Bayesian Decision Tool for the Analysis of Occupational Accidents in the Construction of Embankments

S. Gerassis[1]; J. E. Martín[2]; J. Taboada García[3]; A. Saavedra[4]; and J. Taboada[5]

Abstract: Instability and poor construction practices are responsible for the high accident rate in embankment construction in Spain. Applying a methodology based on data mining and attribute selection and using a 6-year database of accidents, key attributes in accidents associated with the construction of embankments were analyzed. Once the main predictors were identified, Bayesian networks in order to quantify the specific causes of different types of accidents were built. Thus, the main reasons for accidents as a preliminary phase to enhancing safety and embankment stability in mining and civil engineering works can be accurately identified and quantified. **DOI: 10.1061/(ASCE)CO.1943-7862.0001225.** *© 2016 American Society of Civil Engineers.*

Author keywords: Data mining; Attribute selection; Bayesian network; Mine and civil engineering safety; Embankment accidents; Machine learning; Labor and personnel issues.

Introduction

Occupational safety is an issue of vital importance, most particularly in key privatized sectors such as mining and civil engineering works, where the risk of productivity and financial losses arising from poor health and safety policies can have major repercussions on increasingly narrow profit margins (Maffini et al. 2014).

In terms of both human and financial considerations, there is already evidence of safety as a core issue in planning (Chen 2013). Mining is a high-risk occupation where most companies implement health and safety management more to comply with legal obligations than as a beneficial aspect of their business culture. Despite scientific and technological innovations aimed at ameliorating this situation, recent studies and statistical data point to a disproportionately and unacceptably high rate of accidents and injuries in mining and construction compared to other sectors (Coleman and Kerkering 2007; Sanmiquel et al. 2010; Fonseca et al. 2014).

Research in recent years suggests that countries ensuring safe working conditions are also more competitive economies (International Labour Organization 2003; Mainardi 2005), a circumstance associated with the political and financial resources allocated to health and safety. In the mining sector, it has been found that an increase in spending on health and safety has a direct impact on economic growth, with a long-term contribution that can be as high as 7% (Tan et al. 2012).

The construction of embankments is a typical and complex feature of both mining and civil engineering works. Scientific progress in embankment engineering has unquestionably been made in recent years, notably in the development of complex numerical methods to analyze stability (Li 2007), statistical and empirical procedures for assessing potential landslide risk (Alejano et al. 2008), and sophisticated control systems for in situ monitoring (Lienhart 2015).

Nonetheless, embankment-related accidents continue to occur, both during and after construction, mainly as a result of landslides and collapse due to instabilities, to the point that there is little optimism about how to resolve this problem (Swuste et al. 2012). Many different and increasingly complex strategies have been adopted without obtaining the desired results, mainly because many approaches to reducing accidents are too broad in scope, given that they apply to entire sectors, when what is required is a more accurate and realistic analysis of each particular context. This article describes this kind of more focused analysis, implemented for the specific operational context of embankment construction.

The authors describe an approach based on first mining a database of embankment-related accidents in mining and civil engineering works, and then constructing a Bayesian network of the most influential factors so as to identify and quantify specific accident causes.

Data mining is increasingly being used to analyze accidents in the mining industry (Liai and Chou 2012; Sanmiquel et al. 2015) and Bayesian networks have been applied in many disciplines, including the analysis of accidents in general (Matías et al. 2008) and more specific accidents, such as those resulting from falls from a height (Martín et al. 2009), tunnel construction (Zhang et al. 2014; Wu et al. 2015), or fire risk in construction (Wang et al. 2010) among others.

The powerful combination of data mining and Bayesian networks can accurately pinpoint the specific causes of accidents for subsequent remedial measures aimed at enhancing workplace safety (Leu and Chang 2013). The goal is to enhance the quality of safety planning by identifying the impact of specific circumstances on different types of accidents. Ultimately, information is obtained that enables companies to better understand where and how to invest in measures to ameliorate the serious problem of occupational accidents. This should change perceptions of safety,

[1]Ph.D. Student, Dept. of Natural Resources and Environmental Engineering, Univ. of Vigo, Lagoas Marcosende, 36310 Vigo, Spain.

[2]Adjunct Professor, Dept. of Natural Resources and Environmental Engineering, Univ. of Vigo, Lagoas Marcosende, 36310 Vigo, Spain.

[3]Ph.D. Student, Dept. of Natural Resources and Environmental Engineering, Univ. of Vigo, Lagoas Marcosende, 36310 Vigo, Spain.

[4]Associate Professor, Dept. of Statistics and Operational Research, Univ. of Vigo, Lagoas Marcosende, 36310 Vigo, Spain (corresponding author). E-mail: saavedra@uvigo.es

[5]Full Professor, Dept. of Natural Resources and Environmental Engineering, Univ. of Vigo, Lagoas Marcosende, 36310 Vigo, Spain.

Note. This manuscript was submitted on March 2, 2016; approved on July 8, 2016; published online on August 16, 2016. Discussion period open until January 16, 2017; separate discussions must be submitted for individual papers. This paper is part of the ***Journal of Construction Engineering and Management***, © ASCE, ISSN 0733-9364.

which needs to be assumed as an intrinsic and ultimately useful concept in mining and civil engineering works (Hutcheon 1986).

