Risk Analysis, Vol. 28, No. 1, 2008 DOI: 10.1111/j.1539-6924.2008.00998.x

Application of Correspondence Analysis in the Assessment of Mine Tailings Dam Breakage Risk in the Mediterranean Region

Ana Rita Salgueiro,[1*] Henrique Garcia Pereira,[1] Maria-Teresa Rico,[2] Gerado Benito,[3] and Andrés Díez-Herreo[4]

A new statistical approach for preliminary risk evaluation of breakage in tailings dam is presented and illustrated by a case study regarding the Mediterranean region. The objective of the proposed method is to establish an empirical scale of risk, from which guidelines for prioritizing the collection of further specific information can be derived. The method relies on a historical database containing, in essence, two sets of qualitative data: the first set concerns the variables that are observable before the disaster (e.g., type and size of the dam, its location, and state of activity), and the second refers to the consequences of the disaster (e.g., failure type, sludge characteristics, fatalities categorization, and downstream range of damage). Based on a modified form of correspondence analysis, where the second set of attributes are projected as "supplementary variables" onto the axes provided by the eigenvalue decomposition of the matrix referring to the first set, a "qualitative regression" is performed, relating the variables to be predicted (contained in the second set) with the "predictors" (the observable variables). On the grounds of the previously derived relationship, the risk of breakage in a new case can be evaluated, given observable variables. The method was applied in a case study regarding a set of 13 test sites where the ranking of risk obtained was validated by expert knowledge. Once validated, the procedure was included in the final output of the e-EcoRisk UE project (A Regional Enterprise Network Decision-Support System for Environmental Risk and Disaster Management of Large-Scale Industrial Spills), allowing for a dynamic historical database updating and providing a prompt rough risk evaluation for a new case. The aim of this section of the global project is to provide a quantified context where failure cases occurred in the past for supporting analogue reasoning in preventing similar situations.

KEY WORDS: Correspondence analysis; preliminary risk evaluation; qualitative regression; supplementary projection; tailings dam breakage

1. INTRODUCTION

The ancient mining exploitations spread all over the Mediterranean region, together with some active mines (a small fraction of the former), have produced, during their operation, an enormous amount of sludge, dumped into tailing ponds that are in general supported by precarious dams, constructed from locally obtained fills (soil, coarse waste, overburden from mining operation and tailings). Given the hazardous nature of the material contained in such tailing ponds, coupled with the quite unsafe condi-

[1]CERENA—Centro de Reursos Naturais e Ambiente do IST, Lisboa, Portugal.
[2]Instituto Pirenaico de Ecología, CSIC, Zaragoza, Spain.
[3]Institute of Natural Resources, CSIC, Madrid, Spain.
[4]Geological Hazards Unit, Spanish Geological Survey (IGME), Madrid, Spain.
[*]Address correspondence to Ana Rita Salgueiro, CERENA/IST, Av. Rovisco Pais, 1049-001 Lisboa, Portugal; tel: (+351) 218 417 247; fax: (+351) 218 417 389; rita.salgueiro@ist.utl.pt.

tion of most dams, it is felt that, if a breakage should come about, serious damage in the downstream area would be likely to occur.

In spite of this distressing state of affairs, little attention has been paid to this critical issue, since protection agencies are not aware of the risk involved in this type of disaster (and the "responsibility" to prevent it is very diluted among a variety of environmental institutions).

These circumstances were slightly changed after the Aznalcóllar large scale sulfide tailings pond spill on April 25, 1998, in southwest Spain,[1–2] which called public concern to this type of problem in Mediterranean countries. The study of this case after the failure was intensively carried out,[3–5] proving that remediation costs are so high that a rigorous alarm system for catastrophes of this kind should be seriously envisaged. But the singular set of conditions involved in this episode makes it unique and almost unhelpful by itself in establishing a reliable general methodology for coping with a wide-ranging situation.

In addition, as opposed to water reservoir dams, only modest experience is available in tackling this question in a more or less universal way, since it often reaches high levels of complexity (and suffers from a severe lack of reliable data). In fact, not only is the source, generally, poorly characterized in terms of its physical-chemical properties (controlling the mobility of the sludge and its damage potential), but also the dams take a variety of forms and types,[6] whose conditions make it difficult to assess their stability in geotechnical terms, and to model the discharge, if the breakage actually occurs. Hence, a detailed study of this kind, involving the calibration of physical models to forecast the probability of dam failure and the estimation of downstream damage in case of disaster[7–8] is a very costly and time-consuming task (apart from being also strongly site-specific and difficult to generalize).

Running in parallel with the first stages of the above mentioned detailed studies (focused only in four dams where geotechnical parameters were measured to feed physical models), a pilot statistical methodology was devised in the scope of an EU project (e-EcoRisk—A Regional Enterprise Network Decision-Support System for Environmental Risk and Disaster Management of Large-Scale Industrial Spills).

The proposed statistical methodology was developed, validated, and tested in the first stage of the project, giving rise to a general statistical setting where the detailed physical studies were integrated. Furthermore, for the goal of designing the global dedicated system aimed at detecting situations where a spill is prone to occur, the pilot statistical study presented here proved to offer a valuable framework where the study of new cases fits suitably.

2. METHODOLOGY

2.1. Outline of the Proposed Approach

The proposed methodology relies on the application of correspondence analysis as a qualitative prediction tool[9–11] to assess approximately the risk of large-scale spills in tailing dams located in preselected test sites (chosen to guarantee the demonstrativeness of the procedure), on the grounds of an empirical database containing the available set of historical disasters for the same type of situation. Obviously, risk is viewed here by emphasizing its damage dimension, since this study is based on recorded catastrophes that actually occurred in the past.

