Safety Science 75 (2015) 49–55

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/ssci

# Study of Spanish mining accidents using data mining techniques

Lluís Sanmiquel [a,*], Josep M. Rossell [b], Carla Vintró [c]

[a] Chair "Iberpotash" in sustainable mining, Polytechnic University of Catalonia (UPC), Avenue Bases de Manresa, 61-73, 08242 Manresa, Barcelona, Spain
[b] Department of Applied Mathematics III, Polytechnic University of Catalonia (UPC), Avenue Bases de Manresa, 61-73, 08242 Manresa, Barcelona, Spain
[c] Department of Business Management, Polytechnic University of Catalonia (UPC), Avenue Bases de Manresa, 61-73, 08242 Manresa, Barcelona, Spain

ARTICLE INFO

*Article history:*
Received 19 September 2014
Received in revised form 20 January 2015
Accepted 28 January 2015
Available online 11 February 2015

*Keywords:*
Mining accidents
Data mining
Bayesian network
Classification methods

ABSTRACT

Mining is an economic sector with a high number of accidents. Mines are hazardous places and workers can suffer a wide variety of injuries. Utilizing a database composed of almost 70,000 occupational accidents and fatality reports corresponding to the decade 2003–2012 in the Spanish mining sector, the paper analyzes the main causes of those accidents. To carry out the study, powerful statistical tools have been applied, such as Bayesian classifiers, decision trees or contingency tables, among other data mining techniques. Statistical analyses have been performed using Weka software and behavioral patterns based on certain rules have been obtained. From these rules, some conclusions are extracted which can help to develop suitable prevention policies to reduce injuries and fatalities.



## 1. Introduction

Occupational safety has always been a crucial issue, not only for workers but for companies, workers' unions and institutions responsible of national health at work. Despite new laws being introduced to enhance occupational safety, accidents of different nature still occur, including fatal accidents.

Accidents and occupational diseases produce direct and indirect costs. Indirect costs are more difficult to quantify and they account for a higher proportion than direct costs (Hämäläinen et al., 2006). Occupational accidents produce economic losses in the affected companies, as well as injuries and deaths. It is not a coincidence that those countries with safer work conditions report less occupational accidents, together with better percentages of competitiveness (Organización Internacional del Trabajo, 2003).

Prevention of occupational hazards is required in any job. Undoubtedly, mining can be classified as a dangerous activity (Saxen et al., 1996; Mitchell et al., 1998; Gyekye, 2003; Martín et al., 2009; Sanmiquel et al., 2010), revealing that some features such as environmental conditions with a significant presence of humidity, dust or falling rocks have influence on the number and severity of accidents, higher than in other economic sectors. Several studies have indicated that inadequate management can lead to more workplace accidents. Thus, a study performed on truck-related fatal accidents in surface mining shows that the two most common causes of these accidents are pre-operational improper check and poor maintenance. Furthermore, the non-use of seat belts and an inadequate training were also two important factors (Zhang et al., 2014). Research conducted at a Turkish opencast mine concluded that young workers must be trained and staff with more experience should perform the most critical jobs, due to the number of accidents in the range of 18–30 years (Ural and Demirkol, 2008). Another study on equipment-related fatal accidents in U.S. mining operations establishes that less experienced workers appear to be more vulnerable to equipment-related mishaps and should therefore be at the center of intervention strategies for this category (Kecojevic et al., 2007). An elevated frequency of accidents associated with the non-powered hand-tool use has been observed (Groves et al., 2007), concluding that the companies have significant challenges in controlling these risks and the development of new and creative approaches to train miners still needed.

The relation between the quality of safety management and accident rate was analyzed (Sanmiquel et al., 2014). The study concludes that mining industries of industrial and ornamental stone in Spain (quarries and treatment plants) with high incidence rate of accidents had a worse quality of security management. Similar conclusions were reached in previous studies (Mallick and Mukherjee, 1996; Torres and Dinis, 2003; Zhong-Xue et al., 2008; Bottani et al., 2009). Therefore, it is demonstrated in many industries at risk that safety culture and organizational performance are key factors in a safety climate. Two articles determined that the most immediate cause in the genesis of the accidents was related to inadequate conditions of the workplace and the second one due to the behavior of the employees (Williamson and Feyer, 1998; Sanmiquel et al., 2010).

* Corresponding author. Tel.: +34 938777200; fax: +34 938777202.
*E-mail addresses:* sanmi@emrn.upc.edu (L. Sanmiquel), josep.maria.rossell@upc.edu (J.M. Rossell), carla.vintro@upc.edu (C. Vintró).

http://dx.doi.org/10.1016/j.ssci.2015.01.016
0925-7535/

Generally, the factors associated with the workplace conditions are attributable to the management of the business and the human behavior of workers. However, the causes of accidents due to human behavior may also be imputable to the safety management of companies in cases where the accident is the immediate cause of human error owing to the lack of training and/or experience of the worker. In all these studies it is evident that certain deficiencies (employees with inadequate training, inappropriate maintenance, programs machinery, etc.) are attributable to the management of safety and have a great influence on the accident rate in the mining sector.

