An Acad Bras Cienc (2023) 95(Suppl.1): e20220458 DOI 10.1590/0001-3765202320220458
Anais da Academia Brasileira de Ciências | *Annals of the Brazilian Academy of Sciences*
Printed ISSN 0001-3765 I Online ISSN 1678-2690
www.scielo.br/aabc | www.fb.com/aabcjournal

ENGINEERING SCIENCES

Building pathologies caused by failure of Fundão Tailing Dam: A principal component analysis aproach

LADIR ANTONIO DA SILVA JUNIOR & TATIANA B. DOS SANTOS

Abstract: This article presents a study of the impacts caused by Fundão dam failure in buildings from Gesteira district, Barra Longa city, Brazil. The analyzed dataset was built using technical reports from surveys carried out in 152 buildings. Principal component analysis was capable of explain the interdependence of data variable and allowed conditions of understand the consequences and evidenced pathologies. Heavy vehicle traffic caused more damage (57% of buildings) to the studied buildings than the contact with the tailing mud (43% of buildings).

Key words: Heavy vehicle traffic, impacts, pathologies, tailing mud, unsupervised machine learning.

INTRODUCTION

Mineral industry is very important to humanity, being responsible to the production of essential material for various human activities, such as: civil construction, agriculture, etc. Due to the growing of humanity needs, the ore requesting increased, low-content deposits began to be mined and the production of tailings increased. In year of 2010, the tailing production around the world was estimated in 14 billion of tonnes (Adiansyah et al. 2015).

Commonly, tailings generated by mineral processing are stored in dams. These structures have reached large volumes and heights and its management became a geotechnical concern. A tailing dam failure can lead to human, economic and environmental disastrous consequences. The risk of failures is an apprehension of whole world and public concern about the risk and impacts of these structures growing since recent incidents occurred (Rico et al. 2008).

Despite failures of tailing dams being a concern for several decades, Fundão dam from Samarco's iron mine collapsed in year of 2015. More than 32.6 million of m3 of iron ore tailings left the dam reservoir, causing a large environmental damage, affecting 680 km of watercourses from Rio Doce Basin to the Atlantic Ocean (Carmo et al. 2017, SAMARCO 2020). The volume of waste and the extent of ecosystems affected assumed unprecedented proportions, involving the Brazilian Atlantic Forest. The dam was built using upstream method and its height were 130 meters.

The failed volume of tailings caused the death of 19 people, destroyed buildings, houses and the infrastructure of Bento Rodrigues and Paracatu de Baixo districts (city of Mariana, Minas Gerais state), and Gesteira district (city of Barra Longa, Minas Gerais state) (SAMARCO 2020). Researchers analyzed the dispersion of the tailing in Rio Doce basin and the Atlantic Ocean, the contamination of reefs, the affectation of bee pollination, the dam failure catastrophe extent, among others studies of the large environmental and human health impacts caused by the Fundão dam failure (Coimbra et al. 2020, Orlando et al. 2020, Guerra et al. 2017, Marta-Almeida et al. 2016, Quadra et al. 2019).

Due to the impacts caused by one of the world large catastrophes related to tailing dam failures, this topic is still much discussed and researched in the last years. However, studies about the event consequences in terms of building pathologies were not found in the literature. Many buildings from Gesteira district were affected by the event, either by contact with the tailing mud or by the heavy vehicle traffic related to the reconstruction of its infrastructure. The tailing mud reached the district by Carmo river overflow.

Contact of tailing mud with building grounds can cause overloading. An overloading generates stress increase in the structure and it may cause masonry and/or structural elements cracking, foundation settlement and even building collapse (Watt 2007). The heavy vehicle traffic can be a source of vibration, inducing dynamic loads and waves, which reach the foundation of adjacent buildings (Hunaidi & Tremblay 1997).

This article presents an analysis of building pathologies caused by failure of Fundão tailing dam. In Gesteira district (Barra Longa city), 152 buildings were surveyed, analyzed and used to build a dataset. The dataset was analyzed using a unsupervised machine learning, named Principal Component Analysis (PCA). Machine learning is largely used for data analysis and these tools started to be used recently in civil and mining engineering (Santos et al. 2018, 2019, Ali et al. 2019, Coelho et al. 2016, Kang et al. 2021, Qi et al. 2019, Paulo et al. 2020, Yang et al. 2021, Santos & Oliveira 2021).

Principal component analysis was applied to the dataset to understand the interdependence between the surveyed variables. This knowledge allowed determine and evaluate the consequences and pathologies caused by Fundão dam failure to buildings.

MATERIALS AND METHODS

Research methodology

The analysis of pathologies in the buildings damaged by the Fundão dam failure consisted of eight steps. First step consisted of determining the variables to be surveyed in the building and production of technical reports. Second step included the dataset construction, using the variables collected at the field. Dataset knowledge was carried out in the third step. In the fourth step, the correlation matrix was defined and the non-significant variables were removed. Fifth and sixth step consisted of the application of Bartlett's Test and Principal Component Analysis (PCA), respectively. Finally, the seventh and eighth steps comprised the result analysis and conclusion of the research. Figure 1 presents the flowchart of the research methodology.

Determining of the qualitative and quantitative variables and survey of building pathologies

With the goal of analyzing the building pathologies caused by the dam failure, a technical survey was carried out in buildings from Gesteira district, Barra Longa city. In each surveyed building were identified its architectonic and urbanistic characteristics; structural system; pathologies; if the tailing

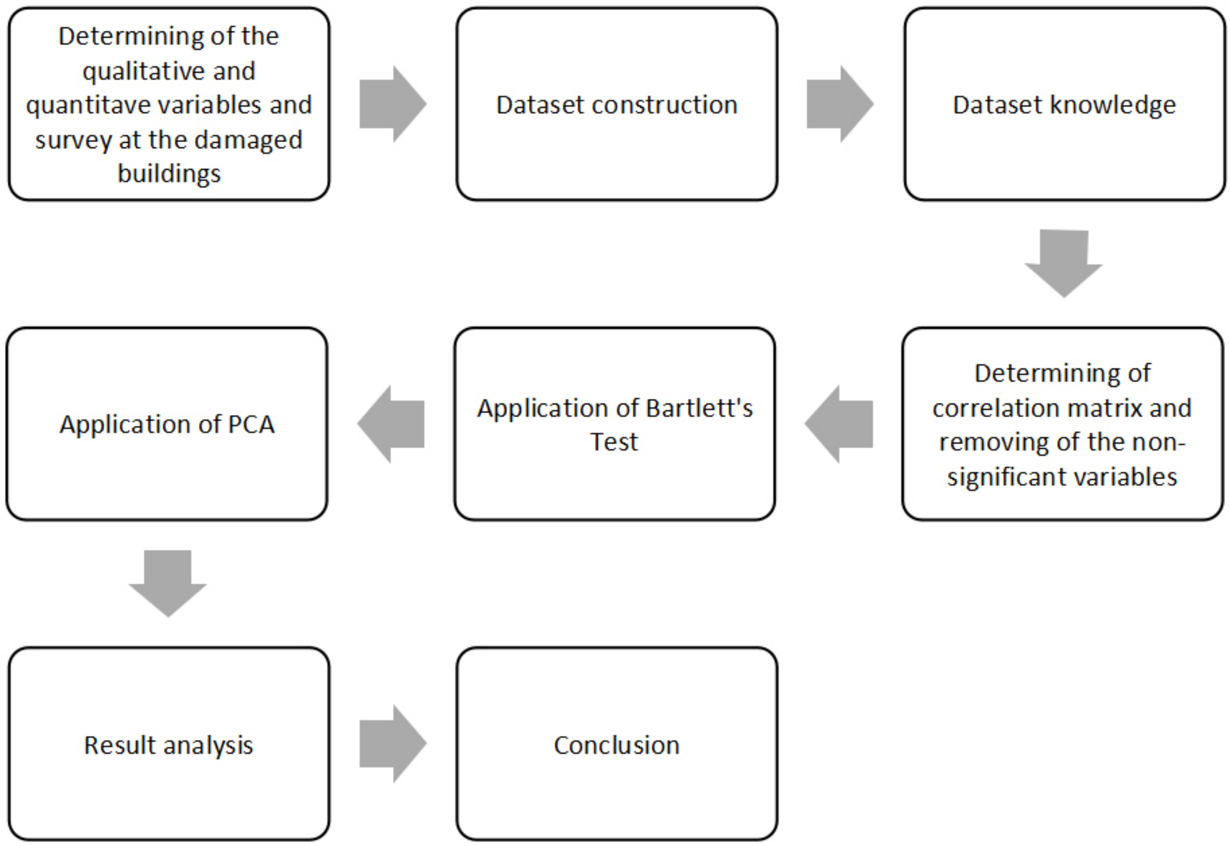

Figure 1. Methodology.

mud reached to the building ground; and the presence of heavy vehicle traffic correlated with the tailing clean-up.

