dissimilarity

Pattern recognition technique requires many times, a measurement of the distance of dissimilarity between two pattern or feature vectors.

Particular classes of dissimilarity functions known as dissimilarity coefficients are required to satisfy some conditions. If $d_{ab}$ is the a measurement of dissimilarity between the object $a$ from object $b$,then

| | |
|---|---|
| $d_{ab} \geq 0$ | for every $a, b$ |
| $d_{aa} = 0 \ 1$ | for every $a$ |
| $d_{ab} = d_{ba}$ | for every $a, b$ |

For numerical variables, different dissimilarity measurements are proposed and listed in table A.1. The selection of a particular way of measurements depend on the application and computational resources. Normally, the use of one other measurement is determined by the best performance in terms of classification error on the test data.

Suppose two patterns $\boldsymbol{x}_i^a = [x_1^a, x_2^a, ..., x_p^a]$ and $\boldsymbol{x}_i^b = [x_1^b, x_2^b, ..., x_p^b]$, each of them with $p$ number of features. The distance between them are defined variously as

Table A.1.: Distance between two single patterns.

| Type of dissimilarity measurement method | Equation |
|---|---|
| Euclidean Method | $d = \sqrt{\sum_{i=1}^{p} (\boldsymbol{x}_i^a - \boldsymbol{x}_i^b)^2}$ |

| Normalized Euclidean distance ($V$ is a diagonal matrix whose diagonal elements are standard deviations) | $d = \sqrt{V \sum_{i=1}^{p} (\boldsymbol{x}_i^a - \boldsymbol{x}_i^b)^2}$ |
|---|---|
| Mahalanobis distance ($C$ is the covariance matrix) | $d = \sqrt{C \sum_{i=1}^{p} (\boldsymbol{x}_i^a - \boldsymbol{x}_i^b)^2}$ |
| City block distance | $d = \sum_{i=1}^{p} \left| \boldsymbol{x}_i^a - \boldsymbol{x}_i^b \right|$ |
| Minkowski distance of order m | $d = \left[ \sum_{i=1}^{p} (\boldsymbol{x}_i^a - \boldsymbol{x}_i^b)^2 \right]^{\frac{1}{m}}$ |
| Chebyshev distance | $d = \max_i \left| \boldsymbol{x}_i^a - \boldsymbol{x}_i^b \right|$ |
| Canberra distance | $d = \sum_{i=p}^{p} \frac{\left| \boldsymbol{x}_i^a - \boldsymbol{x}_i^b \right|}{\boldsymbol{x}_i^a - \boldsymbol{x}_i^b}$ |
| Quadratic distance | $d = \sum_{i=1}^{p} \sum_{j=1}^{p} (\boldsymbol{x}_i^a - \boldsymbol{x}_i^b) Q_{ij} \boldsymbol{x}_j^a - \boldsymbol{j}_i^b \ withQ > 0$ |
| Non-linear distance | $d = H \ if \ d > D \quad and \quad d = 0 \ if \ d \leq D$ |

# Bibliography

[AT00]     C. Aldrich and D. A. Theron. Acoustic Emission for the particle size distribution of sulfide ores in a laboratory SAG mill. *The Journal of The South African Institute of Mining and Metallurgy*, pages 243–248, 2000.

[BBE09]    A. Berezovski, M. Berezovski, and J. Engelbrecht. Waves in inhomogeneous solids. *Applied wave mathematics, Spring-Verlag Berlin Heildelberg*, 2009.

[Bed89]    M. D. Bedworth. Improving upon standard pattern classification algorithms by implementing them as multi-layer perceptrons. Speech Research Unit, SP4, Royal Signal and Radar Establishment, 1989.

[Bez81]    J. C. Bezdek. Pattern recognition with fuzzy objective function algorithms. Plenum Press, New York, 1981.

[Bio62]    M. A. Biot. Generalized theory of acoustic propagation in porous dissipative media. *The journal of the acoustic society of America*, vol 34(9), 1962.

[Bis06]    C. M. Bishop. *Pattern Recognition and Machine Learning.* Springer Science+Business h. Media, LCC, 2006.

[BNGB12]   F. D. Boos, K. Nienhaus, K. Garate, and R. Baltes. Development of acoustic emission (ae) based defect parameters for slow rotating roller bearings. *25th International Congress on Condition Monitoring and Diagnostic Engineering. Journal of Physics:*, 364, 2012.

[Boc85]    H. H. Bock. On some significance tests in cluster analysis. *Journal of Classification*, 2(1):77–108, 1985.

[BSS13]    K. Blazek, C. Stolk, and W.W Symes. A mathematical framework for inverse wave problems in heterogeneous media. *Inverse Problems*, 29(065001):37, 2013.

[CD77]     M. Comninou and J. Dundurs. Reflexion and refraction of elastic waves in presence of separation. In *Proc. R. Soc. Lond*, volume 356, pages 509–528, 1977.

[CH08]     T.W.S. Chow and D. Huang. Data Reduction for Pattern Recognition and Data Analysis. *Studies in computational Intelligence*, vol.115:pp.81–109, 2008.

[CL92]     K. H. Choi and W. K. Lee. Behavior of gas bubbles in a concentric cylindrical airlift column. *Korean J. of Chem. Eng.*, 9(2):66–73, 1992.