Given an array of empirical cases where a disaster occurred, the problem to be approached by the proposed methodology can be summarized in the following steps. The first step aims to extract from the database a set of variables (denoted "predictors") that are observable in a new case where the risk of failure is to be estimated. These variables are nominal in nature (type of dam, country where the pond is located, dam fill material, . . .) or are transformed into ordinals by creating significant classes in quantitative variables, like the dam height or volume. Moreover, in some cases, the order of magnitude of such variables can be guessed, if an appropriate interval is provided; in other cases, the available information does not allow a figure to be assigned to such variables, but to allocate them to a certain interval.

As a result of the above described procedure, the q disparate predictors extracted from the database

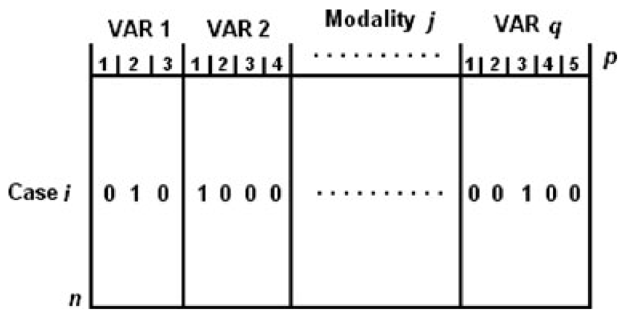

Fig. 1. Complete disjunctive format (matrix **A**).

are arranged in the "complete disjunctive matrix" **A** outlined in Fig. 1 under a common format. Each line of the $n \times p$ matrix **A** represents a failure episode (case i) and each column represents a category of the predictors (modality j).

Each element of the matrix **A** (i, j) is coded as one if modality j occurs in case i, and as zero otherwise.

The second step consists of selecting, from the database, a new set of qualitative variables that are linked to the failure episode and to the damage resulting from the disaster (the two components of the risk, for a new case). These variables are arranged under the same format given in Fig. 1 and give rise to a matrix **B** $(n \times p')$, containing the p' relevant attributes to be predicted in the new cases.

The third step seeks to establish some sort of relationship between **B** and **A.** Obviously, this cannot be achieved by establishing an equation as in ordinary regression, since all variables are qualitative. But a specific kind of graphical relationship between **B**- and **A**-type variables can be obtained if **B** is projected onto the factorial axes resulting from the eigenvalue decomposition of **A** through the correspondence analysis algorithm. This relationship is mediated by the factorial axes, which play the role of a "transfer function" between **B** and **A** (summarizing **A**-type of variables in quantitative coordinates, which are linked, through the same metric, with corresponding **B**-type variables coordinates).

The fourth step consists of using the relationship given by the previous procedure to forecast the modalities where **B**-type attributes fall, for a new matrix **C** $(n' \times p)$ containing only the predictors codified under the same format given in Fig. 1, and referring to the selected n' test sites where the risk is to be assessed.

2.2. Correspondence Analysis as a Qualitative Prediction Tool

The third step of the above described methodology is the crucial point to be dealt with where prediction is concerned. In order to get a satisfactory relationship between matrices **B** and **A**, correspondence analysis should not be used under its most common form (as a descriptive method, cf. Reference 12), but as a qualitative regression tool. This calls for the concepts of "supplementary projection" and "relative contribution." The former places **B**-type variables onto the factorial axes provided by matrix **A**, and the latter measures the quality of the relationship (the analogue of the "correlation coefficient," in ordinary regression). The supplementary projection of modality j' of matrix **B** onto the axis α provided by correspondence analysis of matrix **A** under the complete disjunctive format of Fig. 1 is given by:

$$f^{+}_{j'\alpha} = \frac{1}{n_{j'}\sqrt{\lambda_\alpha}} \sum_{i=1}^{n} \delta_{ij'} f_{i\alpha}, \qquad (1)$$

where

- $n_{j'}$ is the sum of column j', representing the total number of occurrences of each modality
- λ_α is the α-eigenvalue provided by correspondence analysis of matrix **A**

$$\delta_{ij'} \begin{cases} 1 \text{ if modality } j' \text{ occurs in line } i \\ 0 \text{ otherwise} \end{cases}$$

- $f_{i\alpha}$ is the projection of the line i onto the α-eigenvector provided by correspondence analysis of matrix **A**

It is worth noting that, as expected, all terms of Equation (1) depend only on the eigenvalue decomposition of matrix **A**.

Using Equation (1), all modalities of the variables to be predicted are related to predictors (summarized in axes α depending only on matrix **A**).

The relative contribution of axis α to modality j' is given by:

$$C^{r}_{\alpha j'} = \frac{(f^{+}_{j'\alpha})^2}{\left(\sum_{\alpha} f^{+}_{j'\alpha}\right)^2} = \cos^2\beta, \qquad (2)$$

where β is the angle formed by axis α and modality j'.

The more the relative contribution of the axis α to a given modality is close to 1, the more that modality is "associated" with axis α, which in turn relates to a subset of predictors, interpreted in terms of correspondence analysis.

This interpretation is performed on the grounds of "classical" correspondence analysis theory referring only to matrix **A**, by applying the inertia criterion:[12] a given axis is explained by the combination of predictors that exceeds the proportion of the total inertia that would be assigned to these predictors for a hypothetical uniform distribution. Also, the lines of matrix **A**, representing the historical cases, can be projected onto the same axes, as is usual in correspondence analysis, where individuals and properties are displayed in the same graph, as opposed to other factorial techniques like principal components analysis.[12]

Once interpreted, the axes that explain the behavior of predictors are associated with the attributes to be predicted, applying a maximization criterion to relative contributions. Hence, the variables linked to risk provide a new interpretation of the axes, related to dam failure and expected damage.