Materials and Methods

Model Formulation

The main causes of accidents in the construction of embankments were identified and quantified in two main steps:

1. Data mining, aimed at identifying the attributes that impinge on accident occurrence: *Weka* version 3-6-12 data-mining software (Hall et al. 2009), developed at the University of Waikato in New Zealand, was used for attribute selection. This software is a collection of machine learning algorithms, called classifiers, that can be implemented directly in a dataset (Witten et al. 2011). In this case, for example, from an effect attribute, namely, the type of accident, it ranked the most influential attributes predicting the accident. This ranking of attributes allowed a simplified database to be created that was computationally and conceptually accurate in quantifying the probability of an accident; and
2. Bayesian network mapping of the previously identified attributes that most affect accident occurrence: The synthesized network analysed the relationship between cause nodes—reflecting the most influential attributes—converging in a single effect node, namely, the type of accident. This network was constructed using *GeNIe & SMILE* software version 2.0 (Druzdzel 1999), with input from health and safety experts in order to establish topology on the basis of attributes affecting the embankment construction context. The intervention of experts aimed for Bayesian inference to emulate reality as closely as possible in quantifying different types of accidents on the basis of their most likely causes.

Statistical Modeling and Explanatory Power Analysis

Data mining aims to uncover patterns of information in large volumes of data (Fayyad et al. 1996). In recent years it has revolutionized many scientific fields and will undoubtedly acquire even more importance in the future in a world where data volumes are growing apace. Data mining is a particular step in Knowledge Discovery in Databases (KDD), encompassing a broader process of knowledge acquisition from raw data.

Data mining in this case was applied to a database of embankment construction accidents so as to better understand why and how accidents occur. Machine learning as implemented through *Weka* and the statistical procedure known as attribute selection furnished the technical basis for the data-mined information. From the documented accidents, the predictive power of sets of attributes in classifying cases by accident type was analyzed and a smaller subset of attributes was then identified such that predictive capacity was not reduced significantly.

Weka selects attributes in a supervised process based on two algorithmic components:

1. Search Method searches for subsets of attributes. Since evaluating all possible subsets is a computationally intractable problem, algorithms offer different strategies to optimize the search; and
2. Attribute Evaluator determines the subset of attributes for classification purposes and assigns a specific weight to each attribute that reflects its quality.

The *Weka* evaluators (Table 1) evaluate just a single attribute corresponding to the prediction aim, in this case, the type of accident. Searches are conducted by the Ranker algorithm, which provides a weighted ranking of all attributes without exception according to their predictive power.

Tenfold cross-validation was used to run the process and choose the attributes. This involved randomly reordering the dataset and dividing it into sets (folds) of equal size. One fold was used to test the classifier and the remaining folds $(n-1)$ were used to train the classifier.

Once attributes were selected and the attributes that most influenced accidents were identified, a Bayesian network was built to quantify the probabilities of each type of accident occurring.

Bayesian networks are directed acyclic graphs where nodes and arcs symbolize direct dependency relationships between attributes (Charniak 1991). The topological structure of the Bayesian model reflects the dependency relationships between attributes and describes the distribution of probabilities for a particular event occurring in specific conditions. If $X = \{X_1, X_2, \ldots, X_n\}$ is a set of attributes, then a Bayesian network is formally defined as a couplet $X = \langle G, P \rangle$ where

- G is a directed acyclic graph in which each node represents one of the attributes $X_1, X_2, \ldots, X_n$ and each arc represents a direct dependency relationship between attributes. The direction of the arc indicates that a target attribute is dependent on the origin attribute; and
- P is a set of parameters that quanties the network, containing the probabilities for each possible value x_i for each attribute X_i;

Thus, from the decomposition theorem the joint probability P under the hypothesis that each node is independent of its nondescendants can be calculated. The Bayesian network therefore has a single joint probability distribution given by

$$P(X) = P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P[X_i | X_{j(i)}]$$

Table 1. *Weka* Attribute Selection

Attribute evaluator	Description	Search method	Attribute target
ChiSquaredAttibuteEval	Evaluates the worth of an attribute by computing the value of the chi-squared statistic with respect to the class	Ranker	Type of accident
GainRatioAttributeEval	Evaluates the worth of an attribute by measuring the gain ration with respect to the class	Ranker	Type of accident
InfoGainRatioAttributeEval	Evaluates the worth of an attribute by measuring the information gain with respect to the class	Ranker	Type of accident
OneRAttributeEval	Uses the minimum error attribute to predict the quality of each attribute with respect to the class	Ranker	Type of accident
ReliefAttributeEval	Evaluates the worth of an attribute by repeatedly sampling an instance and considering the value of the given attribute	Ranker	Type of accident
SymettricalUncertAttributEval	Measures the symmetrical uncertainty to evaluate the quality of each attribute with respect to the class	Ranker	Type of accident

where $X_{j(i)}$ = set of parent variables of X_i for directed acyclic graph G. Accordingly, application of Bayes' theorem enables to determine the a posteriori probability of the attribute of interest—in this case, the type of accident—from the a priori data obtained for the remaining attributes.

Bayesian network training was divided into two parts, both implemented with *GeNIe & SMILE* software for the construction of theoretical graphical decision models:

1. Structural learning, that is, the creation of the topological structure of the network. In this research, health and safety experts determined the network topology and decided the direction of the arcs (cause–effect). The goal was to obtain a list of attributes that reflected reality—the operational context of embankment construction—as closely as possible. Likewise, the use of specific learning algorithms that might not accurately reflect the nature of the phenomenon was avoided; and
2. Parametric learning. Once the topology of the graph was known, the probabilities corresponding to each node could be obtained through Bayesian inference.

Bayesian networks can be used for reasoning in either direction. Once trained, the network was used—through predictive reasoning—to draw inferences regarding accident type from cause to effect nodes in the direction of the arcs. Intercausal reasoning was also applied in order to analyze how cause nodes were influenced from the same effect node.