Survey of building pathologies in this research was based on a visual inspection. According to the Brazilian entity, named Instituto Brasileiro de Avaliações e Perícias de Engenharia (IBAPE 2012), this type of survey is an inspection Level 1, which consists of an assessment with low technical complexity, normally employed in buildings with very simple or inexistent maintenance plans (case of the buildings located in Gesteira). Inspection Level 1 must to be carried out by qualified professionals through a visual inspection in loco.

Classification of anomalies and cracks and their probable origin factor was carried out in the buildings through in loco inspection. According to recommendations of the standard published by IBAPE, the anomalies were classified as endogenous, exogenous, natural and functional (IBAPE 2012). Cracks were classified as operational, managerial, associated to the planning or associated to execution of the building.

Pathologies were also classified according to their probable risk level. They were classified as critical, medium and minimum. Critical level is correlated to the risk of causing damage to health and security of personal and environment; loss of performance and functionality, possible occupancy stoppage; increase of maintenance and repair costs and significant loss of useful life. Medium level is correlated to the risk of causing the partial loss of building performance and functionality and it does not cause occupancy stoppage. Minimum risk level is associated to a minor damage to the building esthetics or planned maintenance activity, without probability of critical and medium risk occurrence, and low or anyone reduction of real estate value.

Building pathologies have various types, such as cracks, stains, displacements and deformations, among others. Cracking classification was based on its opening (tiny, medium and large). Cracks were also classified according to its possible causes in five groups: caused by overloading, thermal variations, retraction and expansion, excessive deflection of concrete structural elements and foundation settlement (Magalhães 2004).

Using the presented guidelines, an inspection was carried out on 152 buildings affected by Fundão dam failure from Gesteira district. For all buildings, qualitative and quantitative characteristics were surveyed, being subdivided into 19 variables, see Table I.

Table I. Surveyed variables.

Variable	Lable	Type
P1	Tailing reached to the building ground	Qualitative
P2	Heavy vehicle traffic	Qualitative
P3	Nature of building pathology is exogenous	Qualitative
P4	Risk level	Qualitative
P5	Building pathology worsen by the event	Qualitative
P6	Number of cracks caused by settlement	Quantitative
P7	Number of cracks caused by masonry expansion	Quantitative
P8	Number of cracks caused by thermal variations of the masonry	Quantitative
P9	Number of cracks caused by thermal variations of the floor system	Quantitative
P10	Number of cracks caused by floor system deflection	Quantitative
P11	Number of cracks caused by support deflection	Quantitative
P12	Number of cracks caused by superior beam deflection	Quantitative
P13	Number of cracks caused by structural deflection	Quantitative
P14	Number of horizontal cracks caused by overloading	Quantitative
P15	Number of vertical cracks caused by overloading	Quantitative
P16	Number of cracks caused by negative bending in the sills	Quantitative
P17	Number of cracks in openings (doors, windows, vents)	Quantitative
P18	Number of detachments	Quantitative
P19	Number of infiltrations	Quantitative

Dataset construction

The dataset was built using the variables surveyed in buildings of the city, which presented reports of the appearance and/or worsening of pathologies after the dam failure event. In case of quantitative variables (P6 to P19), the counting values were used. For qualitative variables, ordinal values were assigned ranging from 0 to 3 for the variable P1, 0 to 1 for the variables P2, P3 and P5 and a range of 0 to 2 was assigned to variable P4, see Table II. Table III presents the ten first buildings (ID1 to ID10) from the dataset and its surveyed variables (P1 to P19).

Table II. Values assigned to qualitative variables.

Variable	Rating 0	Rating 1	Rating 2	Rating 3
Tailing reached to the building ground – P1	No	Near	Lot	Building ground
Heavy vehicle traffic – P2	No	Yes		
Nature of building pathology is exogenous – P3	No	Yes		
Risk Level – P4	Minimum	Medium	Critical	
Building pathologies worsen by the event - P5	No	Yes		

Table III. Ten first constructions of the dataset.

ID	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19
1	2	1	1	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0
2	3	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
3	2	1	1	1	1	14	0	10	0	0	0	0	0	0	0	6	8	6	0
4	2	1	1	0	1	3	0	7	0	0	0	4	0	0	0	0	1	0	0
5	2	1	1	0	1	1	0	4	0	0	0	0	0	0	0	0	0	0	0
6	0	1	0	0	1	0	0	0	0	0	0	0	0	7	0	0	0	0	1
7	0	1	1	1	1	1	0	14	0	0	0	0	0	0	0	0	8	15	0
8	0	1	0	2	1	12	0	0	0	0	0	0	0	1	0	1	0	0	0
9	0	1	0	2	1	0	0	0	0	0	0	0	0	2	0	0	0	0	0
10	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	2	0

Dataset knowledge

Exploratory analysis of the collected data was carried out with the objective of knowledge by means of an analysis of the category frequency of each variable in the dataset. This investigation provides conditions of understanding the information of the variability contained in the data. Besides, an urban layout of researched district was drawn and the principal churches, bridges and streets were identified. The 152 constructions are located on this area.

Determining of correlation matrix and removing of the non-significant variables

Considering r_{jk} the correlation between the variables X_j and X_k and p the number of variables of a dataset. The data correlation matrix is a symmetric matrix ($r_{jk} = r_{kj}$), given by Equation 1.

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1P} \\ r_{21} & r_{22} & \cdots & r_{2P} \\ & & & \\ r_{P1} & r_{P2} & \cdots & r_{PP} \end{bmatrix} \tag{1}$$

Pearson's correlation r_{jk} is a standard linear measure of relation between the variables X_j and X_k, assuming value ranging from -1 to 1, whose calculation is given by Equation 2.

$$r_{jk} = \frac{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2}\sqrt{\sum_{i=1}^{n}(x_{ik} - \bar{x}_k)^2}} \tag{2}$$

Where:
x_{ij} is the i^{th} value measured in the j^{th} variable;
$\bar{x}_j$ is the average of j^{th} variable;
x_{ik} is the i^{th} value measured in the k^{th} variable;
$\bar{x}_k$ is the average of k^{th} variable;
n is the number of samples.

Table IV presents how to interpreting the correlation values. In this research, the variables which presented correlation values less than 0.3 with all the others variables were removed from the dataset. They do not contribute to the construction of information contained between the variables of the data.

Table IV. Interpretation of correlation values (Callegari-Jaques 2003).

Correlation value	Description
0.1 to 0.3	Weak linear correlation
0.3 to 0.5	Moderate linear correlation
0.5 to 0.7	Strong linear correlation
0.7 to 1.0	Very strong linear correlation

Application of Bartlett's Test

Bartlett's Test (Bartlett 1951) is a multivariate statistical test applied to verifying the presence of significant correlation between all the variables and its application is a requirement for applying PCA or another technique of multivariate data analysis. It consists of a comparison between the correlation matrix and the identity matrix, see Equation 3.

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1P} \\ r_{21} & r_{22} & \cdots & r_{2P} \\ & & & \\ r_{P1} & r_{P2} & \cdots & r_{PP} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{3}$$

Null hypothesis of the test consists of the equality of matrices and the test statistics is given by Equation 4.

$$\chi^2 = -\left[(n-1) - \frac{2p+5}{6}\right]\ln(\det(R)) \tag{4}$$

Where:
n is the sample size;
p is the number of variables;
det(R) is the determinant of the correlation matrix.