As expected from the specific nature of the problem, no "analytical expression" relating **B**- to **A**-type variables is obtained. However, the projections of each modality of **B**-type variables onto the axes are given by Equation (1). Therefore, since the same axes are related to **A**-type variables through their coordinates, the "qualitative regression" is performed in graphical terms, "mediated" by the axes.

Now, the new cases containing only **A**-type variables—arranged under the complete disjunctive format in the matrix **C** ($n' \times p$)—are projected onto the previously obtained axes, according to Equation (3):

$$f^{+}_{i'\alpha} = \frac{1}{q\sqrt{\lambda_\alpha}} \sum_{j=1}^{p} \delta_{i'j} f_{j\alpha}, \qquad (3)$$

where

q is the number of **A**-type variables
λ_α is the α-eigenvector provided by matrix **A**
$\delta_{ij'}$ $\begin{cases} 1 \text{ if modality } i' \text{ occurs in line } j \\ 0 \text{ otherwise} \end{cases}$
$f_{j\alpha}$ is the projection of column j onto the α-eigenvector provided by matrix **A**

Obviously, the total number of variables $q \vee i'$ equals $\sum_{j=1}^{p} \delta_{i'j}$, given the complete disjunctive format of matrix **A** (Fig. 1).

Equation (3) gives rise to a quantitative scale where the test sites can be sorted by ascending order of risk, according to their coordinates in the axes that were interpreted in terms of **B**-type variables (see Fig. 2 for a summary of the entire procedure).

3. CASE STUDY

3.1. Empirical Historical Database Construction

On a worldwide scale, the information used for the historical database construction refers chiefly to 221 tailings dam incidents reported in Reference 13, and to 185 tailings dam incidents during the period 1917–1989 reported in Reference 14. Additionally, The U.S. Environmental Protection Agency,[15] with recent mining and mineral processing damage cases in the United States, and the U.N. Environmental Programme,[16] with 83 major tailings dam failures (updated in March 2006), were also data sources exploited in this phase of the project for the compilation of relevant occurrences of tailings dam breakage and spill.

After a process of revision, cross-checking, and validation, the historical failure database was fed with the above mentioned information. In addition, new data were collected from journals, conference proceedings, reports, published and unpublished dissertations, and web pages.[17–30]

When this compilation, homogenization, and standardization procedure was accomplished, a total set of 147 cases of tailings dam failures was recorded. For each case, all the information related to the tailings pond characteristics, the accident, and its consequences was looked at and documented in a data form. This leads to a database structure composed of five sections or tables that contain the principal attributes of the dam, its failure conditions, and the impacts (as outlined in Table I).

The useful information that could be extracted from Table I is obviously incomplete, scarce, and disparate, calling for a constant updating effort by adding new cases and by improving the reported cases through deeper historical inquiry. In fact, when

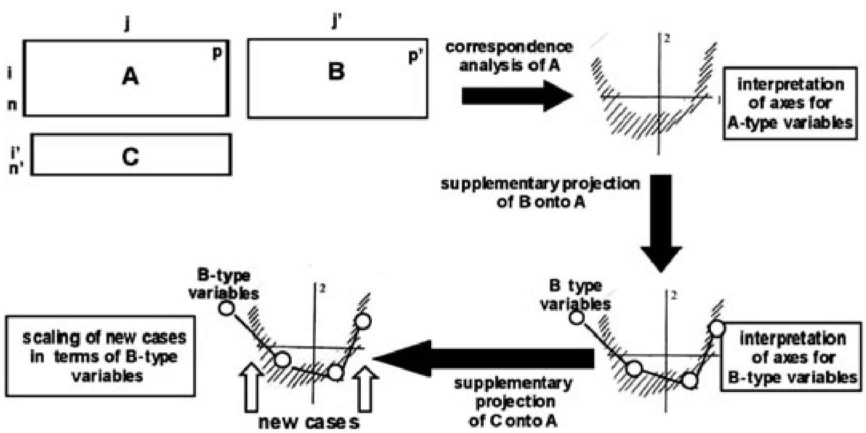

Fig. 2. Outline of the procedure to use correspondence analysis as a qualitative regression tool.

Table I. Database Structure and Main Attributes

Table Name	Components	Main Variables
Dam location	Spatial elements	Mine name, dam name, country, region, province, municipality, and reach name.
Tailings dam characteristics	Physical and chemical characteristics	Physical: dam type, dam situation, type of sequentially raised tailing dam, dam foundation, dam fill material, state of activity, storage volume, tailings dam height, tailings density, and retained water volume. Chemical: ore type, percentage solids by weight and the pond depth to be maintained.
Tailings dam failure	Failure cause(s)	Management operations, seismic liquefaction, rise of the groundwater table, mass movement/slope instability, fluvial undermining, inadequate/insufficient beach or free board, piping/seepage, dam overtopping/overflow, foundation failure, water level rise, snow melt, inadequate decant pipe construction, unusual rainfall event/period.
	Type of failure and duration	Breach, hole, overtopping/overflow.
Sludge characteristics	Sludge and water characteristics and movement	Sludge movement (distance, speed, thickness, and extension), tailings released (e.g., volume, density, and chemical characteristics), and associated water characteristics (e.g., volume, reach length).
Failure impacts	Socioeconomic impacts	Number of fatalities, number of injured, economic losses, affected facilities, etc.
	Environmental impacts	RAMSAR: Classification system for wetlands. Degree of protection at the international (RAMSAR, NATURA 2000, etc.), national, or regional level in the zone affected by the sludge.

an accident actually occurs, it is often difficult to access basic information regarding the tailings dam and its condition prior to the incident (e.g., dam height, tailings volume, water content, etc.). This situation usually stems from the self-interest of mining companies who are prone to hide information about both the accident itself and its consequences. In addition, in some countries where environmental legislation is, or has been, very negligent, governmental inspection bodies do not properly exercise their control duties. In these countries, environmental problems induced by tailings dam failures are often put out of sight of the public.