[CMH09]    D. Crosland, R. Mitra, and P. Hagan. Changes in acoustic emission when cutting difference rock types. Coal Operators Conference. University of Wollongong, 2009.

[Dav02]    C. S. Davis.  Statistical methods for the analysis of repeated measurements. Springer-Verlag New York, Inc, 2002.

[DDJ94]    L. David, I. Donoho, and M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. In *Biometrika*, volume 81, pages 425–455, 1994.

[DGL96]    L. Devroye, L. Gyorfi, and G. Lugosi. A probabilistic theory of pattern recognition. Springer Verlag, Berlin, 1996.

[Dro94]    T.F. Drouillard.  Acoustic Emission. The First Half Century.  12th International Acoustic Emission Symposium, Sapporo, Japan., 1994.

[Dun73]    J. C. Dunn. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3(3):32–57, 1973.

[EGB10]    F. Elamin, F. Gu, and A. Ball. Diesel engine injector faults detection using acoustic emissions technique. *Journal of Modern Applied Science*, 4(9), 2010.

[EJ57]     M. Ewing and W. Jardetzky. *Elastic waves in a layered media*. McGraw-Hill, New York, 1957.

[ELHC06]   C. Ennaceur, A. Laksimi, C. Herve, and M. Cherfaoui. Monitoring of crack growth in preassure vessel steel by the acoustic emission technique and the method of potencial difference. *Int. J. of Pressure vessels and pipeline*, 83(3):227–243, 2006.

[EMM98]    J. Eisenblätter, J. Manthei, and G. Meister.  Monitoring of microcrack formation around galleries in salt rocks. In *Proceeding 6th conference on acoustic emission/microsismic activity in geologic structures and materials*, pages 227–243, 1998.

[Fal87]    K. Fallenius.  Turbulence in flotation cells. *International Journal of Mineral Processing. Elsevier Science Publishers B.V., Amsterdam*, 21(38):01–23, 1987.

[FD91]     J. A. Finch and G. Dobby. Column flotation: A selected review. part i. *International Journal of Mineral Processing*, 33:343–354, 1991.

[Fle88]    R. Fletcher. Practical methods of optimization. Wiley, New York, 1988.

[FM90]     T. Folkestad and K. Mylvaganam. Acoustic measurements detect sand in north sea flow lines. Chr. Michelsen institute. Bergen, Norway, 1990.

[FP93]     K. Fahy and E. Perez. Fast fourier transformation and power spectra in labview. National Instrument Corporation, Application Note 40, 1993.

[FRHE01]   M. Fernandez-Redondo and C. Hernandez-Espinoza. Weight methods for multilayer feedforward. *European Symposium on Artificial Neural Networks*, pages 119–124, 2001.

[Fur97]    S. Furui.  Recent advances in speaker recognition. *Pattern recognition Letters*, vol.18:859–872, 1997.

[FY98]     S. D. Fallls and R. P. Young. Acoustic emission and ultrasonic-velocity methods used to characterize the excavation disturbance associated with deep tunnels in hard rock. *Tectonophysics*, 289:1–15, 1998.

[GK78]  N. Guz and V. D. Kubenko. Diffraction of elastic waves. *Institute of mechanics*, vol 14(8):3–15, 1978.

[GNKB12]  K. Garate, K. Nienhaus, C. Klein, and F. D. Boos. Acoustic emission as tool for material identification in rock cutting. In *21st International Acoustic Emission Symposium ( IAES21).Okinawa-Japan*, pages 199–204, 2012.

[GO10]  C.U. Grosse and M. Ohtsu. *Acoustic Emission Testing*. Springer,Verlag Berlin Heidelberg, 2010.

[Gor96]  A. D. Gordon. Null models in clustering validation.from data to knowledge. *Studies in classification, Data Analysis, and Knowledge Organization*, pages 32–44, 1996.

[Gra77]  K. F. Graff. Wave motion in elastic solids. In *Oxford University Press*, volume 356, pages 509–528, 1977.

[Gra84]  R. M. Gray. Vector quantization. IEEE ASSP Magazine, 1984.

[HMF10]  M. C. He, J. L. Miao, and J. Feng. Rock burst process of limestone and its acoustic emission characteristics under true-triaxial unloading conditions. *Int. J. of rocks Mech. and min. sci*, 47:286–298, 2010.

[Hud80]  J. A. Hudson. *The excitation and propagation of elastic waves*. Cambridge University Press, 1980.

[IKAK12]  K. Ito, H. Kuriki, H. Araki, and S. Kuroda. Noise-resistant method of AE events detection with multiple thresholds and its application to plasma spraying process. In *Progress in Acoustic Emission XVI*, 2012.

[ISO01]  ISO12716. Non-destructive testing. Acoustic Emission inspection. Vocabulary. Technical report, The International Organization for Standardization, 2001.

[JD88]  K. Jain and R. C. Dubes. Algorithms for clustering data. Prentice-Hall, Englewood Cliffs, New Jersey, ISBN 0-13-022278-X, 1988.

[JDM00]  A. K. Jain, R. P. W. Duin, and J. Mao. Statistical pattern recognition: A review. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 22, 2000.