Test statistics follows a chi-square distribution (χ^2) with degrees of freedom equal to $df = [p(p-1)]/2$. Considering 95% confidence level, a p value less than 0.05 indicates significant correlation between the variables. The test was applied to the dataset after removing the variables, which presented weak linear correlation (< 0.30) with all the other variables.

Application of Principal Component Analysis (PCA)

PCA via correlation matrix was carried out in the dataset after application of Bartlett's Test. It consists of an interdependent multivariate technique, whose purpose is to explaining the variance and covariance of a dataset by means of linear combinations of original variables (Hair et al. 2018, Hottelling 1933, Mingoti 2013).

Considering p independent variables denoted by the vector $X = (x_1, x_2, \dots, x_p)$, the vector of standardized variables $Z = (z_1, z_2, \dots, z_p)$ and the correlation matrix R_{pxp}. The i^{th} principal component (PC_i) is given by Equation 5.

$$\mathrm{PC}_i = e_{i1}z_1 + e_{i2}z_2 + \dots + e_{ip}z_p \tag{5}$$

Where $e_i = (e_{i1}, e_{i2}, \dots, e_{ip})$ is the i^{th} eigenvector of matrix R_{pxp}.

Each principal component is capable of explaining a percentage of the original data. The variability explained by each principal component is given by the proportion of the eigenvalue correspondent to the eigenvector used to build the linear combination, see Equation 6.

$$\frac{\mathrm{Var}[Y_i]}{\mathrm{Var}[X]} = \frac{\lambda_i}{\sum_{i=1}^{p} \lambda_i} \tag{6}$$

Where $Var[Y_i]/Var[X]$ is the proportion of the variance explained by the i^{th} principal component, λ_i is i^{th} eigenvalue associated to the i^{th} eigenvector.

Geometrically, principal component analysis consists of the coordinates of the samples in an axis system obtained by rotating the original axis system, in the direction of maximum variability. The orthogonality between the axis must to be kept.

Analyses were carried out using R software version R-4.3.1 (R Core Team 2015). After performing the analyses, the interdependence between the variables were analyzed and grounded using the literature. Conclusions of the consequences of Fundão dam failure in Barra Longa city were obtained using the information contained in the principal components generated in the analysis.

RESULTS AND DISCUSSION

Exploratory analysis

According to exploratory analysis, 37.5% of the surveyed buildings had contact with the tailing mud in the lot and/or the building ground (P1). This information is correlated to the geographical position of Gesteira and Carmo river. The district is located in one side of the river bank (see Figure 2a), then a majority of buildings is located above the flooding level. They do not had contact with the mud. Considering heavy vehicle traffic, 84.0% of the buildings had in its surroundings (P2). Figure 2 (b) presents the traffic route of these vehicles, which coincides with the river bank and the location of the damaged buildings, respectively.

A percentage of 69.7% of the building pathologies are endogenous and 30.3% are exogenous (P3). During technical survey, the identified architecture of the district was characterized by a combination of traditional constructive techniques (the city born in colonial period) and contemporary techniques. Moreover, the self-construction phenomenon is the predominant way of construction. It consists of constructive technical passed from generation to generation, which the building owner is the responsible by the construction, often without assistance of qualified professionals (Balthazar 2012). This fact can explain the high percentage of endogenous pathologies. Despite 69.7% of endogenous pathologies, dwellers of affected buildings attest that 90.8% of pathologies worsened due to the event (P5), showing the importance of the event in the increase of the building damages.

Pathologies caused by masonry expansion (P7), thermal variation of the system floor (P9), floor system deflection (P10), support deflection (P11), superior beam deflection (P12), structural deflection (P13), overloading (P14), negative bending in the sills (P16) were found in 5.3% of surveyed buildings. Considering the cracks caused by settlement (P6), 46.7% of the studied buildings presented one or more of this type of cracking. This value is greater than 37.5% of the buildings which have contact with the tailing mud. It can be justified by the high number of buildings with endogenous pathologies (69.7%). Probably, part of the surveyed buildings presented cracking before the dam failure and this cracking could have been worsened by the event.

A percentage of 16.4% of the surveyed buildings presented detachment (P18), 17.1% presented infiltrations (P19), 27.6% presented cracks caused by thermal variations in the masonry (P8) and 47.4% presented cracks in openings (P17). These types of pathologies are endogenous and they also could have been worsened by the event.

Seventeen residences (11.2%) presented a high-risk level (P4). Among them, fourteen residences had heavy vehicle traffic in its surroundings, two residences had contact with the tailing mud and

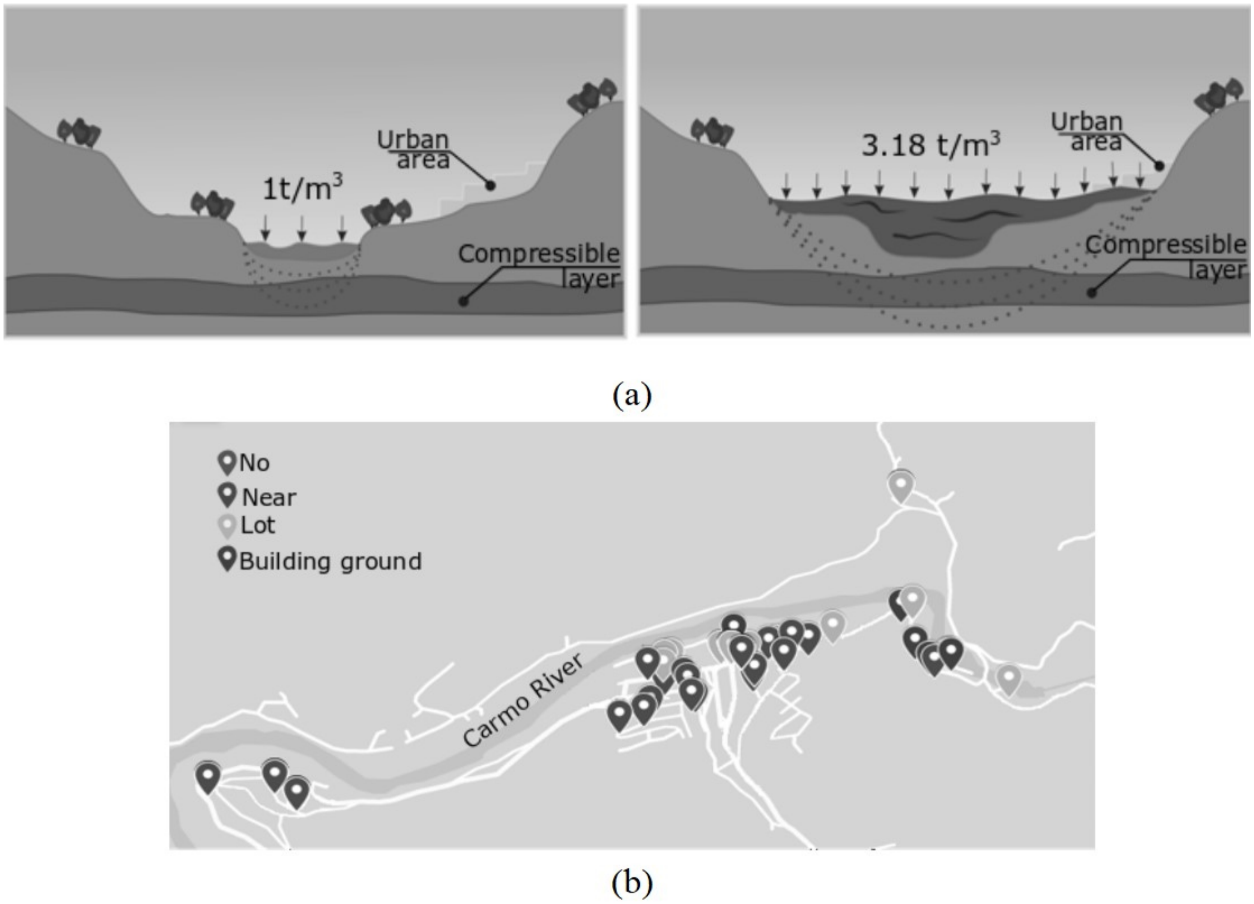

Figure 2. Cross profile of the river before and after the dam failure and location of the majority of the damaged buildings.

one residence without contact with the mud and the heavy vehicle traffic. The last one, presented endogenous pathologies, then its pathologies do not were generated by the dam failure. They were generated by the intrinsic characteristics of its construction process and degree of deterioration.