In 2012, according to the Ministry of Employment and Social Safety of Spain, the number of accidents per 100,000 workers (*incidence index*) was 4.8 times higher than the total for all economic sectors. If the Spanish incidence index is compared with those of other countries, its value is also much more elevated. Specifically, in 2012 the annual Spanish incidence index was 7.0 times higher than in the United States and 6.5 times higher than in Australia in accordance with the National Institute for Occupational Health and Safety in United States and Safe Work Australia, respectively. Extensive research has been developed in this field to discover the main causes of accidents.

Annual incidence index of the Spanish mining sector, considering the number of occupational accidents, is analyzed in the paper. To do this, statistical techniques of Bayesian networks with Weka, well-known and powerful software for statistics that enables the user to preselect the most representative variables for the genesis of accidents, as well as to obtain behavioral patterns based on classification and association rules and decision trees, are used.

Bayesian networks have been applied in several scientific fields, such as medicine (Antal et al., 2004); ecology (Adriaenssens et al., 2004); environmental assessment impact (Baran and Jantunen, 2004; Marcot et al., 2001; Matías et al., 2006); business risk and product life-cycle analysis (Zhu and Deshmukh, 2003); geological engineering (Rivas et al., 2007); road traffic safety (Flask and Schneider, 2013); workplace risk area (Galán et al., 2007); workplace tasks (Martín et al., 2009); construction and mining accidents (Matías et al., 2008; Rivas et al., 2011); interrelation between hygienic workplace conditions and occupational accidents (García-Herrero et al., 2012), among others. Analysis of simulated data indicates that, in general, methods based on Bayesian networks are more sensitive in detecting associations among categorical variables than other statistics methods. Reliable conclusions for the decision-making process in safety issues are extracted.

Data Mining Techniques is very important for mining accidents analysis and statistics. However, also it is very important can carry out an accurate data acquisition for mining accidents analysis and their prevention. In this way, there are different technologies very useful for it, such as Remote Sensing and Global Positioning System. The remote sensing can play an important role for regional-scale analysis and effective management of environmental impacts induced by surface mining activities (Demirel et al., 2011). Also, this technology can be useful for an adequate mine planning that can help in the prevention of some type of accidents. The Global Positioning System has been used in many surfaces mining activities, mainly in the prevention of the collision accidents (Wu et al., 2013).

## 2. Material and methodology

### 2.1. Objectives

The goal of this work is to analyze the main factors that could be causes of Spanish mining accidents with the purpose to develop future prevention policies minimizing the rates of occupational hazards. The study is conducted in four parts:

(1) A selection of mining data is made, taking into account its relevance in our study, in order to work with a reasonable data source. The variable *Type of Accident*, chosen a priori, serves as output attribute.
(2) By means of powerful attribute evaluators and data mining tools provided by Weka software version 3-7-8, a ranking of the most important predictors for the output variable *Type of Accident* is established.
(3) Association rules from decision trees for the predictor variables are obtained. These association rules allow the most relevant factors to be extracted to enable the output variable to be evaluated.
(4) After that, another variable named *Lost Work Days* serves to measure the severity of the different kinds of accidents given by *Type of Accident*.

### 2.2. Study population

The study population is constituted by 69,869 instances of occupational accidents recorded in the Spanish mining sector for the period 2003–2012. Data was obtained from the annual digital database on accidents of the Spanish Ministry of Employment and Social Safety. The accidents considered in the study are those that occurred in the mining workplace, within regular working hours and causing to the injured employees the loss of, at least, one working day. Accidents called "in itinere" were not taken into account.

### 2.3. Variables

Although 58 variables were collected, only 15 of them were selected by using criteria such as our previous experience and other results published on this topic. Two variables were considered as response and the remaining thirteen variables as predictors. To facilitate the further study, the variables were grouped into numerical classes or categories.

The 13 predictor variables are the following:

**Age (A)**: Age of the injured worker when the accident happened. Seven classes are considered: [16,24], [25,29], [30,34], [35,39], [40,44], [45,54], [55,→].

**Experience (E)**: Experience in this kind of work (in months) of the injured employee. The following seven groups: [0,12], [13,30], [31,60], [61,120], [121,180], [181,240], [241,→] have been created.

**Size (S)**: Number of employees of the company. The groups are: [0,9], [10,19], [20,49], [50,99], [100,499], [500,→].

**Contract (C)**: Type of employment contract. Thus, 1 = permanent and full-time; 2 = permanent and part-time; 3 = temporary and full-time; 4 = temporary and part-time.

**Previous Causes (PC)**: Existence of a previous cause before the accident. The causes have been grouped into seven categories: 1 = electric problem, explosion, fire, overflow, overturn, leak, spill, vaporization or emanation; 2 = fracture, slip, fall or collapse; 3 = loss of control (total or partial) of the working machinery; 4 = falls/tumbles of a person; 5 = body movement without physical effort; 6 = body movement with physical effort; 7 = other causes.

**Place (P)**: Place where the injury occurred. Five groups are considered: 1 = repair area, storage area, loading and unloading area; 2 = general constructions or demolitions; 3 = opencast mine or quarry; 4 = underground mine; 5 = other places. Moreover, the variable *Place* determines if the accident occurred inside or outside of the mine.

**Physical Activity (PA)**: Type of physical activity that the worker was making at the time of the accident. It has been grouped into seven categories: 1 = machine operations; 2 = working with hand tools; 3 = driving or being in a conveyance; 4 = manipulation of