[Kai50]  Joseph Kaiser. *Untersuchungen über das Auftreten Geräuschen beim Zugversuch*. PhD thesis, Technische Hochschule München, 1950.

[Kin82]  L. Kinler. Fundamentals of acoustic. 3rd edition. New York, chichester: Wiley, 1982.

[KMP07]  N. Kazakis, A. Mouza, and A. Paras. Experimental study of bubble formation at porous spargers. 6th International Conference on Multiphase Flow, ICMF, 2007.

[Kru64]  J. B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, vol.29(1), 1964.

[Lab11]  LabVIEW. Example: time frequency spectrogram. national instruments, 2011.

[Leo00]    M. Leonard. Comparison of manual and automatic onset time picking. Bulletin of
           the Seismological Society of America, 90, 6, pp. 1384,1390, 2000.

[Lin73]    R. F. Ling. A probability theory of cluster analysis. *Journal of American Statistical
           Association*, 68(341):159–164, 1973.

[LM12]     R. J. Larsen and M. L. Max. *An Introduction to Mathematical Statistics and its
           Application*. Prentice Hall, 2012.

[Lo12]     I. Lee and others. Evaluation of rock integrity using Fourier and wavelet transforms.
           *Tunnelling and Underground Space Technology*, 28:304–314, 2012.

[Lub89]    S. D. Lubetkin. Measurement of bubble nucleation rates by an acoustic method.
           *Journal of applied electrochemistry*, 19:668–676, 1989.

[Lyo04]    R. Lyons. Understanding digital signal processing. *2nd ed. Prentice Hall, New
           Jersey*, pages 603–611, 2004.

[M+99]     R. K Miller et al. A reference standard for the development of acoustic emission
           pipeline leak detection technique. In *NDT and E International*, volume 32, pages
           1–8, 1999.

[McL92]    G. J. McLachlan. *Discriminant analysis and statistical pattern recognition*. New
           York, Wiley, ISBN: 0-471-61531-5, 1992.

[MFD+01]   R. Manasseh, R. F. Fontaine, J. Davy, I. Shepherd, and Y. Zhu. Passive acoustic
           bubble sizing in sparged systems. *Experiments in Fluids*, 30(6):672–682, 2001.

[MGN12]    C. Molina, K. Garate, and K. Nienhaus. Acoustic emission in planetary gearboxes:
           Mathematical model and analysis of experimental measurements. *21st International
           Acoustic Emission Symposium ( IAES21).Okinawa-Japan*, pages 49–54, 2012.

[Min93]    M. Minnaert. On musical air bubbles and the sound of running water. *Philosophical
           Magazine*, 16:235–248, 1993.

[MLMC05]   A. Mesgouez, G. Lefeuve-Mesgouez, and A. Chambarel. Simulation of Transient
           Mechanical Wave Propagation in Heterogeneous Soils. *ICCS, LNCS*, 3415:647–624,
           2005.

[Mou09]    O. Mouraille. *Sound propagation in dry granular materials: discrete elements simu-
           lations, theory, and experiments*. PhD thesis, Doctoral thesis University of Twente,
           Enschede, 2009.

[MYR98]    S. C. Maxwell, R. P. Young, and R. S. Read. A micro-velocity tool to assess the
           excavation damaged zone. *International journal of rock mechanics and mining
           sciences*, 35(2):235–247, 1998.

[Nis72]    R. Y. Nishi. Ultrasonic detection of bubbles with doppler flow transducers. *Ultra-
           sonic*, 10:173–179, 1972.

[NKBB12]   K. Nienhaus, C. Klein, F. D. Boos, and R. Baltes. Acoustic Emission and pick
           forces in rock cutting. In *Aachen International Mining Symposia (AIMS)*, 2012.

[NKBR13]  M. Nienhaus, Warcholik, K., F. D. Boos, and D. Röllinger. Emission analysis of cutting tools with regard to material identification. In *Aachen international mining symposia. Fourth International Symposium Mineral Resources and Mine Development (AIMS)*, 2013.

[Oco90]  C. T. Oconnor. Measurement of the effects of physical and chemical variables on bubble size. *International Journal of Mineral Processing*, 28:139–149, 1990.

[OJ99]  J.O. Owino and L.J Jacobs. Attenuation measurements in cement-based material using laser ultrasonic. *Journal of Engineering Mechanics*, vol 125:637–647, 1999.

[PDFR02]  B. Pyke, J. Duan, D. Fornasiero, and J. Ralston. From turbulence and collision to attachment and detachment: A general flotation model. flotation and flocuation. from fundamentals to applications. Hawaii, 2002.

[Pea01]  K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, vol.2:pp.559–572, 1901.

[PFTV92]  W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes. The Art of Scientific Computing. 2th Edn.* Cambridge University Press, 1992.

[Pla98]  J. C. Platt. Sequential minimal optimization: A fast algorithm for training support vector machines. Microsoft Research. Technical Report MSR-TR-98-14, 1998.

[PS05]  N. H. Pontoppidan and S. Sigurdsson. Independent components in acoustic emission energy signals from large diesel engines. *International Journal of COMADEM*, 2005.