Figure 3 presents the histograms of category frequency of the surveyed variables.

Data correlation matrix and principal component analysis (PCA)

After data exploratory analysis, the correlation matrix was obtained, see Table V. Variables that presented only weak linear correlations (< 0.3) with all the other variables were removed from the data before application of PCA. The variables number of cracks caused by masonry expansion (P7), number of cracks caused by thermal variations of the floor system (P9), number of cracks caused by floor system deflection (P10), number of cracks caused by support deflection (P11), number of cracks caused by superior beam deflection (P12), number of horizontal cracks caused by overloading (P14), number of vertical cracks caused by overloading (P15), number of infiltration (P19) were removed from the analysis. PCA analysis identify the information contained in the interdependence between variables. Variables without correlation with others will not present independence with others and any information will be found. The analyses were carried out in the standardized data.

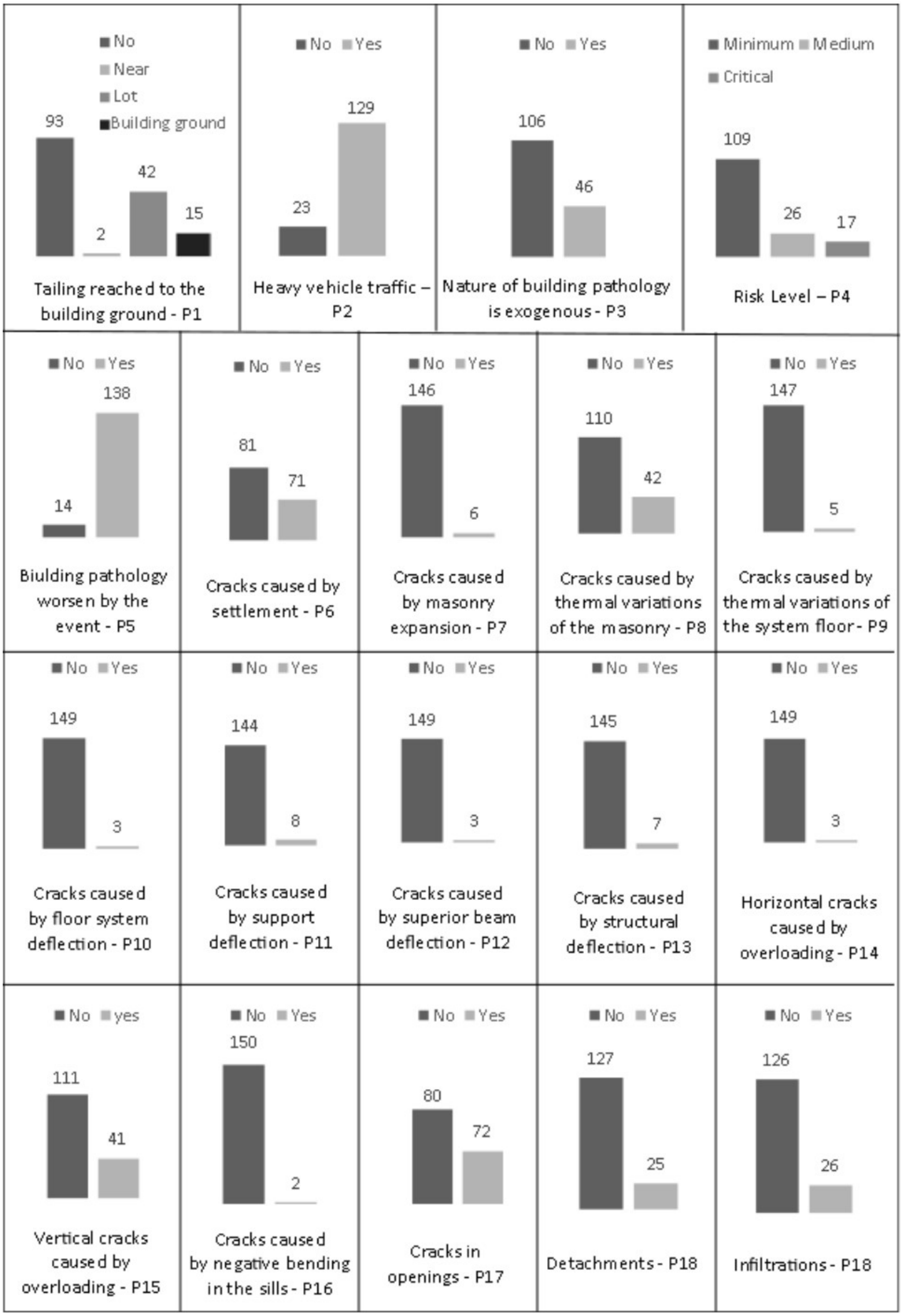

Figure 3. Analysis of the category frequency of each variable in the dataset.

Table V. Correlation matrix.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19
P1	11.00	-0.15	0.58	0.06	0.24	0.23	-0.14	0.06	-0.03	-0.10	0.05	0.13	0.16	0.17	-0.09	0.09	0.05	-0.08	0.11
P2	-0.15	11.00	-0.04	0.00	0.63	0.05	0.08	0.12	0.07	-0.13	0.06	0.00	0.07	-0.13	0.04	0.04	0.13	0.11	-0.11
P3	0.58	-0.04	11.00	0.12	0.21	0.48	-0.12	0.25	0.14	-0.09	-0.04	0.18	0.24	-0.02	-0.11	0.14	0.01	0.11	0.16
P4	0.06	0.00	0.12	11.00	0.08	0.44	-0.05	0.01	-0.06	-0.08	0.04	-0.07	-0.05	0.07	-0.11	0.06	0.12	0.29	-0.06
P5	0.24	0.63	0.21	0.08	11.00	0.14	0.02	0.06	0.06	-0.19	0.04	0.04	0.05	0.04	-0.01	0.03	0.05	0.09	-0.08
P6	0.23	0.05	0.48	0.44	0.14	11.00	-0.10	-0.01	0.00	-0.04	-0.07	0.01	-0.07	-0.04	-0.15	0.25	-0.03	0.17	-0.05
P7	-0.14	0.08	-0.12	-0.05	0.02	-0.10	11.00	-0.05	-0.03	-0.02	-0.03	-0.02	-0.03	-0.02	-0.07	-0.02	0.05	0.00	0.05
P8	0.06	0.12	0.25	0.01	0.06	-0.01	-0.05	11.00	0.01	-0.06	0.00	0.18	0.18	-0.06	-0.13	0.37	0.17	0.42	0.28
P9	-0.03	0.07	0.14	-0.06	0.06	0.00	-0.03	0.01	11.00	-0.02	-0.02	-0.02	-0.03	-0.02	-0.01	-0.02	0.04	-0.03	-0.03
P10	-0.10	-0.13	-0.09	-0.08	-0.19	-0.04	-0.02	-0.06	-0.02	11.00	-0.02	-0.02	-0.02	-0.06	-0.01	-0.08	-0.04	-0.05	0.05
P11	0.05	0.06	-0.04	0.04	0.04	-0.07	-0.03	0.00	-0.02	-0.02	11.00	-0.02	0.00	-0.02	-0.06	-0.01	-0.06	-0.03	-0.05
P12	0.13	0.00	0.18	-0.07	0.04	0.01	-0.02	0.18	-0.02	-0.02	-0.02	11.00	-0.02	-0.02	-0.05	-0.01	-0.04	-0.03	0.03
P13	0.16	0.07	0.24	-0.05	0.05	-0.07	-0.03	0.18	-0.03	-0.02	0.00	-0.02	11.00	0.08	0.08	0.01	0.16	0.04	-0.01
P14	0.17	-0.13	-0.02	0.07	0.04	-0.04	-0.02	-0.06	-0.02	-0.06	-0.02	-0.02	0.08	11.00	-0.05	0.13	-0.06	-0.04	0.17
P15	-0.09	0.04	-0.11	-0.11	-0.01	-0.15	-0.07	-0.13	-0.01	-0.01	-0.06	-0.05	0.08	-0.05	11.00	-0.10	-0.11	-0.14	-0.04

Bartlett's Test was carried out with the objective of verifying the suitability of the data for application of PCA. Table VI presents the test result. Once, the p value is less than 0.05, it is possible to affirm that the data is suitable.