Database Description

This study was based on a total of 353 mining and civil engineering accidents that occurred between 2009 and 2014 in companies operating under the Spanish legislative framework. The core database was created from accident reports supplied by the companies. However, in order to ensure a more robust model, further data were collected via questionnaires and personal and telephone interviews with relevant company employees. Contact was also made with the prevention technicians responsible for accident investigation who, along with the workers, provided relevant information regarding the attributes that needed to be defined in relation to embankment construction accidents.

This research also took into account accidents during maintenance operations for embankments built in the past that, with the passage of time, are more prone to instability and so need reinforcement. This activity is especially risky, because workers need to operate at heights and in conditions of possible instability.

Definition of Attributes

The information obtained through questionnaires/interviews and from prevention experts was used to define attributes reflecting the full range of the causes of embankment construction accidents. It was necessary, for instance, to define attributes reflecting worker experience and training as common causes of accidents (Bird et al. 2004) and also to include specific attributes related to the embankment construction technique, such as the height, slope, and materials used. All the attributes considered in this study are listed subsequently. Only those identified by data mining as having greater predictive power were considered for the construction of the Bayesian network. Each attribute had two or three possible responses that reflected real scenarios, reflected in turn in the Bayesian network as different states. The considered attributes are

1. Shift work
   - S1. No
   - S2. Yes
2. Shift duration
   - S1. 8 or fewer hours
   - S2. More than 8 h
3. Machinery age
   - S1. Under 10 years
   - S2. More than 10 years
4. Machinery maintenance
   - S1. Satisfactory
   - S2. Unsatisfactory
5. State of loading and unloading zones and haul roads
   - S1. Satisfactory
   - S2. Unsatisfactory
6. Embankment geology
   - S1. Rock
   - S2. Soil
   - S3. Mixed
7. Embankment slope
   - S1. Below 45°C
   - S2. Subvertical to 45°
   - S3. Vertical
8. Embankment height
   - S1. Under 10 m
   - S2. 10 to 40 m
   - S3. Over 40 m
9. Anchorage system drilling
   - S1. Drilling carriage
   - S2. Manual
10. Outsourcing in the same loading and unloading zone
    - S1. No
    - S2. Yes
11. Signaling and signposting
    - S1. Satisfactory
    - S2. Unsatisfactory
12. Temperature
    - S1. Over 30°C
    - S2. 0 to 30°C
    - S3. Below 0°C
13. Order and cleanliness
    - S1. Satisfactory
    - S2. Unsatisfactory
14. Works completion schedule
    - S1. On time
    - S2. Delayed
15. Type of operation
    - S1. Mesh installation
    - S2. Cabling, bolting, and/or shotcreting
    - S3. Others
16. Time of accident
    - S1. Morning
    - S2. Afternoon
    - S3. Overtime
17. Day of accident
    - S1. Monday
    - S2. Tuesday to Thursday
    - S3. Friday or Saturday
18. Seniority of operator
    - S1. Longer than 1 year
    - S2. 3 months to 1 year
    - S3. Under 3 months
19. Operator training
    - S1. Training before
    - S2. Training during
    - S3. No training

20. Injured operator age
    S1. 18 to 30 years old
    S2. 31 to 55 years old
    S3. Over 55 years old
21. Injured worker's usual job?
    S1. Yes
    S2. No
22. Injured operator nationality
    S1. Spanish
    S2. Other
23. Employment contract type
    S1. Indefinite
    S2. Temporary

Types of Accident

As a further attribute, in addition to the 23 above, a predictor whose states were different types of accidents associated with embankment construction was included. These were as follows:

- Falls from the same/different height,
- Detachment, collapse or handling-induced falls of loose objects,
- Stepping on objects,
- Collisions with mobile or stationary objects,
- Blows from objects or equipment,
- Projection of fragments and particles,
- Entrapment by/between objects or upturned machinery,
- Overexertion,
- Thermal contact,
- Exposures to electrical contacts,
- Exposure to caustic and corrosive substances,
- Fires, and
- Accidents involving vehicles.

Results and Discussion

The most influential attributes in embankment construction accidents were identified and ordered using data-mining techniques and the most likely accidents were quantified and analyzed using a Bayesian network model that pinpointed the influence of specific attributes in causing accidents. The results obtained are described in the next subsections.

Data-Mining Results

Data mining using *Weka*'s attribute selection option established a classification of the most influential attributes causing accidents. Table 2 shows the classification obtained by the InfoGainAttributeEval attribute evaluator and the Ranker search method, along with average merit and average rank for each attribute and the corresponding standard deviations. Average merit reflects correlations measured by InfoGainAttributeEval in the 10 cross-validation cycles, whereas average rank indicates the average order of each attribute in each of these cycles. Standard deviations of merit and rank are given at the right of each average.

As can be observed, V12 (temperature), V15 (type of operation), V7 (embankment slope), and V2 (shift duration) were, in that order, the four attributes most affecting the occurrence of accidents, whereas V3 (machinery age), V4 (machinery maintenance), and V22 (injured operator nationality) were the three attributes that least affected the occurrence of accidents.

The results were similar for all search methods implemented, with unanimity regarding the least-predictive attributes up to V23 (employment contract type). From V13 (order and cleanliness), the different methods slightly modified the ranking of some of the attributes, with InfoGainAttributeEval representing the general trend in ranking.