[RAC01]  R. B. Randall, J. Antoni, and S. Chobsaard. The relationship between spectral correlation and envelope analysis in the diagnostics of bearing faults and other cyclostationary machine signals. *Mechanical Systems and Signal Processing*, vol 15(5):945–962, 2001.

[Ran71]  W. H. Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850, 1971.

[RHW86]  D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning internal representations by error propagation. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, MIT Press, Cambridge, MA*, vol.1:318–362, 1986.

[Ron79]  C. Rong. Acoustic emission of rocks under triaxial compression with various stress paths. *Int. J. Rock Mech. Min. Sci. and Geomech*, 16:401–405, 1979.

[S$^+$11]  S. J. Spencer et al. Monitoring of jemeson cell flotation performance by passive acoustic emission. *The 5th international conference: flotation 11. Cape town*, 2011.

[Sam69]  J. W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers*, vol.C-18(5), 1969.

[Sat98]      S. R. S. Satri. Column flotation: Theory and practice. *IIME, Jamshedpur*, pages 44–63, 1998.

[SBK58]      B.H. Schofield, R.A. Bareiss, and A.A. Kyrala. Acoustic emission under applied stress. WADC Technical Report 58-194, Lessells and associates, Boston, Massachusetts, 1958.

[SCWL99]     S. J. Spencer, J. J. Campbell, K. L. Weller, and Y. Liu. Acoustic emission monitoring of sag mill performance. In *Intelligent Processing and Manufacturing of Materials (IPMM99)*, volume 2, pages 939–946, 1999.

[Sha48]      C. E. Shannon. A mathematical theory of communication. In *Bell System Technical Journal*, volume 3, pages 379–423, 1948.

[SHEE05]     T. Spies, J. Hesser, J. Eisenblätter, and G. Eilers. Measurements of acoustic emission during backfilling of large excavations. In *Proceeding 6th symposium rockbursts and seismicity in mines (RaSiM 6)*, pages 379–383, 2005.

[SHK97]      H. Shen, H. Hardy, and A. Khair. Laboratory study of acoustic emission and particle size distribution during rotary cutting. *Int. J. Rock Mech. and Min. Sci.*, 34(121):3–4, 1997.

[SK75]       L. Z. Shuck and T. W. Keech. Monitoring of acoustic emission from propagating fractures in petroleum reservoir rocks. Conference on acoustic emission (microseismic activity) in geologic structures and materials, University Park, PA, USA., 1975.

[SMTD01]     W.K. Sakamoto, P. Marin, D. Tunnicliffe, and D.K. Das. Lead zirconate titanate/polyurethane (pzt/pu) composite for acoustic emission sensor, 2001.

[SW13]       W. Sikorski and K. Walczak. *Power Transformer Diagnostics Based on Acoustic Emission Method, Acoustic Emission - Research and Applications*. InTech, DOI: 10.5772/55211, 2013.

[TZA08]      P. Theobald, B. Zeqiri, and J. Avison. Couplants and their influence on AE sensor sensitivity. *Journal of Acoustic Emissiong*, 26, 2008.

[Web02]      R. Webb. *Statistical Pattern Recognition*. John Wiley and Sons, ISBN 0-470-84514-7, 2002.

[WF05]       S.M. Walley and J.E. Field. *Elastic Wave Propagation in Materials*. p.p. 1-7, 2005.

[Wil02]      C. K. I. Williams. On a connection between kernel pca and metric multidimensional scaling. *Machine Learning*, vol.46:11–19, 2002.

[WNM06]      B. A. Willis and T. J. Napier-Munn. *Mineral Processing Technology*. 7th Edition. Butterworth-Heinemann, 2006.

[WP89]       H. Wu and G. K. Patterson. Laser-doppler measurements of turbulent-flow parameters in a stirred mixer. *Chemical Engineering Science*, 44:2207–2221, 1989.

[Z⁺06]    H. Zejli et al. Detection of broken wires in the cables hidden parts (anchorings) by acoustic emission. *Advanced Materials Research*, 13:345–350, 2006.

[ZSC12]   W. Zhang, S. J. Spencer, and P. Coghill. An acoustic technique for measurement of bubble solids mass loading : (a) fundamental study of single bubble. *Minerals Engineering*, 36(38):36–38, 2012.

# List of Figures

# List of Tables

# Curriculum Vitae

**Personal data**

| | |
|---|---|
| Name: | Kenny Alexander Gárate Peñaranda |
| Date of birth: | 21 February 1983 |
| Place of birth: | Arica, Chile |
| Nationality: | Chilean |

**Educational Career**

| | |
|---|---|
| 01/2011 - 03/2014 | Doctoral Studies at the Institute for Mining and Metallurgy |
| 07/2007 - 08/2008 | Exchange Student at the RWTH Aachen University |
| 03/2001 - 03/2007 | Mechanical Engineer, University of Tarapacá, Arica, Chile |

**Internship**

| | |
|---|---|
| 07/2006 - 12/2006 | Trainee Engineer, EMIN S.A., Calama, Chile |
| 01/2006 - 03/2006 | Trainee Engineer, Shipyard S.A., Arica , Chile |