Table VI. Bartlett's Test.

χ^2	df	p value
428.01	55	3.57×10^{-59}

Principal component analysis (PCA) was performed. Table VII presents the proportion of the explained variance of the original data by each generated principal component. Kaiser's criterion was used to determining the number of retained principal components. This criterion consist of keeping in the analysis the principal components with eigenvalue greater than 1.0 (Kaiser 1960). The first five principal components (PC_1 to PC_5) were kept and they are capable of explaining 75% of the original data information, see Table VII.

Table VII. Proportion of the variance explained of the original data by each generated principal component.

	PC_1	PC_2	PC_3	PC_4	PC_5	PC_6
Proportion of explained variance	0.23	0.15	0.14	0.13	0.10	0.07
Cumulative proportion	0.23	0.38	0.52	0.65	0.75	0.82
	PC_7	PC_8	PC_9	PC_{10}	PC_{11}	
Proportion of explained variance	0.06	0.04	0.03	0.02	0.02	
Cumulative proportion	0.88	0.92	0.96	0.98	1.00	

The first five principal components are given by Equations 7, 8, 9, 10 and 11, where Z_i is the i^{th} standardized variable.

$$\begin{aligned} \mathrm{PC}_1 =& 0.3z_1 + 0.2z_2 + 0.4z_3 + 0.3z_4 + 0.3z_5 \\ &+ 0.4z_6 + 0.3z_8 + 0.2z_{13} + 0.3z_{16} + 0.2z_{17} + 0.3z_{18} \end{aligned} \tag{7}$$

$$\begin{aligned} \mathrm{PC}_2 =& 0.4z_1 - 0.4z_2 + 0.4z_3 + 0.1z_4 - 0.1z_5 \\ &+ 0.4z_6 - 0.3z_8 - 0.2z_{13} - 0.1z_{16} - 0.4z_{17} - 0.3z_{18} \end{aligned} \tag{8}$$

$$\begin{aligned} \mathrm{PC}_3 =& 0.1z_1 + 0.5z_2 + 0.6z_5 - 0.3z_8 - 0.3z_{16} - 0.2z_{17} \\ &- 0.3z_{18} \end{aligned} \tag{9}$$

$$\begin{aligned} \mathrm{PC}_4 =& 0.3z_1 - 0.2z_2 + 0.2z_3 - 0.4z_4 - 0.3z_6 - 0.6z_{13} \\ &- 0.1z_{16} + 0.3z_{17} - 0.4z_{18} \end{aligned} \tag{10}$$

$$\begin{aligned} \mathrm{PC}_5 =& 0.1z_1 + 0.2z_3 - 0.6z_4 - 0.2z_6 + 0.4z_8 - 0.3z_{13} \\ &+ 0.3z_{16} - 0.5z_{17} \end{aligned} \tag{11}$$

The five retained principal components were interpreted and interdependence variable information were obtained. The analysis of the two first principal components (PC_1 and PC_2) are presented in Figure 4. The last three principal components (PC_3, PC_4 and PC_5) do not presented interpretability.

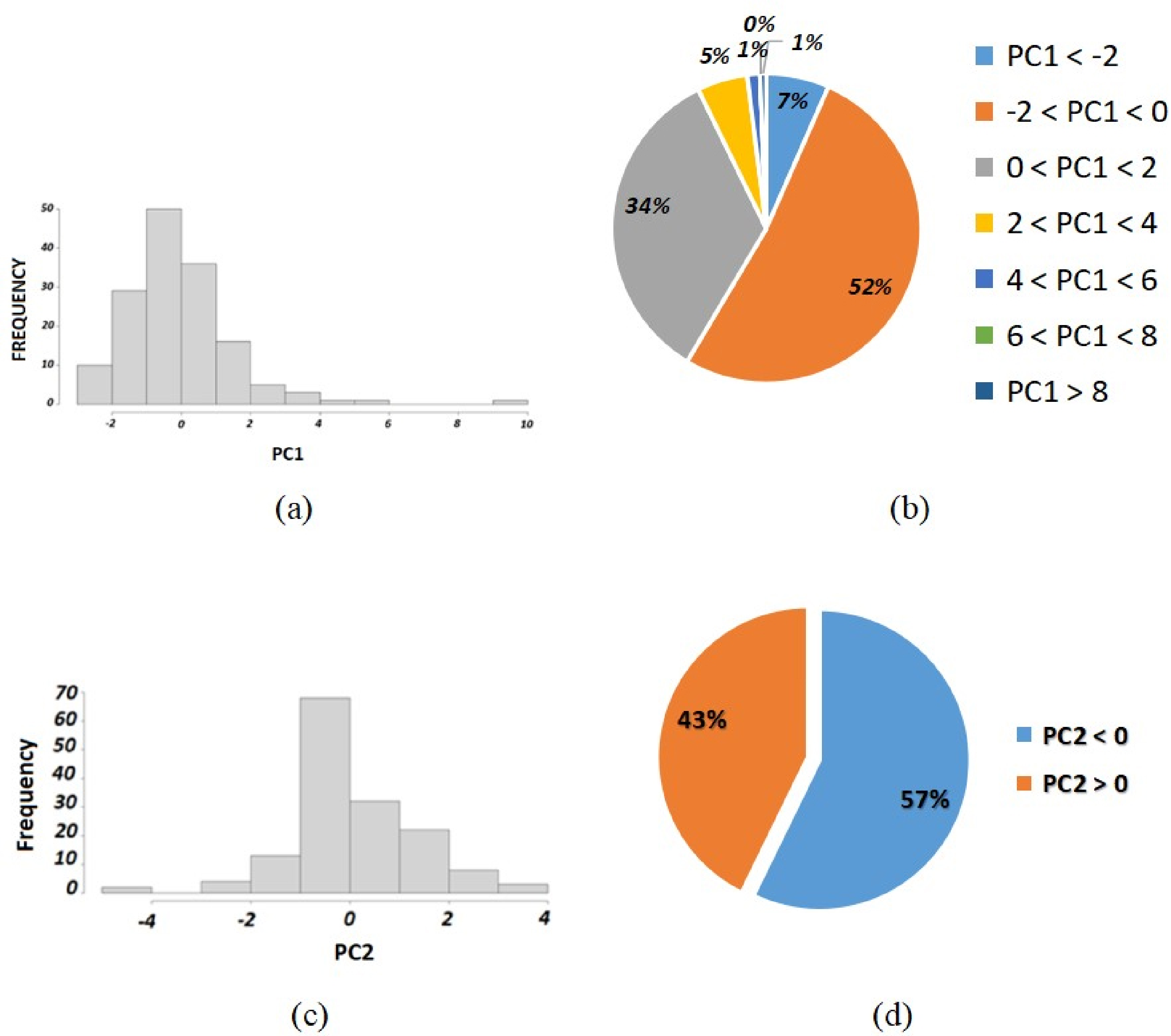

Figure 4. PC1 and PC2 results.

Principal component 1 (PC_1) consists of a Global Index. Buildings with highest values of PC_1 are the ones most damaged by the dam failure. The buildings with lowest values of PC_1 were less or not even damaged by the event. Figure 4a-b the histogram and the pie chart of PC_1 values for the 152 buildings.

Most of buildings presented values of PC_1 in a range of -2 to 2, where 52% of constructions presents value of PC_1 between -2 to 0 and 34% of constructions has values between 0 and 2. Table VIII shows the modal values of variables P1 to P5 for the PC_1 ranges defined in the histogram of Figure 4a.

Table VIII. Variable modes of the ranges of PC1 values.