Table 2. Classification Obtained by the Info GainAttributeEval Attribute Evaluator and the Ranker Search Method for Establishing the Most Influential Attributes Causing Accidents

Average merit	Average rank	Attribute
0.067 ± 0.007	1 ± 0	Temperature
0.054 ± 0.006	3 ± 1.26	Type of operation
0.051 ± 0.006	3.8 ± 1.83	Embankment slope
0.049 ± 0.006	4.6 ± 1.2	Shift duration
0.044 ± 0.006	5.7 ± 2.19	Time of accident
0.043 ± 0.005	5.9 ± 1.76	Injured operator age
0.043 ± 0.005	6 ± 1.95	Day of accident
0.039 ± 0.005	7.4 ± 1.8	Embankment height
0.035 ± 0.006	9.1 ± 1.51	Embankment geology
0.034 ± 0.005	9.4 ± 2.24	Seniority of operator
0.032 ± 0.003	10.5 ± 1.28	Operator training
0.025 ± 0.004	13.5 ± 2.11	Anchorage system drilling
0.024 ± 0.003	13.7 ± 1.95	State of loading and unloading zones and haul roads
0.023 ± 0.003	14.7 ± 1.49	Order and cleanliness
0.022 ± 0.004	14.8 ± 1.83	Employment contract type
0.02 ± 0.004	16.1 ± 2.59	Works completion schedule
0.018 ± 0.005	17.2 ± 2.56	Outsourcing in the same loading and unloading zone
0.017 ± 0.003	17.8 ± 2.14	Injured worker's usual job?
0.015 ± 0.003	18.7 ± 2.19	Signaling and signposting
0.015 ± 0.003	19.1 ± 2.47	Shift work
0.013 ± 0.002	20.8 ± 1.33	Injured operator nationality
0.012 ± 0.002	21.5 ± 1.2	Machinery maintenance
0.011 ± 0.004	21.7 ± 2.37	Machinery age

Bayesian Model Results: Simplification and Interpretation

Attribute selection from the most influential attributes in the expert-built network pointed to the attributes on which efforts to reduce accidents should be focused.

This selection capability is useful as a support tool in preliminary safety studies conducted in the planning stage of embankment construction, when a large number of attributes need to be considered. The attributes that statistically have the greatest influence on safety can be used to create a simple Bayesian network. This simplification does not invalidate a subsequent, more detailed preventive study that analyzes all attributes selected for the embankment construction scenario.

An application example is described next based on the construction of a Bayesian network for the entire set of attributes (Fig. 1). This general network graphically quantified accident risk in terms of all possible recorded attributes.

Although this network yielded valuable information from the preventive and technical points of view, the true potential of the model was reflected in a simplified network (Fig. 2) that zeroed in on the most influential accident risk predictors.

In choosing the number of attributes to use, V13 (order and cleanliness) was chosen as the cutoff point from which there were slight changes in ranking. Moreover, order and cleanliness has been documented as a key issue that negatively affects productivity if deficient (Chinchilla 2002; Sacristan 2005).

Attributes immediately below V13 (order and cleanliness) were ranked the same by all the search methods and could safely be excluded since there was little doubt as to their poor predictive capacity.