**Work Experience**

| | |
|---|---|
| Since 04/2014 | Project Director, AMT Solutions Ltda., Chile |
| 01/2011 - 03/2014 | Research Assistant at IMR, RWTH Aachen University, Germany |
| 10/2008 - 06/2009 | Project Engineer, Alto Norte Foundry, Xstrata copper, Antofagasta, Chile |
| 06/2009 - 12/2010 | Planner Engineer, Joy Global Chile, Antofagasta, Chile |

# Schriftenreihe des Instituts für Maschinentechnik der Rohstoffindustrie

**ABAR**     **A**achener **B**eiträge zur **A**ngewandten **R**echnertechnik

Band 1     Weidemann J. (1992).
Rechnerunterstützte Maschinenüberwachung als Instrument der vorausschauenden Instandhaltung in Walzwerken am Beispiel von Profilwalzgerüsten; ISBN 3-86073-051-7

Band 2     Cer, H. (1992).
Softwarekonzept zur Analyse und echtzeitfähigen Simulation des dynamischen Verhaltens elektromechanisch gekoppelter Antriebe; ISBN 3-86073-058-4

Band 3     Ruhnau, S. (1992).
Automatisierung des Darstellungsprozesses bei der Variantenkonstruktion von Gewindebohrwerkzeugen unter Verwendung einer graphischen Programmiersprache am Beispiel eines mittelständischen Zulieferunternehmens der Montanindustrie; ISBN 3-86073-067-3

Band 4     Schniering, B. (1993).
Entwiclung und Umsetzung einer praxiorientierten CIM-Strategie als Informationssystem für kleine Unternehmen der Präzisionswerkzeugindustrie am Beispiel eines indirekten Zulieferunternehmens der Montanindustrie; ISBN 3-86073-153-X

Band 5     Streichfuss, M. (1993).
Maschinendiagnose an dieselgetriebenen Transportfahrzeugen im deutschen Steinkohlenbergbau; ISBN 3-86073-154-8

Band 6     Friedhelm, K. (1993).
Konzeption eines Informationssystems für bergmännische und markscheiderische Planung und Dokumentation; ISBN 3-86073-166-3

Band 7        Borstell, D. (1993).
Entwicklung eines Softwarekonzeptes für die rechnergestützte maschinentechnische Detailplanung im deutschen Steinkohlenbergbau; ISBN 3-86073-163-7

Band 8        Linnartz, A. (1993).
Integration von Zustandsdiagnose u. Instandhaltungsplanung, -steuerung und -analyse für dieselgetriebene Transportfahrzeuge im deutschen Steinkohlenbergbau; ISBN 3-86073-290-0

Band 9        Müller, T. (1994).
Zustandüberwachung und Verfügberkeitsprognose in Großanlagen der Stahlindustrie; ISBN 3-86073-291-9

Band 10      Kessler, H.-W. (1994).
Entwicklung und Untersuchung eines Diagnosesystems zur Wälzlagerüberwachung hydrodynamischer Getriebe; ISBN 3-86073-292-7

Band 11      Peschers, H. (1994).
Funktionalität und Aufbau eines Lietsystems zur Verbesserung der Materiallogistik von Steinkohlenbergwerken; ISBN 3-86073-293-5

Band 12      Dummoulin, L. (1995).
Die berührungslose, digitale Messung von Drehmoment und Drehzahl; ISBN 3-86073-293-3

Band 13      Brodurius, T. (1995).
Konzeption eines objektbasierten Programmsystems zur markscheiderischen Bearbeitung und abbauplanung plattenförmiger lagerstätten; ISBN 3-86073-295-1

Band 14      Geropp, B. (1995).
Schwingungsdiagnose an Wälzlagern mit Hilfe der Hüllkurveanalyse; ISBN 3-86073-296-X

Band 15      Plaster, A. (1996).
Weiterentwicklung des wettertechnischen CAE-Arbeitsplatzes für den deutschen Steinkohlenbergbau; ISBN 3-86073-297-8

Band 16      Ahrens, M. (1996).
Entwicklung eines Konzeptes zur Auslegung und zeitoptimalen Berechnung metallischer Zylinderkopfdichtung mit Hilfe der Finite Elemente Methode; ISBN 3-86073-298-6

Band 17    Zhang, M. (1996).
           Messstellenreduktion an elektromechanischen Antriebssträngen mittels prozess-
           geführter digitaler Simulation - ein Beitrag zur Maschinenüberwachung und voraus-
           schauenden Instandhaltung; ISBN 3-86073-299-4

Band 18    Mackel, J. (1996).
           Die anwendungorientierte Elastizitätsberechnung von Walzgerüsten unterschiedlicher
           Bauart; ISBN 3-86073-530-6

Band 19    Seeliger, A. (1996).
           AKIDA - Aachener Kolloquium für Instandhaltung, Diagnose und Anla-
           genüberwachung - Tagungsband des Kolloquiums vom 30.-31. Mai 1996 in Aachen;
           ISBN 3-86073-531-4

Band 20    Standiullo, H.-J. (1996).
           Ein Beitrag zur Entwiklung eines graphikorientierten Informationssystems für die
           maschinentechnische Betriebsmitteleinsatzplanung im deutschen Steinkohlenbergbau;
           ISBN 3-86073-532-2

Band 21    Tao, X. (1997).
           Entwicklung eines Berechnungsmodells zur Ermittlung des Brennraumdruckverlaufs
           aus der Winkelgeschwindigkeit der Kurbelwelle; ISBN 3-86073-533-0