	Tailing reached to the building ground (P1)	Heavy vehicle traffic (P2)	Nature of building pathology is exogenous (P3)	Risk level (P4)	Building pathology worsen by the event (P5)
$PC_1 < -2$	No	No	No	Minimum	No
$-2 < PC_1 < 0$	No	Yes	No	Minimum	Yes
$0 < PC_1 < 2$	Lot	Yes	Yes	Minimum	Yes
$PC_1 > 2$	Lot	Yes	Yes	Critical	Yes

Buildings with value of PC_1 smaller than -2 have not contact with the tailing mud and heavy vehicle traffic. The pathologies of these buildings are endogenous, were not worsened by the failure dam event and present a minimum risk level. They correspond to 7% of the 152 surveyed buildings. The buildings with PC_1 between -2 and 0 also have not contact with the tailing dam, they are also endogenous and present a minimum risk level. Nonetheless, they had contact with heavy vehicle traffic and this event could worsen its preexisting pathologies. They correspond to 52% of the surveyed buildings.

Buildings with PC_1 value between 0 and 2 present exogenous pathologies and they have been worsened by the event. These buildings had contact with the tailing mud at least on the lot ground and contact with heavy vehicle traffic, but they present a minimum risk level. They correspond to 34% of the 152 buildings. Buildings with PC_1 greater than 2 had contact with the tailing mud at least in the lot, and contact with heavy vehicle traffic. Its pathologies are exogenous and they were worsened by the event. In this case, these buildings have a critical level risk ant they correspond to 7% of the data.

Table IX presents the variables and range of PC_1 of the buildings where the tailing mud reached the building ground.

Building where the tailing mud reached the building ground presented values of PC_1 in a range of -2 to 2. These buildings mostly presented minimum risk level, had contact with heavy vehicle traffic and had exogenous pathologies. Besides, building pathology was worsen by the event in all cases, eight buildings had contact with heavy vehicle traffic and seven had no contact. Buildings with PC_1 between -2 and 0 had no contact with heavy vehicle and the nature of building pathology is endogenous. This information shows that the mud reach or not reach the building ground is not an ensure of worst building performance and the heavy vehicle traffic must to be carefully considered and analyzed.

Principal component 2 (PC_2) consists of an index, which determines the type of building pathology and its cause, see Equation (8) and Table X. In this principal component, P1 (tailing reached to the building ground), P3 (nature of building pathology is exogenous), P4 (risk level) e P6 (number of cracks caused by settlement) has positive coefficients. Observing Table II, the closer the mud is to the building, the greater the value of variable P1. So, exogenous pathologies will have greater values of P2 and the greater the degree of risk, the greater the value of variable P4. Then, the closer the mud is to the building, the greater the probability of generating greater overloads to the building, which can cause greater cracks by settlement, which are exogenous pathologies and, finally, increasing the degree of risk of the building. Wherefore, positive values of principal compnent 2 respresents pathological problems related to the direct contact with the mud (P1).

Table IX. Buildings that have contact with the tailing mud.

ID	Tailing the building reached to ground (P1)	Heavy vehicle traffic (P2)	Nature of building pathology is exogenous (P3)	Risk level (P4)	Building pathology worsen by the event (P5)	Range of PC1
93	Building ground	Yes	Yes	Minimum	Yes	$0 < PC_1 < 2$
90	Building ground	Yes	Yes	Minimum	Yes	$0 < PC_1 < 2$
91	Building ground	Yes	Yes	Minimum	Yes	$0 < PC_1 < 2$
142	Building ground	No	Yes	Minimum	Yes	$0 < PC_1 < 2$
5	Building ground	Yes	No	Regular	Yes	$0 < PC_1 < 2$
8	Building ground	Yes	Yes	Minimum	Yes	$0 < PC_1 < 2$
21	Building ground	Yes	Yes	Minimum	Yes	$0 < PC_1 < 2$
77	Building ground	No	No	Minimum	Yes	$-2 < PC_1 < 0$
78	Building ground	Yes	No	Minimum	Yes	$0 < PC_1 < 2$
2	Building ground	No	Yes	Minimum	Yes	$0 < PC_1 < 2$
57	Building ground	No	Yes	Minimum	Yes	$0 < PC_1 < 2$
66	Building ground	No	Yes	Minimum	Yes	$0 < PC_1 < 2$
69	Building ground	No	No	Minimum	Yes	$-2 < PC_1 < 0$
94	Building ground	Yes	Yes	Minimum	Yes	$0 < PC_1 < 2$
95	Building ground	No	Yes	Minimum	Yes	$0 < PC_1 < 2$

Variables P2 (heavy vehicle traffic), P5 (building pathology worsen by the event), P8 (number of cracks caused by thermal variations of the masonry), P13 (number of cracks caused by structural deflection), P16 (number of cracks caused by negative bending in the sills), P17 (number of cracks in openings - doors, windows, vents) e P18 (number of detachments) have negative coefficients in principal component 2. Observing Table II, P1 is equal to 0 when the building do not have contact with the mud, P3 is equal to 0 when the pathology is endogenous and P4 have lower values for lower risk level. Then, negative values of principal component 2 represent the building pathologies that not are related to the direct contact with the mud (P1), but are related to heavy vehicle traffic (P2). The pathologies measured in P8, P13, P17 e P18 variables are fissures, which have endogenous nature and the vibrations induced by heavy vehicle traffic can cause increase of these fissures (this subject will be more discussed further in the article).

Still evaluating principal component 2 (PC_2), P5 (building pathologies worsen by the event) has a negative coefficient with smallest magnitude (value of 0.1 – see equation 8). Regardless of the origin of the pathology being due to direct contact with the mud (P1) or heavy vehicles traffic (P2), both were capable of worsen the pathologies, justifying this lower magnitude. The number of buildings with contact with heavy vehicle traffic (129 buildings) is much higher than the number of buildings that had direct contact with the mud (15 buildings), see Figure 3. Then, P5 has more relation with P2 than P1, due to the greater number of recurrence, thus having its negative coefficient. Finally, P5 has relation with P1 (positive coefficient) and P2 (negative coefficient) and, therefore, it has a lower absolute value in principal component 2.

Table X. Variables and the signal of its coefficients in PC2.

Variables with positive coefficients	Variables with negative coefficients
Tailing reached to the building ground (P1)	Heavy vehicle traffic (P2)
Nature of building pathology is exogenous (P3)	Building pathologies worsen by the event (P5)
Risk level (P4) is exogenous (P3)	Number of cracks caused by thermal variations of the masonry (P8)
Number of cracks caused by settlement (P6)	Number of cracks caused by structural deflection (P13)
	Number of cracks caused by negative bending in the sills (P16)
	Number of cracks in openings (P17)
	Number of detachments (P18)

Therefore, the buildings with positive values of PC_2 were majority damaged by the tailing proximity (Tailing reached to the building ground - P1) and the associated pathology was cracks caused by settlement (P6) and buildings with negative values of PC_2 were majority damaged by the heavy vehicle traffic (P2) and the associated pathologies were cracks caused by thermal variations of the masonry (P8), cracks caused by structural deflection (P13), cracks caused by negative bending in the sills (P16), cracks in openings (P17) and detachments (P18). Figure 4c-d present the histogram and pie chart of PC2 values for the 152 buildings.

Contact with the tailing mud cased damage to 43% of the buildings and 57% of the buildings were damaged by heavy vehicle traffic, see Figure 4c-d. PC_2 opposes the contact with the tailing mud and the pathologies generated by its overload (settlement) with the heavy vehicle traffic and the pathologies generated by vibration. Pathologies caused by settlement are correlated to driving forces greater than the soil strength capacity in the foundations and, it is knowing that, the driving forces are increased by the extra overload of tailing mud. The heavy vehicle traffic (trucks, heavy machinery and compaction equipment) has been could generate vibrations on buildings and, depending on the vibration level, fissures could manifest (Andrews et al. 2013).

Figure 5 shows the graph of PC_1 versus PC_2, knowing as a Biplot graph, where is possible to evaluate the variable interdependence discussed in the topics above.