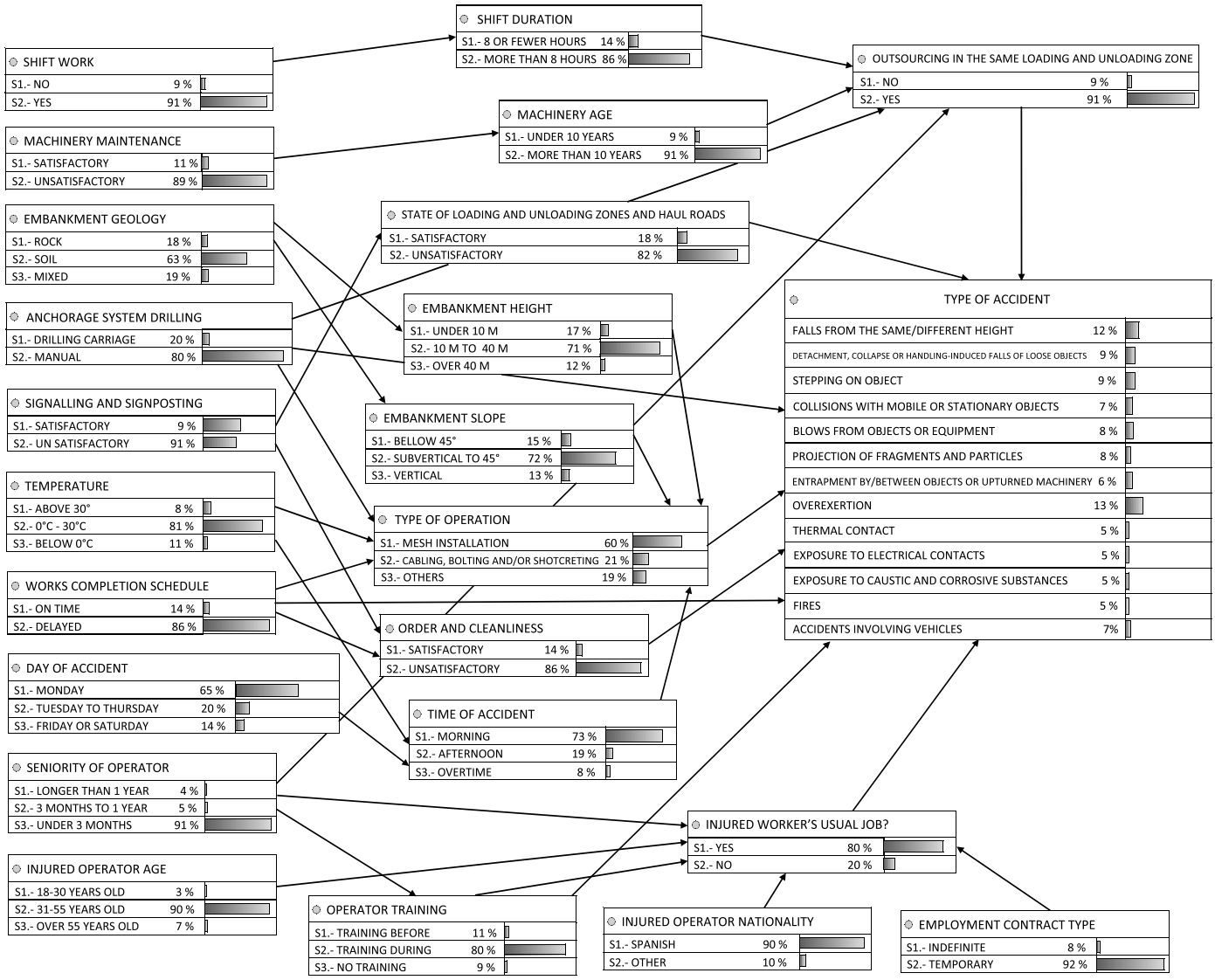

Fig. 1. Bayesian network for the analysis of embankment construction accidents

The 23 attributes were reduced to 14 cause nodes converging in a common effect node, type of accident. Probabilities and predictive capacity of the simplified network were virtually identical to the original network.

From the simplified Bayesian network, the following results (also observable in the original network) were obtained.

The most likely type of accident was overexertion (13%), followed by falls from the same/different height (12%), accidents caused by detachment, collapse or handling-induced falls of loose objects (9%), and stepping on object (9%). The least-likely accidents were those resulting from exposure to thermal or electrical contacts, contact with caustic or corrosive substances, and fire.

Other interesting results were as follows:

- A total of 60% of accidents occurred during mesh installation and 21% during cabling, bolting, and shotcreting. That leads to 81% of accidents occurring during reinforcement operations. These accident rates are explained by a high degree of manual drilling (80%), also related, in turn, to the most likely type of accident, namely, overexertion (13%);
- Seniority in the job was under 3 months in 91% of operators associated with accidents; 90% of these operators were aged 31–55 years and only 11% had been trained before commencement of the activity. The impact of lack of experience coupled with the lack of training on risky behaviour is corroborated as a common cause of accidents by other studies (Paul and Maiti 2007);
- Another key attribute in accidents was general order and cleanliness, which was deficient in 86% of cases. Loading and unloading zones and haul roads also presented shortcomings in 82% of cases and the temperature typically associated with accidents was between 0 and 30°C. Monday was the most accident-prone day of the week (65%) and 73% of accidents occurred in the morning. Furthermore, long working days, in 86% of cases more than 8 h, evidently led to fatigue and significantly added to accident risk; and
- Regarding technical aspects, the geological material most likely to be associated with an accident was soil (63%)—far more than rock (18%). Embankment slope was another major attribute, responsible for 72% of accidents when the angle was subvertical to 45°. As for embankment height, 71% of accidents occurred at between 10 and 40 m compared to only 17% below 10 m.

Besides the aforementioned results, Bayesian networks can be used to exploit the conceptual field by applying intercausal reasoning