Band 22    Petit, E. (1997).
           Entwicklung eines neuen Verfahrens zur Nassaufbereitung von Bauschutt; ISBN 3-
           86073-534-9

Band 23    Levin, C. (1997).
           Konzeption und Realisierung eines rechnergestützten Planungs- und Informationssys-
           tems für die Projektplanung im deutschen Steinkohlenbergbau; ISBN 3-86073-535-7

Band 24    Kaub, R. (1997).
           Konzeption und Realisierung eines Informationssystems zur Dokumentation und
           Vorhersage der Ausgasunng im deutschen Steinkohlenbergbau; ISBN 3-86073-536-5

Band 25    Schumacher, T. (1998).
           Konzeption und Entwicklung eines integrierten maschinentechnischen Planungs- und
           Informationssystems für Rohstoffgewinnungsbetriebe; ISBN 3-86073-537-3

Band 26    Rensmann, F. (1998).
           Entwicklung einer kleinbauenden Gruben-Diesellokomotive mit optimiertem hydraulis-
           chen Antriebsystem unter Berücksichtigung der wirtschaftlichen Rahmenbedingungen
           im deutschen Steinkohlenbergbau; ISBN 3-86073-538-1


Band 27    Asch, A. (1998).
           Teschnische Diagnose am stoßaufgeladenen Dieselmotor mit Hilfe leistungskorrelierter
           Meßgrößen; ISBN 3-86073-539-X


Band 28    Koch, S. (1998).
           Schwingungsimulation industrieller Antriebe mit Drehfeldmaschinen großer Leistung;
           ISBN 3-86073-700-7


Band 29    Wischnewski, R. (1998).
           Entwicklung eines modelaren, wissensbassierten Systems zur Diagnose von diesel-
           getriebenen Transportfahrzeugen im Deutschen Steinkohlenbergbau;ISBN 3-86073-701-
           5


Band 30    Seeliger, A. (1998).
           AKIDA - 2.   Aachener Kolloquium für Instandhaltung, Diagnose und Anla-
           genüberwachung - Tagungsband des Kolloquiums vom 3.-4 Juni 1998 in Aachen; ISBN
           3-86073-531-4


Band 31    Burgwinkel, P. (1998).
           Das  Transportleitsystem  TLS  -  ein  zentrales  Planungs-,  Steuerung-  und
           Überwachungssystem zur Verbesserung der Materiallogistik von Steinkohlenberg-
           werken; ISBN 3-86073-703-1


Band 32    Lange, C. (1998).
           Die Entwicklung eines Planung- und Informationssystems für die bergmännische Pla-
           nung im deutschen Steinkohlenbergbau; ISBN 3-86073-704-X


Band 33    Tao, X. (1997).
           Konzeption einer Regelung zur automatischen Kohlen injektion in Drehstromlichtbo-
           genöfen; ISBN 3-86073-705-8


Band 34    Georges, D. (1999).
           Konzeption, Entwicklung und Einsatz eines Überwachungssystems zur Erkennung und
           Vermeidung von qualitätsmindernden Störschwingungen an Kaltwalzgerüsten; ISBN
           3-86073-706-6

Band 35    Lorbach, J. (1999).
           Konzeption und Implementirung eines objekt-orientierten Datenmodells zur integri-
           erten Nutzung von Prozessdaten für den Bereich Wettertechnik; ISBN 3-86073-707-4

Band 36    Reimers, J. (1999).
           Konzeption eines Rohrleitungsplanungssystems für den deutschen Steinkohlenbergbau
           ISBN 3-86073-708-2

Band 37    Arefzadeh, S. (1999).
           Klassifizierung der Belastung von Walzanlagen mittels Softcomputing-Methoden als
           Grundlage zur Stichplan- und Restlebensdaueroptimierung; ISBN 3-86073-709-0

Band 38    Bauer, B. (2000).
           Körperschallanalyse von Verzahnungen im Rahmen der zurstandsorientierten Instand-
           haltung; ISBN 3-86073-670-1

Band 39    Baumann, S. (2000).
           Entwicklung eines Diagnosesystems für hydrostatische Antriebe von dieselgetriebenen
           Transportfahrzeugen im deutschen Steinkohlenbergbau; ISBN 3-86073-671-X

Band 40    Mandelartz, J. (2000).
           Der Sekundäraluminium-Stoffstrom in Deutschland unter transporttechnischen und lo-
           gistischen Gesichtspunkten; ISBN 3-86073-672-8

Band 41    Schneider, S. (2000).
           Methode zur Zustandsüberwachung von Kreuzgelenkwellen in den Hauptantrieben von
           walzgerüsten; ISBN 3-86073-673-6

Band 42    Quacken, G. (2000).
           Bestimmung von produktspezifischen, störungsbedingten Qualitätskosten bei der
           Maschinendiagnose; ISBN 3-86073-674-4

Band 43    Türk, M. (2001).
           Optimierung des Walzprozesses und der Anlagenauslastung durch moderne Diagnosev-
           erfahren und vernetzte Überwachungssysteme; ISBN 3-86073-675-2