Induced vibrations in buildings caused by heavy vehicle traffic

The vibration level generated by vehicle traffic depends on the vehicle weight and road conditions. Irregular pavement roads, such as prismatic blocks or damaged roads, are those responsible to generate greater vibratory energy (Hunaidi & Tremblay 1997). Considering the type of the vehicle, according to California Department of Transportation (Andrews et al. 2013), the vibration caused by construction and maintenance equipment directly influence in the ground and, consequently, in foundation buildings.

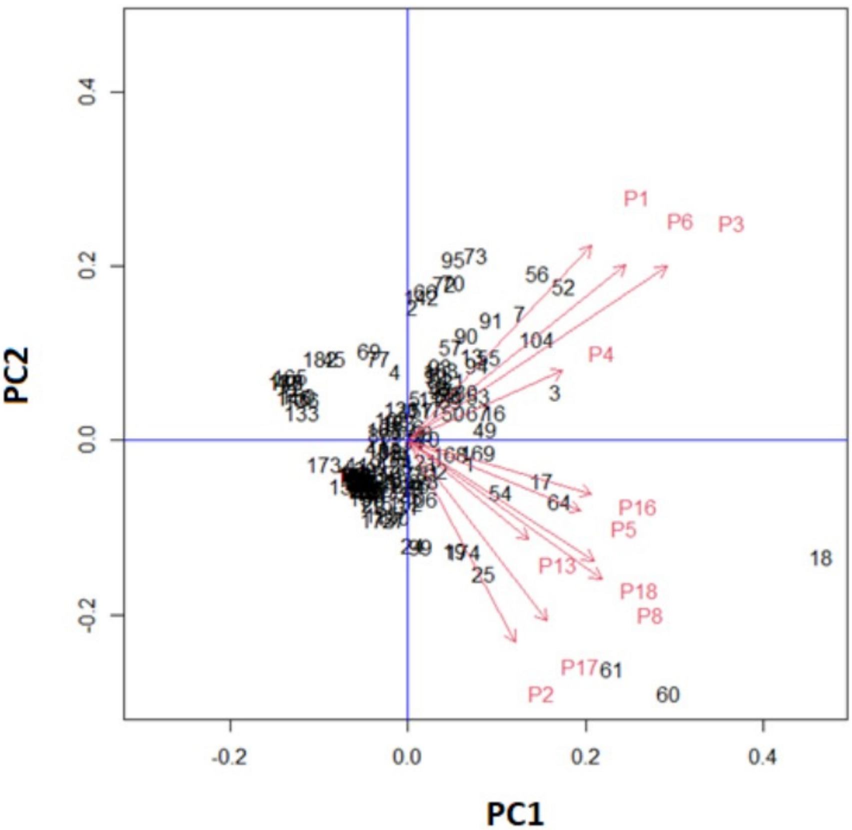

Figure 5. Biplot of PC1 versus PC2.

Peak Particle Velocity (PPV) transmitted from vehicle to building is an index used to verify the influence of induced vibration of heavy vehicle traffic in urban roads given by Equation 12 (Andrews et al. 2013). This index is compared with the allowable PPV limit values of buildings, see Table XI.

$$\mathrm{PPV}_{\text{heavy vehicle traffic}} = \mathrm{PPV}_{\mathrm{Ref}}\left(\frac{25}{D}\right)^{n}(in/sec) \tag{12}$$

where:
PPV_{Ref} is the reference PPV at 25 ft., see Table XI;
D is the distance from equipment to the receiver (building) (ft) and
n is the value related to the attenuation rate through ground.

According to California Department of Transportation (Andrews et al. 2013), the suggested value of n is equal to 1.1, considering the studied site. The equipment vibration originates near to ground surface, then the determination of n is not influenced by the soil type.

Considering the values of Table XII and the Equation 12, PPV of the heavy vehicle traffic used to clean up the tailing mud and reconstruct the district was calculated. A distance D equal to 2.5 meters (distance of the equipment gravity of center to the building facade) was considered. The obtained results are presented in Table XIII.

Table XI. Guideline Vibration Damage Potential Threshold Criteria - Maximum PPV (in/sec) (Andrews et al. 2013).

Structure and Condition	Transient Sources	Continuous/Frequent Intermittent Sources
Extremely fragile historic buildings, ruins, ancient monuments	0.12	0.08
Fragile buildings	0.2	0.1
Historic and some old buildings	0.5	0.25
Older residential structures	0.5	0.3
New residential structures	1.0	0.5
Modern industrial/commercial buildings	2.0	0.5

Table XII. Vibration Source Amplitudes for Construction Equipment (Andrews et al. 2013).

Equipament	Reference PPV at 25 ft. (in/sec)
Vibratory roller	0.210
Large bulldozer	0.089
Caisson drilling	0.089
Loaded trucks	0.076
Jackhammer	0.035
Small bulldozer	0.003
Crack-and-seat operations	2.4

Comparing the obtained results of PPV (Table XIII) with the allowable PPV limit values of buildings (Table XI), considering transient sources, it is possible to notice that some PPV generated by the heavy vehicle in Barra Longa city are very close to the limit values from old buildings and others are greater from fragile buildings. Observing on the results, it is possible to conclude that the heavy vehicle traffic may have worsen preexisting pathologies of the buildings, considering that buildings from Gesteira are old, constructed using self-construction techniques.

CONCLUSION

Principal component analysis was capable of explain the interdependence of data variable and allows conditions of understand the consequences caused by Fundão dam failure in Gesteira district, city of Barra Longa. Through the analysis, it was possible to defining the existence of relationship between the heavy vehicle traffic with the appearance of several pathologies related to fissures. Besides, the contact with the tailing mud, according to the analysis, caused pathologies related to settlement. These conclusions are expected of technical point of view, but the analysis provides way of determining the coverage of these consequences in the studied location. This type of conclusion would not possible through univariate statistical analysis.

According to the obtained results of PC_1, 7% of analyzed building were very damaged by the dam failure, where new pathologies with critical level of risk tended to appear on them. 34% of the buildings

Table XIII. PPV results, considering the distance D equal to 2.5 meters (8,2 ft).

Equipment	Reference PPV at 25 ft. (in/sec)
Vibratory roller	0.716
Large bulldozer	0.303
Loaded trucks	0.259
Small bulldozer	0.010

were moderately damage, where new pathologies with minimum risk level tended to appear on them. 52% of the studied buildings were slightly damaged, where preexisting pathologies tended to present minimum risk of level and were worsen by the failure event. Finally, 7% of building were not damaged by the event, where the pathologies tended to be endogenous and were not worsen by the event. According to obtained results of PC_2, 43% of building were damaged by the contact with the tailing mud and 57% of them were damaged by the heavy vehicle traffic, which were validated by the results of PPV analysis.

Observing the obtained results, it possible to conclude that the heavy vehicle traffic associated to reconstruction of Gesteira infrastructure caused more damage to the studied buildings than the contact with the tailing mud. Is result is related to the reach of the tailing mud, which were limited to the buildings located next to the river (9.9% of the studied buildings), whereas the heavy vehicle traffic occurred along several streets of the district. In case of future events with the same characteristics, the concern with the vibration level of heavy vehicle traffic involved in the infrastructure reconstruction must to be seriously considered, because they can be more harmful for the building than the mud flood.

Acknowledgments

The authors wish to thank Universidade do Estado de Minas Gerais, Universidade Federal de Ouro Preto, Grupo de Pesquisa em Ciência de Dados aplicada à Engenharia (CIDENG), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG APQ00501-21) to support this study and Alteto Engenharia to provide the data used in this research.

REFERENCES

ADIANSYAH JS, ROSANO M, VINK S & KEIR G. 2015. A framework for a sustainable approach to mine tailings management: disposal strategies. J Clen Prod 108: 1050-1062. DOI: https://doi.org/10.1016/j.jclepro.2015.07.139.

ALI Y, IRFAN M, BULLER AS, KHAN HA & GUL HMF. 2019. A binary logistic model for predicting the tertiary stage of permanent deformation of conventional asphalt concrete mixtures. Constr Build Mater 227: 116608. DOI: https://doi.org/10.1016/j.conbuildmat.2019.07.334.