As an end result of the above mentioned issues, which often inhibit the dissemination of relevant data on tailings dam failures, only 37% of the reported instances fit the requirements of the proposed statistical analysis. In fact, out of the 147 episodes contained in the historical database, only 55 can be used in the sequel, given that the complementary subset does not contain a common group of **A**-type variables, that is, the observable tailings dam characteristics mentioned in the first section of Table I, which are the backbone of the proposed methodology.

The fraction of the entire database containing the above mentioned 55 cases is the corpus to be inputted for the application of the proposed methodology to the e-EcoRisk test sites. It is composed of both qualitative (e.g., failure causes) and quantitative (e.g., tailings dam height, storage volume) data. In order to use the same algorithm for both types of variables, they were put under a common format, shown in Fig. 1. Hence, quantitative variables were categorized by splitting their histogram into significant classes (e.g., the information on tailings dam height (m) was divided into three possible categories, $\leq$15 m, 15–30 m, or $\geq$30 m).

3.2. Application to a Set of Test Sites in the Mediterranean Region

By exploring the database outlined in Table I, the two sets of variables that construct matrices **A** and **B** were extracted for the 55 instances where a common set of information exists regarding predictors and consequences. According to the previously described methodology, the eigenvalue decomposition of matrix **A** gives rise to the axes displayed in Fig. 3, explaining 56% of the information contained in the **A**-type variable modalities.

Even though a relatively small fraction of the total inertia is conveyed by axis 1 (35%), this is not the crucial point of the application of correspondence analysis as a qualitative regression tool. In fact, what matters is not the "importance" of the axis in terms of fitting the cloud of variables (as is the case when correspondence analysis is used as a descriptive

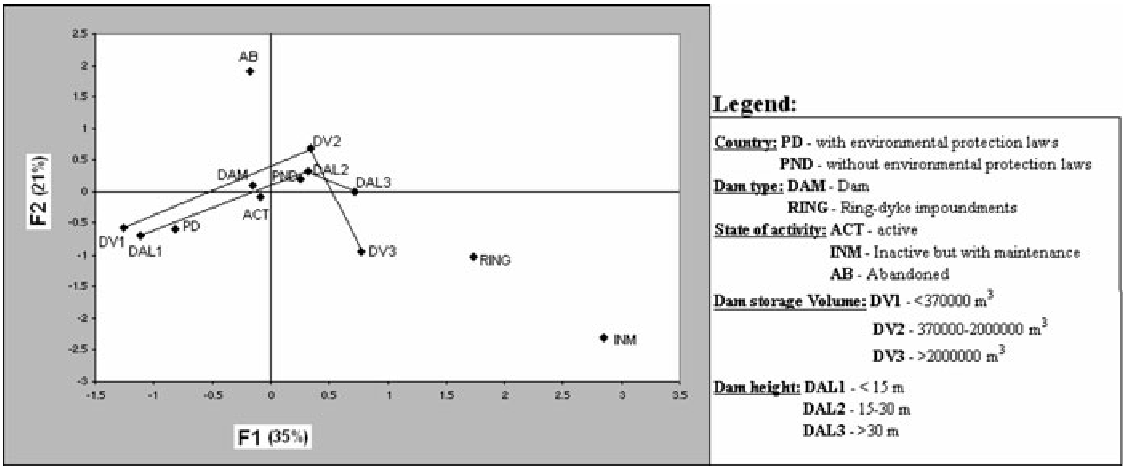

Fig. 3. Correspondence analysis results for matrix **A**.

procedure), but how such an axis may serve as a "mediator" between **A**- and **B**-type variables. The important goal at this stage of the methodology is that axis 1 allows the summary of a certain subset of predictors used in the sequel for linking them to **B**-type variables. This subset of predictors was selected by the inertia criterion given in the methodological section (specifically in Section 2.2), retaining those variables whose linkage to axis 1 exceeds the uniform distribution threshold.

When axis 1 is interpreted in terms of **A**-type variables, it shows a clear sequence (from left to right) in what concerns an increasing dam height (DAL1>DAL2>DAL3) and storage volume (DV1>DV2>DV3).

To the negative side of the axis is associated PD (location in an environmentally regulated country) and to the positive side are associated the RING and INACTIVE (IMM) types of dam located in an environmentally unregulated country.

Hence, axis 1 shows the opposition between small dams in environmentally regulated countries (negative semi-axis) versus big, inactive, and ring dams in environmentally unregulated countries (positive semi-axis).

Now, looking at the **B**-type modalities projected as supplementary attributes according to Equation (1), Fig. 4 is obtained, where the axes are the same as in Fig. 3. The first conclusion drawn from Fig. 4 is that only axis 1 is relevant for interpretation, according to relative contributions calculated by Equation (2). Therefore, axis 2, revealing the opposition between abandoned dams (AB) and inactive dams (INM), can be disregarded from the qualitative