Band 44    Teschers, R. (2001).
           Konzeption und Entwicklung eines Geamtinformationssystems in einem interdiszi-
           plinären Forschungsprogramm am Beispiel eines Arbeitsbereichs; ISBN 3-86073-676-0

Band 45       Shan, J. (2001).
Zustandsdiagnose von Planetengetrieben bei stationärem und instationärem Betrieb mit Hilfe der Schwingungsanalyse; ISBN 3-86073-677-9

Band 46       Seeliger, A., Burgwinkel, P. (2002).
AKIDA - 4. Aachener Kolloquium für Instandhaltung, Diagnose und Anlagenüberwachung- Tagungsband des Kolloquiums vom 6.-7. November 2002 in Aachen; ISBN 3-86073-677-9

Band 47       Lorbach, J. (1999).
Entwicklung und Erprobung eines neuartigen integrierten Informations-, Kommunikations- und Automatisierungssystems für den Untertagebau; ISBN 3-86073-679-5

Band 48       Nahrath, T. (2003).
Untersuchung zur Erfassung des Säulendurchmessers bei der Düsenstrahlinjektion zur Baugrundfestigung; ISBN 3-89653-981-7

Band 49       Heim, G. K. (2003).
Grenzen und Möglichkeiten einer EDV-Unterstützung gruppenbezogener Planungsarbeit am Beispiel bergbaulicher Produktionsbetriebe; ISBN 3-86130-210-1

Band 50       Vollmer, A. (2003).
Technische und organisatorische Lösungen zur Einführung von Instandhaltungsplanungs- und -steuerungssystemen; ISBN 3-86130-192-X

Band 51       Markhöfer, J. (2003).
Konzeption und Realisierung eines modellgestützten Instrumentariums zur Optimierung der Transportlogistik für den primären und sekundären Kupferstoffstrom; ISBN 3-86130-211-X

Band 52       Hoppe, H. (2003).
Konzeption und Entwicklung eines technischen Berichtsystems zur weltweiten Anlagenüberwachung am Beispiel von Gutbett-Walzenmühlen; ISBN 3-86130-211-X

Band 53       Weyres, S. (2003).
Mobile Computing in der Instandhaltung? Realisierung medienbruchfreier Arbeitabläufe; ISBN 3-86130-212-8

Band 54    Schaaf, C. (2003).
           Entwicklung eines Reportgenerators zur prozessoptimierten Messsignalauswertung an
           Walzanlagen; ISBN 3-86130-213-6


Band 55    Reitz, K. (2004).
           Betrachtung der Körperschallemission von Wälzlagern zur Wälzlagern zur Verfeinerung
           der Zustandsdiagnose; ISBN 3-86130-214-4


Band 56    Gantevoort, T. (2004).
           Integration der Körperschalldiagnose in das kommunikations- und informationstechnis-
           che System des deutschen Untertagebergbaus; ISBN 3-86130-216-0


Band 57    Jansen, S. (2004).
           Modellierung einer Redistributionslogistik metallhaltiger Gebrauchsgüter am Beispiel
           von Elektronikaltgeräten; ISBN 3-86130-216-0


Band 58    Seeliger, A., Burgwinkel, P. (2004).
           AKIDA - 5.   Aachener Kolloquium für Instandhaltung, Diagnose und Anla-
           genüberwachung - Tagungsband des Kolloquiums vom 9.-10.  November 2004; ISBN
           3-86130-217-9


Band 59    Schmacher, G. (2004).
           Entwicklung und Erprobung eines nassmechanischen Bauschuttsortierers; ISBN 3-
           86130-218-7


Band 60    Küpper, T. M. E. (2005).
           Konzeption und Entwicklung eines intranetbasierten Informationssystems für den
           deutschen Steinkohlenbergbau; ISBN 3-86130-219


Band 61    Aguilar C. A. (2005).
           Konzeption und Implementierung einer komponentenbasierten Softwareentwick-
           lungsplattform für raumbezogene internetfähige Informationssysteme; ISBN 3-86130-
           220-9

**ASRE**    **A**achener **S**chriften zur **R**ohstoff- und **E**ntsorgungstechnik

Band 62    Balke, A. (2006).
           Lebenszyklusrechnung von IPS-Systemen in der Auswahl-, Einführungs-und Betrieb-
           sphase; ISBN-10: 3-9810344-2-2; ISBN-13: 978-3-9810344-3-1

Band 63     Seeliger, A., Burgwinkel, P. (2006).
AKIDA - 6 Aachener Kolloquium für Instandhaltung, Diagnose und Anlagenüberwachung - Tagungsband des Kolloquiums vom 14.-15. November 2006 in Aachen; ISBN-10: 3-9810344-3-0; ISBN-13: 978-3-9810344-3-1

Band 64     Charlier, F. (2006).
Neuentwicklung eines schlagwettergeschützten Dieselmotors der Leistungsklasse bis 100kw mit Abgasturbolader; ISBN-10: 3-9810344-4-9; ISBN-13: 978-3-9810344-4-8

Band 65     Bencze, A. (2007).
Entwicklung eines miniaturisierten Drehmomentmesssystems - $\mu$-ETC; ISBN 978-3-9810344-5-5

Band 66     Brümmer, G. (2008).
Konzeption, Entwicklung und Einführung eines grafikbasierten Qualitätssicherungs-System für den mobilen Einsatz; ISBN 978-3-9810344-7-9