ANDREWS J, BUEHLER D, GILL H & BENDER WL. 2013. Transportation and construction vibration. [S. l.]: California Depatment of Transportation.

BALTHAZAR RDS. 2012. A permanência da autoconstrução: um estudo de sua prática no município de Vargem Grande Paulista. Master’s thesis. Universidade do Estado de São Paulo. URL https://www.teses.usp.br/teses/disponiveis/16/16137/tde-02072012-132335/pt-br.php.

BARTLETT MS. 1951. The effect of standardization on a chi-square approximation in factor analysis. Biometrika 38: 337-344.

CALLEGARI-JAQUES SM. 2003. Bioestatística: princípios e aplicações. São Paulo: Artemed, 264 p.

CARMO F ET AL. 2017. Fundão tailings dam failures: the environment tragedy of the largest technological disaster of Brazilian mining in global context. Perspect Ecol Conserv 15(3): 145-151. DOI: https://doi.org/10.1016/j.pecon.2017.06.002.

COELHO M, SENA-CRUZ J, NEVES L, PEREIRA M, CORTEZ P & MIRANDA T. 2016. Using data mining algorithms to predict the bond strength of NSM FRP systems in concrete. Constr Build Mater 126: 484-495. DOI: https://doi.org/10.1016/j.conbuildmat.2016.09.048..

COIMBRA K, ALCÂNTARA E & FILHO CS. 2020. Possible contamination of the Abrolhos reefs by Fundao dam tailings, Brazil – New constraints based on satellite data. Sci Total Environ 733: 138101. DOI: https://doi.org/10.1016/j.scitotenv.2020.138101.

GUERRA M, TEANEY B, MOUNT BJ, ASUNSKIS DJ, JORDAN BT, BARKER RJ, SANTOS EE & SCHAEFER C. 2017. Post-catastrophe Analysis of the Fundão Tailings Dam Failure in the Doce River System, Southeast Brazil: Potentially Toxic Elements in Affected Soils. Water Air Soil Pollut 228: 1-12. DOI: https://doi-org.ez28.periodicos.capes.gov.br/10.1007/s11270-017-3430-5.

HAIR JF, BLACK WC, BABIN BJ & ANDERSON RE. 2018. Multivariate Data Analysis. Georgia: Kennesaw State University, 833 p.

HOTTELLING H. 1933. Analysis of a complex of statistical variables into principal components. J Educ Psychol 24(6): 417-441. URL http://dx.doi.org/10.1037/h0071325.

HUNAIDI O & TREMBLAY M. 1997. Traffic induced building vibrations in Montréal. Canad J Civil Eng 24(5): 736-753. DOI: https://doi.org/10.1139/l97-023.

IBAPE. 2012. Norma De Inspeção Predial Nacional. URL http://www.ibape-nacional.com.br.

KAISER HF. 1960. The application of electronic computers to factor analysis. Educ Psychol Meas 20: 141-151.

KANG MC, YOO DY & GUPTA R. 2021. Machine learning-based prediction for compressive and flexural strengths of steel fiber-reinforced concrete. Constru Build Mater 266: 121117. DOI: https://doi.org/10.1016/j.conbuildmat.2020.121117.

MAGALHÃES EF. 2004. Fissuras Em Alvenarias: Configurações Típicas E Levantamento De Incidências No Estado Do Rio Grande Do Sul. Master's thesis. Universidade Federal do Rio Grande do Sul. URL https://lume.ufrgs.br/handle/10183/10135.

MARTA-ALMEIDA M, MENDES R, AMORIM FN, CIRANO M & DIAS JM. 2016. Fundão Dam collapse: Oceanic dispersion of River Doce after the greatest Brazilian environmental accident. Mar Pollut Bull 112(1-2): 359-364. DOI: https://doi.org/10.1016/j.marpolbul.2016.07.039.

MINGOTI SA. 2013. Análise de dados através de métodos de estatística multivariada. Belo Horizonte: Editora UFMG, 297 p.

ORLANDO MD ET AL. 2020. Tracing iron ore tailings in the marine environment: An investigation of the Fundão dam failure. Chemosphere 257: 127184. DOI: https://doi.org/10.1016/j.chemosphere.2020.127184.

PAULO E, PEREIRA C, SANTOS TB & OLIVEIRA RM. 2020. Avaliação das barragens de rejeito brasileiras por meio da análise de agrupamentos k médias. Res Soc Develop 9(9): e731997811-e731997811. DOI: http://dx.doi.org/10.33448/rsd-v9i9.7811.

QI C, CHEN Q, FOURIE A, TANG X, ZHANG Q, DONG X & FENG Y. 2019. Constitutive modelling of cemented paste backfill: A data-mining approach. Constr Build Mater 197: 262-270. DOI: https://doi.org/10.1016/j.conbuildmat.2018.11.142.

QUADRA GR ET AL. 2019. Far-reaching cytogenotoxic effects of mine waste from the Fundão dam disaster in Brazil. Chemosphere 215: 753-757. DOI: https://doi.org/10.1016/j.chemosphere.2018.10.104.

R CORE TEAM. 2015. R: A language and environment for statistical computing. Disponível em: https://www.rproject.org.

RICO M, BENITO G, SALGUEIRO AR, DÍEZ-HERRERO A & PEREIRA HG. 2008. Reported tailings dam failures. A review of the European incidents in the worldwide context. J Hazard Mater 152(2): 846-852. DOI: https://doi.org/10.1016/j.jhazmat.2007.07.050.

SAMARCO. 2020. Rompimento da barragem de Fundão. URL https://www.samarco.com/rompimento-dabarragem-de-fundao/.

SANTOS AEM, LANA MS, CABRAL IE, PEREIRA TM, NAGHADEHI MZ, SILVA D & SANTOS TB. 2019. Evaluation of Rock Slope Stability Conditions Through Discriminant Analysis. Geotech Geol Eng 37: 775-802. DOI: https://doi.org/10.1007/s10706-018-0649-x.

SANTOS TB, LANA MS, PEREIRA TM & CANBULAT I. 2018. Quantitative hazard assessment system (Has-Q) for open pit mine slopes. Int J Min Sci Technol 29(3): 419-427. DOI: https://doi.org/10.1016/j.ijmst.2018.11.005.

SANTOS TB & OLIVEIRA R. 2021. Failure risk of brazilian tailings dams: a data mining approach. An Acad Bras Cienc 93: e20201242. DOI: https://doi.org/10.1590/0001-3765202120201242.

WATT D. 2007. Building Pathology: Principles and Practice. Hoboken: J Wiley & Sons, 320 p.

YANG E, TANG Y, LI L, YAN W, HUANG B & QIU Y. 2021. Research on the recurrent neural network-based fatigue damage model of asphalt binder and the finite element analysis development. Constr Build Mater 267: 121761. URL https://doi.org/10.1016/j.conbuildmat.2020.121761.

How to cite

SILVA JUNIOR LA & SANTOS TB. 2023. Building pathologies caused by failure of Fundão Tailing Dam: A principal component analysis aproach. An Acad Bras Cienc 95: e20220458. DOI 10.1590/0001-3765202320220458.

Manuscript received on May 24, 2022;
accepted for publication on September 27, 2022

LADIR ANTONIO DA SILVA JUNIOR[1]
https://orcid.org/0000-0002-0193-7060

TATIANA B. DOS SANTOS[2]
https://orcid.org/0000-0001-5484-6675

[1]Universidade do Estado de Minas Gerais, Departamento de Engenharia Civil, Av. Brasília, 1304, Baú, 35930-314 João Monlevade, MG, Brazil
[2]Universidade Federal de Ouro Preto, Departamento de Mineração da Escola de Minas, Campus Universitário, Rua Henri Gorceix, 241-321, Morro do Cruzeiro, 35400-000, Ouro Preto, MG, Brazil

Correspondence to: **Ladir Antonio da Silva Junior**

E-mail: ladir.jr@uemg.br

Author contributions

Ladir Antonio da Silva Junior: Conception or design of the work; Data analysis and interpretation; Drafting the article. Tatiana Barreto dos Santos: Conception or design of the work; Data analysis and interpretation; Drafting the article.