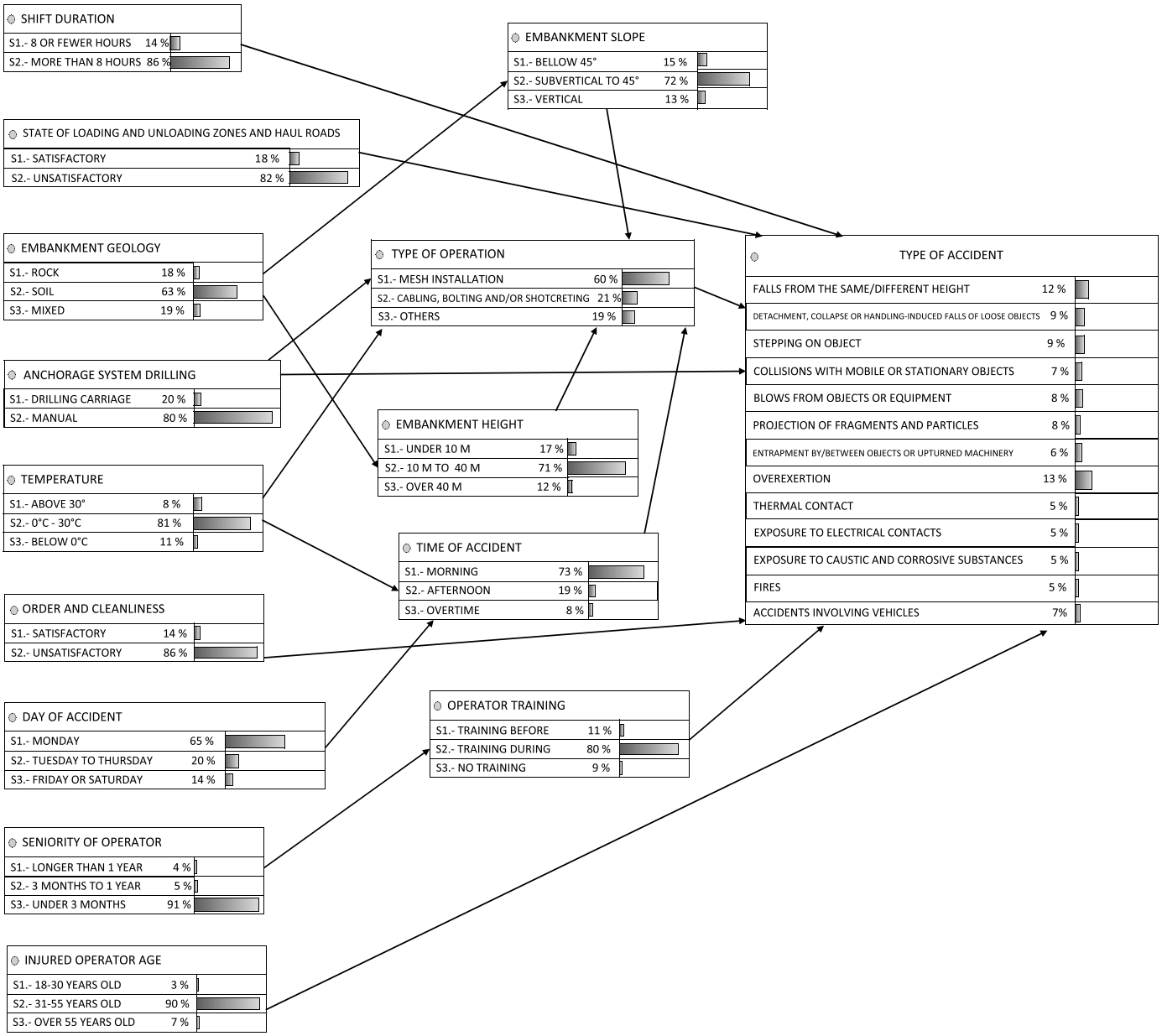

Fig. 2. Simplified Bayesian network for the analysis of embankment construction accidents

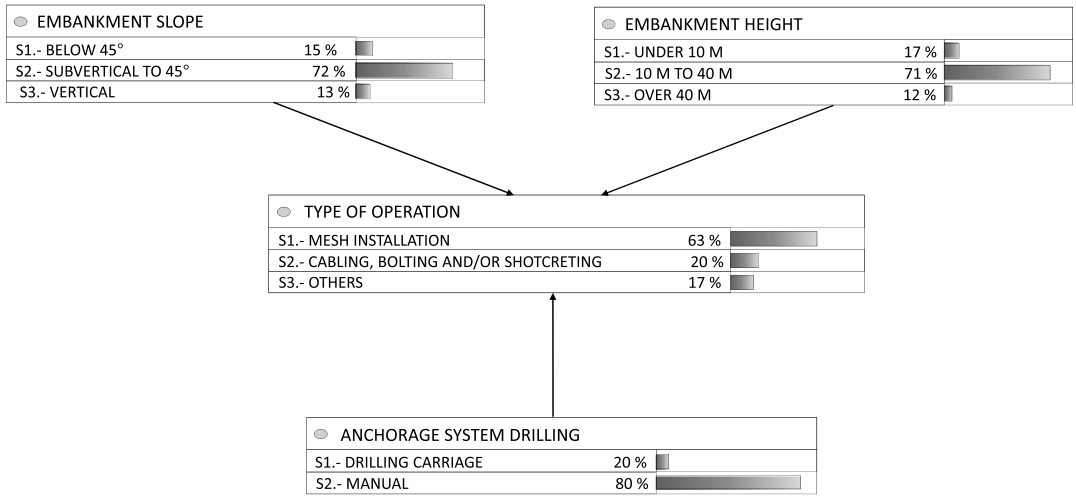

Fig. 3. Analysis of factors predictive of an embankment construction accident: typical scenario

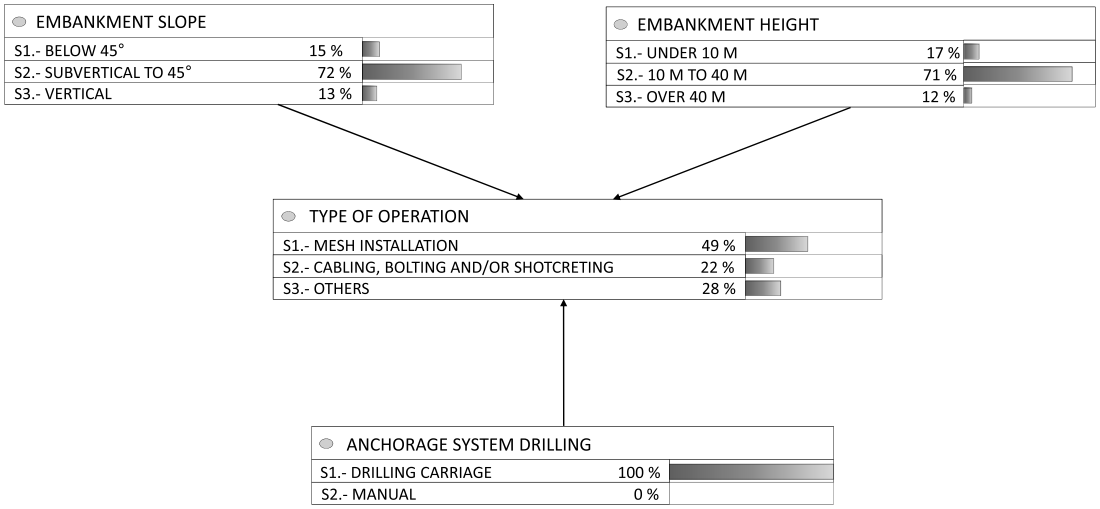

Fig. 4. Analysis of factors predictive of a specific embankment construction accident: anchorage system implemented mechanically using a drilling carriage