Band 67     Steinhusen, C. (2008).
Konzept zur Überwachung elektronischer Komponenten in Windenergieanlagen mit Hilfe konditionierter Simulation und begleitender Messung; ISBN 978-3-9810344-9-3

Band 68     Meßner, A. (2008).
Verfahren zur Schadensdetektion bei langsam oszillierenden Kreuzgelenkwellenlagerungen in Walzwerksantrieben; ISBN 978-3-941277-00-7

Band 69     Lachmann, J. M. (2008).
Entwicklung eines simulationsgestützten Condition-Monitoring-Systems zur Onlineüberwachung des mechanischen Antriebsstranges von Multimegawattwindenergieanlagen; ISBN 978-3-941277-01-4

Band 70     Seeliger, A., Burgwinkel, P. (2008).
AKIDA - 7. Aachener Kolloquium für Instandhaltung für Instandhaltung, Diagnose und Anlagenüberwachung - Tagungsband des Kolloquiums vom 18.-19. November 2008 in Aachen; ISBN 978-3-941277-02-1

Band 71     Buttgereit, D. A. (2009).
Konzeption und Entwicklung eines WLAN-basierten Kommunikations- Informationssystems für den deutschen Steinkohlenbergbau; ISBN 978-3-941277-04-5

Band 72     Molina Vicuna, C. (2010).

Contributions to the Analysis of Vibrations and Acoustic Emissions for the Condition Monitoring of Epicyclic Gearboxes; ISBN 978-3-941277-06-09

Band 73     Nienhaus, K., Burgwinkel, P. (2010).

AKIDA - 8. Aachener Kolloquium für Instandhaltung, Diagnose und Anlagenüberwachung- Tagungsband des Kolloquiums vom 17.-18. November 2010 in Aachen; ISBN 978-3-941277-08-03

Band 74     Neye, E. (2010).

Integration von Informationstechnologie und Grubengebäude bei der Betriebmittelplanung von Steinkohlenbergwerken; ISBN 978-3-941277-09-0

Band 75     Vreydal, D. (2001).

Entwicklung und Einsatz numerischer Simulationsmodelle zur Restlebensdauerabschätzung bei Antriebssträngen von Hubwindsystemen unter Nutzung eines Zweiparametrigen Kollektivierverfahrens; ISBN 978-3-941277-10-6

Band 76     Mavroudis, F. (2011).

Infrarotsensorik zur Grenzschichterkennung - Entwicklung und Einsatz eines bildgebenden Infrarotsensorsystems bei der Automatisierung von Walzenladern im Untertagebergbau; ISBN 978-3-941277-12-0

Band 77     Becker, S. (2011).

Weiterentwicklung und Erprobung eines Kollisionsschutzsystems für Personen und Fahrzeuge im Untertagebau; ISBN 978-3-941277-13-7

Band 78     Gaastra, M. (2012).

Online-Elementanalyse in der Rohstoffgewinnung - Entwicklung und Konzeptionierung von Anwendungen der laserinduzierten Plasmaspektroskopie zur Echtzeitbestimmung mineralischer Rohstoffe am Beispiel unter- und übertägiger Gewinnungsgeräte; ISBN 978-3-941277-14-4

Band 79     Vintila, C. R. (2012).

Entwicklung und Erprobung eines internetfähigen komponentenbasierten Informationssystems zur Echtzeitüberwachung von Schienenfahrzeugen; ISBN 978-3-941277-15-1

Band 80     Vijayakumar, N. (2012).

Computer based Prototype of Heavy Duty Drive Systems by Electromechanical Cosimulation; ISBN: 978-3-941277-16-8

Band 81     Nienhaus K.; Burgwinkel, P.(2012).
            AKIDA - 9.   Aachener Kolloquium für Instandhaltung, Diagnose und Anla-
            genüberwachung - Tagungsband des Kolloquiums vom 14. -15. November 2012; ISBN
            : 978-3-941277-17-5


Band 82     Baltes, R. (2013).
            Entwicklung und Erprobung eines dieselelektrischen Hybridantriebs; ISBN: 978-3-
            941277-18-2


Band 83     Hahn, M. H. (2013).
            Abbildendes Radarsystem für die Rohstoffindustrie; ISBN: 978-3-941277-19-9


Band 84     Nienhaus, K.; Burgwinkel, P. (2014).
            AKIDA - 10.   Aachener Kolloquium für Instandhaltung, Diagnose und Anla-
            genüberwachung - Tagungsband des Kolloquiums vom 19.-20. November 2014 ISBN:
            978-3-941277-21-2


Band 85     Zingsheim, J. M. (2015).
            Inertiale Navigation für die Rohstoffindustrie; ISBN: 978-3-941277-23-6


Band 86     Bitzen, S. (2015).
            Konzeption und Entwicklung eines 3D-Modell-basierten Prozessüberwachungssystems
            für die Rohstoffgewinnung; ISBN: 978-3-941277-24-3


Band 87     Boos, F. D. (2015).
            Acoustic Emission bei der Maschinen- und Prozessüberwachung - Neue Analysemeth-
            oden und Anwendungsgebiete; ISBN: 978-3-941277-25-0