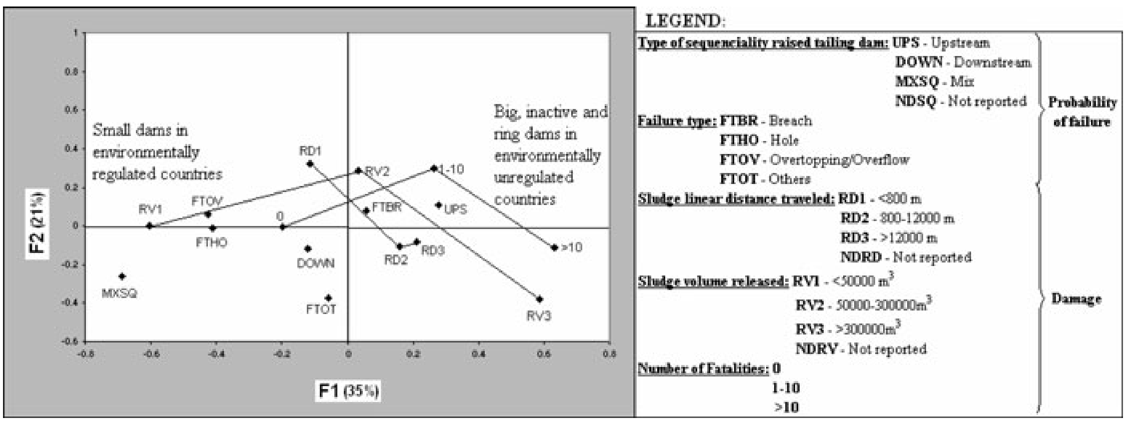

Fig. 4. Supplementary projection of **B**-type attributes onto the axes provided by matrix **A** eigenvalue decomposition.

Table II. Relative Contribution of Significant Variable Modalities for Negative and Positive Semi-Axis 1

Negative Semi-Axis:		Positive Semi-Axis:	
Sludge volume released <50,000 m^3	(0.75)	Upstream type of sequenciality raised tailing dam	(0.63)
Mix type of sequenciality raised tailing dam	(0.62)	Sludge linear distance traveled >12,000 m	(0.51)
No fatalities	(0.55)	Sludge volume released >300,000 m^3	(0.47)
Failure type: hole	(0.20)	> 10 fatalities	(0.40)
Failure type: overtopping/overflow	(0.16)	Sludge linear distance traveled between 800 and 12,000 m	(0.15)
Downstream type of sequenciality raised tailing dam	(0.12)	1–10 fatalities	(0.11)

regression analysis, since it is not linked to any relevant variable indicating the importance of the breakage episode in terms of reported damage.

Hence, both negative and positive semi-axis 1, which were already interpreted in terms of an apposition between modalities of **A**-type variables, can now be associated with the modalities of **B**-type variables given in Table II, sorted by descending order of their relative contribution, calculated by Equation (2).

Table II and Fig. 4 show clearly that axis 1 corresponds to a scale of risk (increasing from left to right). This empirical scale of risk provides a quantified graphical relationship between the modalities of **B**- and **A**-type variables. In fact, the "importance" of such modalities is given by their coordinate in axis 1, which is thus the analogue of the "coefficients" applied to each variable in the equation linking two sets of real continuous variables for the "classical" regression framework.

Now, given the test sites located in Fig. 5, they can be sorted by ascending order of risk (Fig. 6), when matrix **C** is projected as a set of supplementary lines onto the same axis 1, according to Equation (3).

Furthermore, each test site is associated in Fig. 6 with the historical cases projected onto axis 1 within a given neighborhood, allowing for analogy reasoning on the grounds of the hypothesis that a particular test site is "similar" to the historical cases contained in such a vicinity. These historical cases are displayed in Fig. 6 as reference numbers whose meaning is given in Table III.

Fig. 5. Location of test sites.

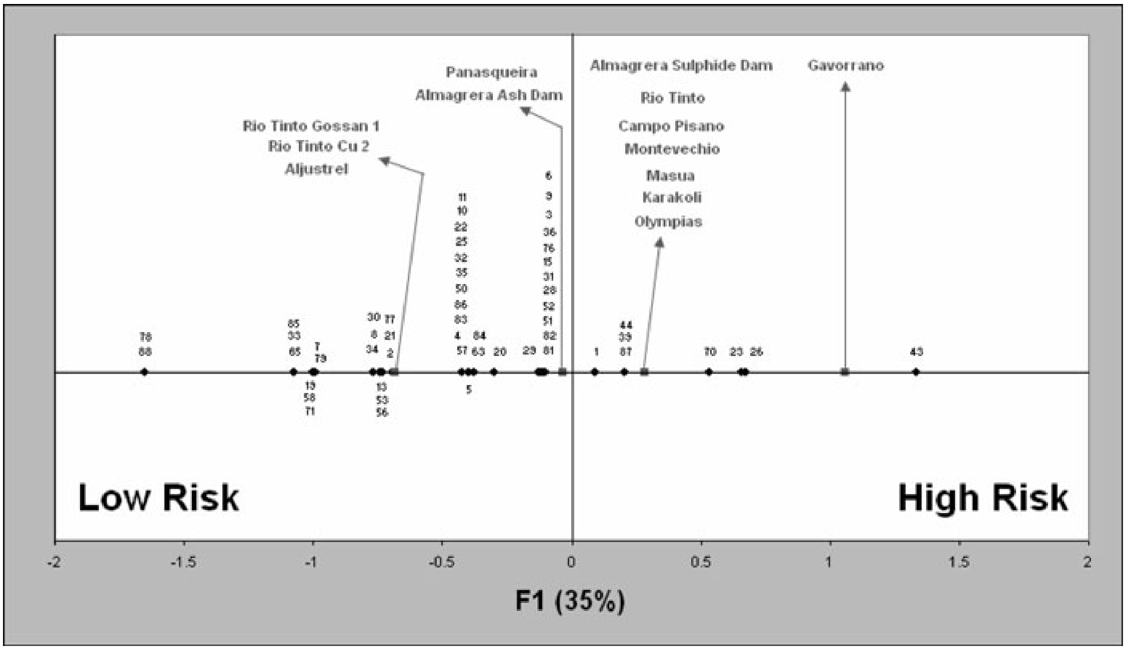

Fig. 6. Supplementary projection of test sites onto axis 1 provided by matrix **A**.