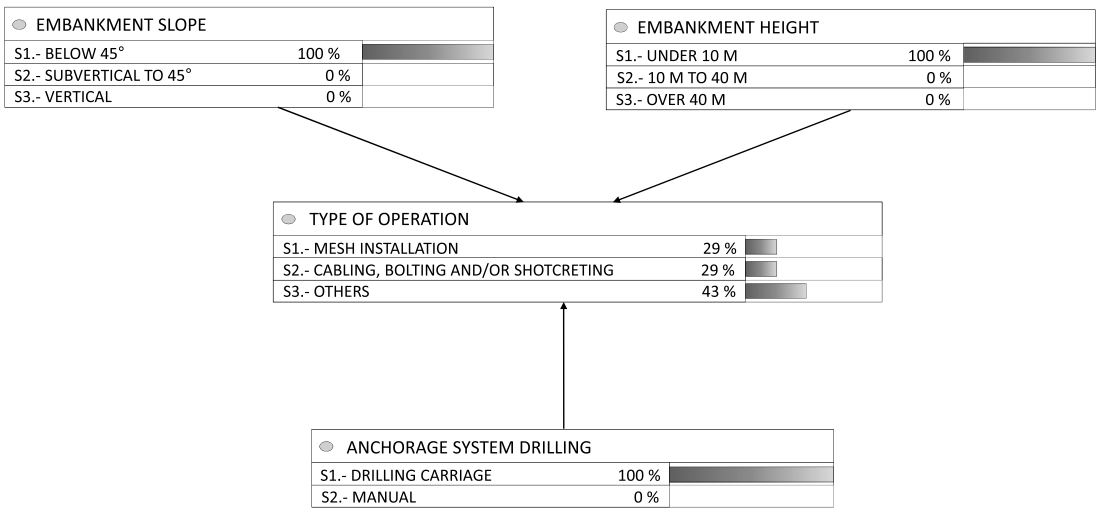

Fig. 5. Analysis of factors predictive of an embankment construction accident: conservative scenario

so as to determine the probability relationship between casual attributes producing particular types of accidents.

The simplified network (Fig. 3) highlighted a particularly striking result, namely, that 83% of accidents occurred during embankment reinforcement operations, that is to say: mesh installation and cabling, bolting, and/or shotcreting. It was found that, in this very typical operation, anchorage system drilling was manual in 80% of the cases, embankment height exceeded 10 m in 83% of cases and slope was typically between subvertical and 45°.

Mechanical drilling was simulated while holding other attributes constant (Fig. 4) in order to see whether this would improve outcomes. The inference result showed that reinforcement accidents (mesh installation and cabling, bolting, and/or shotcreting) were reduced to 71%, with those associated with mesh installation dropping from 63 to 49%, in other words dropping by 14 percentage points. For a correct interpretation of this result, it is important to keep in mind that the network only simulates cases where an accident occurred. Therefore, this indicates that mesh installation is a more dangerous operation when executed manually than by drilling carriage.

Finally, simulating a highly conservative construction scenario—embankment height under 10 m and embankment slope less than 45°—led to accidents associated with mesh installation and with cabling, bolting, and shotcreting dropping significantly to 29% each (Fig. 5). Thus, a value of 58% can be obtained for embankment reinforcement operations by adding those latest percentages.

This practical illustration using attributes in the built network demonstrates the potential of data mining combined with Bayesian networks.

Conclusions

The method described represents a practical and easily implemented tool for identifying and quantifying the causes of accidents associated with the construction of embankments in mining and civil engineering works.

The simplified version of the Bayesian network, demonstrated to be capable of quantifying different types of accidents and their most influential predictive attributes, can assist in optimizing

spending at the design and preliminary analysis stages of embankment construction.

Both the original and simplified networks can assist in devising health and safety preventive strategies that are adapted to existing material and human resources in the company and that are aligned with the weight of each attribute. The inferential capacity can also highlight two-way links between causal attributes. Going one step forward, they can be used to analyze the influence of various factors on the risk of a certain type of accident and how factors are interrelated. Unlike traditional methods of analysis, Bayesian networks also enable a spatial vision of the heterogeneity of the embankment-related accident risk scenario.

These results would suggest that this type of network can be reliably used as an analytical tool with potential for applications in other areas. Future applications could include evaluating changing environments in the construction or mining sectors, where it is difficult to assess work posts and conditions using traditional techniques. This kind of networks can also be used to predict the probability of different types of accident and can also be used to perform sensitivity studies, that is, to quantify the impact that a small change in one factor produces in other factors.

However, the ultimate goal of this research was to bring to light the different attributes responsible for accidents that tend to go unnoticed by experts. This study contributes to knowledge of earthmoving accidents and so enhances occupational health and safety from the perspective of the productivity-safety binomial.

Acknowledgments

The authors acknowledge the funding received by the Xunta de Galicia and FEDER funds through the GRC2014/020 project.

References

Alejano, L. R., Stockhausen, H. W., Alonso, E., Bastante, F. G., and Ramírez, P. (2008). "ROFRAQ: A statistics-based empirical method for assessing accident risk from rockfalls in quarries." *Int. J. Rock Mech. Min. Sci.*, 45(8), 1252–1272.

Bird, F. E., Germain, G. L., and Clark, M. D. (2004). *Practical loss control leadership*, International Loss Control Institute, Loganville, GA.