3.3. Validation and Discussion of the Case Study Results

From the results of the above described case study, the following preliminary conclusions can be drawn.

(i) The initial correspondence analysis of the table containing only predictors leads to a straightforward interpretation of the first axis produced by the eigenvalue decomposition of matrix **A**. This axis, denoted by F1 in Figs. 3, 4, and 6, leads to the identification of two main clusters of basic characteristics of historical episodes tailings dam failure: (a) small dams in countries with strict environmental laws at the time of the dam construction, and (b) large, inactive, and ring-type dams in countries with poor records of environmental regulation.

(ii) Supplementary attributes related to the conditions of failure (type of dam raise, failure category) and incurred damage (sludge distance and volume, and number of casualties) were incorporated in the analysis. A positive relationship was found among upstream raised tailing dams, the largest category of sludge travel distance (>12 km) and sludge volume released (>300,000 m^3), and the largest number of casualties (>10 deaths). Comparing Figs. 3 and 4, it is clear that modalities indicating a sequence from low to high risk follow axis F1 from its negative to the positive side, allowing the association of them with predictors that project onto F1 in the same areas, according to the clusters identified in (i).

(iii) An evaluation of the potential risk regarding the dams chosen as test sites for the e-EcoRisk project (located in European Mediterranean countries) was obtained by the proposed methodology. The greatest potential hazard is at the Gavorrano tailings dam (Italy) of ring-dyke type of height class 2 (15–30 m), volume class 3 (>2,000000 m^3), and that is currently inactive, though subject to maintenance and control tasks. The high degree of potential disaster for this particular dam stems from the combination of its size with its type, conditions that are referred to in literature as particularly prone to induce the highest failure risk.[17] The lowest potential hazard occurs at the Rio Tinto Copper 1 and Rio Tinto Gossan 1 (Spain) and Aljustrel tailings dams, which are currently active and, therefore, have a greater degree of maintenance. Intermediate risk positions comprise the remaining cases, where dam type embankments, although inactive, are currently subject to ongoing maintenance.

Table III. List of the 55 Cases that Compose the Historical Database

Ref. Number	Name	Country	Year of the Incident	Ref. Number	Name	Country	Year of the Incident
1	Los Frailes	Spain	1998	36	El Cobre Old Dam	Chile	1965
2	Aitik	Sweden	2000	39	Fort Meade	USA	1971
3	Baia Borsa	Romania	2000	43	Harmony, Merriespruit	South Africa	1994
4	Baia Mare	Romania	2000	44	Hokkaido	Japan	1968
5	Sgurigrad (Zgorigrad)	Bulgaria	1966	50	Itabirito	Brazil	1986
6	Maritsa Istok 1	Bulgaria	1992	51	Jinduicheng	China	1988
7	Stava	Italy	1985	52	La Patagua New Dam	Chile	1965
8	Balka Chuficheva	Russia	1981	53	Los Maquis	Chile	1965
9	Zletovo	Macedonia (Yugoslavia)	1976	56	Mike Horse	USA	1975
10	Maggie Pie	United Kingdom	1970	57	Mochikoshi No. 1	Japan	1978
11	Bilbao	Spain	1964	58	Montcoal No. 7, Raleigh County	USA	1987
13	Derbyshire	United Kingdom	1966	63	Olinghouse	USA	1985
15	Madjarevo	Bulgaria	1975	65	Omai	Guyana	1995
19	Middle Arm	Tasmania	Unknown	70	Placer, Surigao del Norte	Philippines	1995
20	Partizansk, Primorski Krai	Russia	2004	71	Riverview	USA	2004
21	Huelva	Spain	1998	76	Sipalay	Philippines	1982
22	Amatista, Nazca	Perú	1994 or 1966	77	Stancil	USA	1989
23	Arcturus	Zimbawe	1978	78	Sullivan mine	Canada	1991
25	Bafokeng	South Africa	1974	79	Tennessee Consolidated No. 1	USA	1988
26	Bellavista	Chile	1965	81	Unidentified	SW USA	1973
28	Buffalo Creek	USA	1972	82	Veta de Agua No. 1	Chile	1985
29	Cerro Negro	Chile	2003	83	Barahona,	Chile	Unknown
30	Cerro Negro No. 4	Chile	1985	84	Bonsal	USA	Unknown
31	Cerro Negro No. 3	Chile	1695	85	Mochikoshi n2	Japan	Unknown
32	Church Rock	USA	1979	86	Phelps-Dodge	USA	Unknown
33	Deneen Mica	USA	1974	87	Silver King	USA	1974
34	Unidentified, East Texas	USA	1966	88	Unidentified	USA	Unknown
35	El Cobre New Dam	Chile	1965				

(iv) A crucial aspect that emerged from the analysis is the importance of the environmental regulations of countries where tailing dams are located. Since such regulations have changed substantially over the last two decades, in particular for the Mediterranean region where test sites of the e-EcoRisk project are located, it is useful to refer to Table IV for analyzing the results of the case study in the context of EU environmental norms.