Charniak, E. (1991). "Bayesian networks without tears." *AI Mag.*, 12(4), 50–63.

Chen, J. K. C., and Zorigt, D. (2013). "Managing occupational health and safety in the mining industry." *J. Bus. Res.*, 66(11), 2321–2331.

Chinchilla, R. (2002). *Salud y seguridad en el trabajo*, Editoral Universidad Estatal a Distancia, Madrid, Spain (in Spanish).

Coleman, P., and Kerkering, J. (2007). "Measuring mining safety with injury statistics: Lost workdays as indicators of risk." *J. Saf. Res.*, 38(5), 523–533.

Druzdzel, M. J. (1999). "SMILE: Structural Modeling, Inference, and Learning Engine and GeNIe: A development environment for graphical decision-theoretic models (intelligent systems demonstration)." *Proc., 16th National Conf. on Artificial Intelligence (AAAI-99)*, AAAI Press/MIT Press, Menlo Park, CA.

Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). "From data mining to knowledge discovery in databases." *AI Mag.*, 17(3), 37–54.

Fonseca, A., McAllister, M. L., and Fitzpatrick, P. (2014). "Sustainability reporting among mining corporations: A constructive critique of the GRI approach." *J. Cleaner Prod.*, 84(1), 70–83.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. A. (2009). "The WEKA data mining software: An update." *SIGKDD Explor.*, 11(1), 10–18.

Hutcheon, I. C. (1986). "Intrinsic safety rules OK. Control science and technology for development." *Proc., IFAC/IFORS Symp.*, Y. Jiachi, ed., Beijing, People's Republic of China, 135–137.

International Labour Organisation. (2003). "La seguridad en cifras." ⟨https://www.ilo.org/legacy/english/protection/safework/worldday/report_esp.pdf⟩ (Jul. 8, 2016).

Leu, S.-S., and Chang, C.-M. (2013). "Bayesian-network-based safety risk assessment for steel construction projects." *Accid. Anal. Prev.*, 54(1), 122–133.

Li, X. (2007). "Finite element analysis of slope stability using a nonlinear failure criterion." *Comput. Geotech.*, 34(3), 127–136.

Liai, S.-H., and Chou, S.-Y. (2012). "Data mining techniques and applications—A decade review from 2000 to 2011." *Expert Syst. Appl.*, 39(12), 11303–11311.

Lienhart, W. (2015). "Case studies of high-sensitivity monitoring of natural and engineered slopes." *J. Rock Mech. Geotech. Eng.*, 7(4), 379–384.

Maffini, C., Marques, J., Kruglianskasb, I., Barbieri, L. A., and Schoproni, R. (2014). "Management for sustainability in companies of the mining sector: An analysis of the main factors related with the business performance." *J. Cleaner Prod.*, 84(1), 84–93.

Mainardi, S. (2005). "Earnings and work accident risk: A panel data analysis on mining." *Resour. Policy*, 30(3), 156–167.

Martín, J. E., Rivas, T., Matías, J. M., Taboada, J., and Argüelles, A. (2009). "A Bayesian network analysis of workplace accidents caused by falls from a height." *Saf. Sci.*, 47(2), 206–214.

Matías, J. M., Rivas, T., Martín, J. E., and Taboada, J. (2008). "A machine learning methodology for the analysis of workplace accidents." *Int. J. Comput. Math.*, 85(3-4), 559–578.

Paul, P. S., and Maiti, J. (2007). "The role of behaviour factors on safety management in underground mines." *Saf. Sci.*, 45(4), 449–471.

Sacristan, F. R. (2005). *Las 5S. Orden y limpieza en el puesto de trabajo*, FC Editorial, Madrid, Spain.

Sanmiquel, L., Freijo, M., Edo, J., and Rossell, J. M. (2010). "Analysis of work related accidents in the Spanish mining sector from 1982-2006." *J. Saf. Res.*, 41(1), 1–7.

Sanmiquel, L., Rossell, J. M., and Vintró, C. (2015). "Study of Spanish mining accidents using data mining techniques." *Saf. Sci.*, 75(1), 49–55.

Swuste, P., Frijters, A., and Guldenmund, F. (2012). "Is it possible to influence safety in the building sector?: A literature review extending from 1980 until the present." *Saf. Sci.*, 50(5), 1333–1343.

Tan, H., Wang, H., Chen, L., and Ren, H. (2012). "Empirical analysis on contribution share of safety investment to economic growth: A case study of Chinese mining industry." *Saf. Sci.*, 50(7), 1472–1479.

Wang, T., Liao, B., Ma, X., and Fang, D. (2010). "Using Bayesian network to develop a probability assessment approach for construction safety risk." *China Civ. Eng. J.*, 43(2), 384–391.

Weka [Computer software]. Univ. of Waikato, Waikato, New Zealand.

Witten, I. H., Frank, E., and Hall, M. (2011). *Data mining: Practical machine learning tools and techniques*, 3rd Ed., Morgan Kaufmann, Burlington, MA.

Wu, X., Liu, H., Zhang, L., Skibniewski, M. J., Deng, Q., and Teng, J. (2015). "A dynamic Bayesian network based approach to safety decision support in tunnel construction." *Reliab. Eng. Syst. Saf.*, 134(1), 157–168.

Zhang, L., Wu, X., Skibniewski, M. J., Zhong, J., and Lu, Y. (2014). "Bayesian-network-based safety risk analysis in construction projects." *Reliab. Eng. Syst. Saf.*, 131(1), 29–39.