It is worth noting that the above outlined conclusions refer only to the specific array of variables selected as predictors for this particular case study. These were the common set of attributes that were found in the database, which generated the instances allowing the illustration of the methodology for a statistically significant corpus. Given the general content of this set of predictors, it is not surprising that practical results obtained in the case study may appear as "trivial" (or expectable) when experts in

Table IV. Dam Construction and End of Activity at the Case Study Sites

e-EcoRisk Test Sites	Construction Year	End of Activity
Karakoli (Greece)	1980–1985	2003
Olympias (Greece)	1985–1987	1995
Gavorrano (Italy)	1908	1981
Masua (Italy)	1965	1987
Montevecchio (Italy)	1930	1990
Campo Pisano (Italy)	1960	1991
Panasqueira (Portugal)	1905–1910	1996
Almagrera Ash (Spain)	1984	2002
Almagrera Sulfide (Spain)	1985	2002
Aljustrel (Portugal)	1845	1992*
Rio Tinto (Spain)	1873	1995

* Aljustrel is now restarting activity.

the field discuss them. Paradoxical as it may seem, this trait demonstrates the coherence of the methodology. In fact, expert opinion on results obtained for the case study confirmed both the positioning of individual test sites and historical occurrences in the empirical scale of risk produced by the methodology. This constitutes a sound basis for the validation of the essence of methodology in general terms, independently of the set of variables that are in play. Moreover, in the e-EcoRisk project context, the methodology proved to be very useful for a preliminary risk alert, providing a rough automatic risk estimation, prior to any expert intervention.

Should the variables contain a more detailed content, it is likely that the methodology would give rise to more relevant results, from a practical point of view. For instance, if a significant set of common geotechnical variables were collected and integrated in the database, it is clear that the qualitative regression problem could be solved using the same methodology but leading to a more positive solution, in which novel and unforeseen results are concerned.

A critical analysis of the outcome reached in this particular case study permits the assertion that this application's main limitations stem from the lack of comparable information regarding historical tailings dam breaks. Steps are to be taken to maximize the formal homogeneity of predictors, along the lines already pursued when categorization of quantitative variables allowed their use alongside qualitative ones in the same algorithm. But such a homogenization purpose should be achieved under the constraint of avoiding the oversimplification that emerges from the general nature of the case study presented. This leads to a tradeoff between two contradictory goals, which does not affect the essence of the methodology, but only the selection of predictors. However, given the robustness of the correspondence analysis algorithm, which the methodology relies on, it is plausible that the improvement of results achieved by the addition of a new set of relevant variables balances the loss in statistical significance incurred by using the smaller number of historical instances where the expanded set of variables is available.

4. CONCLUSIONS

When approaching the problem of tailings dam breakage statistical modeling, the proposed methodology proved to be effective in providing an empirical scale of risk, depending only on observable prefailure attributes (both quantitative and qualitative). This scale of risk, validated by expert knowledge, was constructed on the grounds of the information contained in a database of tailings dam disasters, where "risk" is equated with "damage." Such a database was purposely compiled to test the methodology in the scope of the e-EcoRisk project (available to "e-EcoRisk System" users).

Such a database, which is dynamically updated and contains 147 occurrences to date, is a valuable tool for analogy reasoning when considering a new case where some of the prefailure attributes can be observed. When coupled with the proposed statistical prediction methodology, this database may provide a prompt procedure to evaluate approximately the risk of failure of a given tailings dam, in order to envisage prevention measures and guide the collection of further information. Also, for a set of new cases, remediation practices to avoid breakage may be planned according to objective priorities based on similar historical contexts provided by the empirical scale of risk.

Further research is performed in the scope of the e-EcoRisk project to combine in the same system architecture a set of detailed physical models of failure (which are strongly information demanding) with the statistical approach proposed here (which is a much more parsimonious approach, where input data are concerned). Hence, the automatic system incorporating both approaches is led by the information available for each case: if such information is scarce, the proposed statistical methodology is applied; if detailed parameters are accessible (e.g., in the geotechnical domain), physical models are to be used. For the intermediate circumstances defining a new case, a combination of the two approaches is offered, the first one being a reliable basis for early alert/intervention prioritization, and the second a complete framework for complementary information gathering.

ACKNOWLEDGMENTS

This research has been funded by the European Commission through the project "A Regional Enterprise Network Decision-Support System for Environmental Risk and Disaster Management of Large-Scale Industrial Spills," e-EcoRisk Project (Contract EVG1-2002-0068) and by the Spanish Ministry of Science and Education (HP2006-0072) in conjunction with the Council of Deans of Portuguese Universities (E-58/07). The authors are very grateful to

Varyl Thorndycraft for the critical review of the original article, and for his very useful comments and suggestions.

REFERENCES

1. López-Pamo, E., Barettino, D., Antón-Pacheco, C., Ortiz, G., Arránz, J. C., Gumiel, J. C., Martínez-Pledel, B., Aparicio, M., & Montouto, O. (1999). The extent of the Aznalcóllar pyritic sludge spill and its effects on soils. *Science of the Total Environment*, *242*(1–3), 57–88.
2. Benito, G., Benito-Calvo, A., Gallart, F., Martín-Vide, J. P. Regües, D., & Bladé, E. (2001). Hydrological and geomorphological criteria to evaluate the dispersion risk of waste sludge generated by the Aznalcollar mine spill (SW Spain). *Environmental Geology*, *40*(4/5), 417–428.
3. CMA. (1998). Corredor Verde del Guadiamar. Balance de actuaciones para la corrección de los efectos del vertido minero de Aznalcóllar. Oficina Técnica para la Recuperación del Guadiamar, Consejería de Medio Ambiente, Junta de Andalucía. CD format.
4. Grimalt, J. O., & Macpherson, E. (Eds.). (1999). The environmental impact of the mine tailing accident in Aznalcóllar (South-West Spain). *Science of the Total Environment*, *242*(Special issue), 337.
5. Ayora, C., Barettino, D., Carrera, J., Manzano, M., & Mediavila, C. (Eds.). (2001). Las aguas y los suelos tras el accidente de Aznalcóllar. *Boletín Geológico y Minero*, *112*, Especial, 294.
6. Vick, S. G. (1983). *Planning, Design, and Analysis of Tailings Dams*. New York: Willey-Interscience Publication. Willey Series in Geotech.
7. Vick, S. G. (1999). Considerations for estimating structural response probabilities in dam safety risk analysis, U.S. Bureau of Reclamation, Technical Service Center, Denver, CO.
8. Dise, K. M., & Vick, S. G. (2000). Dam safety risk analysis for Navajo Dam. *Trans. 20th International Congress Large Dams*, *1*, 337–350.
9. Pereira, H. G. (1988). Case study on application of qualitative data analysis techniques to an Uranium mineralization. In F. Chung *et al*. (Eds.), *Qualitative Analysis of Mineral and Energy Resources*, (pp. 617–624). Dordrecht: Reidel Publishing Company.
10. Pereira, H. G., Ribeiro, L., Soares, A., Patinha, P., Pereira, M. J., & Dias, M. (1997). Geostatistical modelling of environmental variables at mine sites. In Baafi & Schofield (Eds.), *Geostatistics Wallangong'96*. *2*, 951–959.
11. Pereira, H. G., Sousa, A. J., Ribeiro, J., Albuquerque, T., Luís, G., Lamberto, V., & Saraiva, J. (2000). Case studies on the application of geostatistics to in situ natural stone characterization. In Kleingeld & Krige (Eds.), *Geostatistics 2000 Cape Town*, *2*, 757–766.
12. Greenacre, M. J. (1984). *Theory and Applications of Correspondence Analysis*. New York: Academic Press.
13. ICOLD. (2001). *Tailings Dams—Risk of Dangerous Occurrences, Lessons Learnt from Practical Experiences, Bulletin 121*. United Nations Environmental Programme (UNEP). Division of Technology, Industry and Economics (DTIE) and International Commission on Large Dams (ICOLD), Paris.
14. USCOLD. (1994). *Tailings Dam Incidents, U.S. Committee on Large Dams—USCOLD* Denver, CO.
15. EPA. (1997). *Damage Cases and Environmental Releases from Mines and Mineral Processing Sites*. Washington, U.S. Environmental Protection Agency, Office of Solid Waste.
16. UNEP. (1996). Environmental and safety incidents concerning tailings dams at mines: Results of a survey for the years 1980-1996 by Mining Journal Research Services. A report prepared for United Nations Environment Programme, Industry and Environment. Paris.
17. Blight, G. E. (1997). Destructive mudflows as a consequence of tailings dyke failures. *Proceedings of the Institution of Civil Engineers*, *125*, 9–18.
18. Bourcy, S. C., & Weeks, R. E. (2000). Stream morphology and habitat restoration of Pinto Creek, Gila County, Arizona. In Balkema (Ed.), *Tailings and Mine Waste 00* (pp. 467–475). Rotterdam.
19. Brink, D. (1998). The long-term repair of the Merriespruit tailings dam. In Balkema (Ed.), *Tailings and Mine Waste 98* (pp. 953–957). Rotterdam.
20. Chandler, R. J., & Tosatti, G. (1995). The Stava tailings dams failure, Italy, July 1985. *Proc. Instn. Civ. Engrs.*, *113*, 67–79.
21. Dobry, R., & Alvarez, L. (1967). Seismic failures of Chilean tailings dams. *Proc. ASCE, Jnl SM&F Division*, *93*(SM6), 237–260.
22. Fourie, A. B., Papageorfiou, G., & Blight, G. E. (2000). Static liquefaction as an explanation for two catastrophic tailings dam failures in South Africa. In Balkema (Ed.), *Tailings and Mine Waste 00* (pp. 149–158). Rotterdam.
23. Hansen, L. A., LaFronz, N. J., & Yasin, M. B. (2000). Stabilization of the Pinto Valley tailings impoundment slide. In Balkema (Ed.), *Tailings and Mine Waste 00* (pp. 477–487). Rotterdam.
24. Jeyapalan, J. K., Duncan, J. M. & Seed, H. B. (1983). Investigation of flow failures of tailings dams. *Journal of the Geotechnical Engineering Division*, *109*(2), 72–189.
25. Kostaschuk, R. N., & Wilkinson, J. M. T. (1999). Predicting the deformation of tailings dams resulting from earthquake liquefaction. In Balkema (Ed.), *Tailings and Mine Waste 99* (pp. 237–248). Rotterdam.
26. McPhail, G., Stuart, R. J., & Venter, D. (1998). The disaster of Merriespruit and its consequences: Remediation-making safe. In Balkema (Ed.), *Tailings and Mine Waste 98* (pp. 959–957). Rotterdam.
27. Miller, D. J. (1999). Failure modes and effects analyses of potential ground movements at Golden Sunlight Mine. In Balkema (Ed.), *Tailings and Mine Waste 99* (pp. 335–344). Rotterdam.
28. Van Niekerk, H. J., & Viljoen, M. J. (2005). Causes and consequences of the Merriespruit and other tailings-dam failures. *Land Degradation & Development*, *16*, 201–212.
29. Wagener, F. V. M., Craig, H. J., Blight, G., McPhail, G., Williams, A. A. B., & Strydom, J. H. (1998). The Merriespruit tailings dam failure. In Balkema (Ed.), *Tailings and Mine Waste 98* (pp. 925–952). Rotterdam.
30. Willow, M., & tenBraak, C. (1999). Survey of three hard-rock acid drainage treatment facilities in Colorado. In Balkema (Ed.), *Tailings and Mine Waste 99* (pp. 759–767). Rotterdam.