

Cluster analysis for slope geotechnical prioritization of intervention for the Estrada de Ferro Vitória-Minas

<http://dx.doi.org/10.1590/0370-44672017710173>

Denise de Fátima Santos da Silva

Mestre

Universidade Federal de Ouro Preto – UFOP

Escola de Minas

Núcleo de Geotecnia

Ouro Preto – Minas Gerais - Brasil

denisefsilva61@gmail.com

Allan Erlichman Medeiros Santos

<http://orcid.org/0000-0003-1241-8423>

Doutorando

Universidade Federal de Ouro Preto – UFOP

Escola de Minas

Departamento de Engenharia de Minas

Professor do Instituto Federal Goiano - Catalão - Goiás

Ouro Preto – Minas Gerais - Brasil

allanboni@hotmail.com

Bruno Trindade Ferreira

Mestrando

Universidade Federal de Ouro Preto – UFOP

Escola de Minas

Departamento de Engenharia de Minas

Ouro Preto – Minas Gerais - Brasil

btrindade.ferreira@gmail.com

Tiago Martins Pereira

Professor-Adjunto

Universidade Federal de Ouro Preto - UFOP

Instituto de Ciências Exatas e Biológicas – ICEB

Departamento de Estatística

Ouro Preto – Minas Gerais - Brasil

tiago.martin@gmail.com

Rosyelle Cristina Corteletti

Professora Colaboradora

Universidade Federal de Ouro Preto – UFOP

Escola de Minas

Núcleo de Geotecnia

Ouro Preto - Minas Gerais - Brasil

rosyellecorteletti@gmail.com

Abstract

This article proposes the geotechnical prioritization of intervention of slopes with landslide scars for the Estrada de Ferro Vitória-Minas by cluster analysis and also the proposition of a relationship between area and volume in landslide scars. Cluster definition helps the decision-making associated to containment measures, mapping and study of landslides for the Estrada de Ferro Vitória-Minas. The database is composed of the variables: slope's height, inclination, scar area and scar volume. The distance measure used was Gower's index, with Ward's methods to build the clusters. Eight characteristic groups were identified. It was possible to identify stretches that need attention in relation to the propensity of landslides, such as Group 7, stretches 362+600, 093+xxxE and 419+000. Group 7 presented high values for the scarred area and volume, such as maximum area $9.75 \times 10^4 \text{ m}^2$ and minimum area $7.49 \times 10^4 \text{ m}^2$, and maximum volume $9.20 \times 10^5 \text{ m}^3$ and minimum volume $4.08 \times 10^5 \text{ m}^3$. Group 7 presented high ranges for slope height and inclination. The set of results about Group 7 can be interpreted as stretches with a predisposition for landslides. In relation to intervention measures, Group 7 presents the sections with priority. The relationship between area and volume of landslide scars obtained by the research was compared with the relationships established in literature.

Keywords: geotechnical prioritization of intervention of railway slopes, Estrada de Ferro Vitória-Minas, cluster analysis, multivariate statistics applied to geotechnical database, landslide scars.

1. Introduction

Predicting the area and volume of rock mass in landslides is a matter of extreme complexity. Some cases allowed equating landslides in different parts of the world. The empirical relationships allow estimat-

ing parameters of landslides as a function of characteristics that can be determined, such as slope inclination and height. The empirical relationships allow the assessment and delimitation of areas of risk, allowing

conviviality with natural phenomena. In addition, they also allow the preparation of programs to prevent, alert, and implement containment construction (Polanco, 2010).

This study was developed taking as

reference the railroad line of the Estrada de Ferro Vitória-Minas (EFVM), Vale S.A., which is inserted in the states of Minas Gerais and Espírito Santo. The railroad has 929 km of extension and the trajectory of the trunk line is 540 km. The geology that covers the study area, includes dominant lithologic classes in the trunk line of the EFVM, such as granites, gneisses, gneisses with mafic intrusives, metasedimentary rocks with presence of granitic intrusions, shales and gneisses and unconsolidated sediments.

In 1997, the company Vale S.A.

received, through a contract signed with Brazil's government, the concession for the complete operation of its rail transportation services. EFVM handled 119.2 million tons of ore, in addition to other cargo transported to other companies, such as coal and agricultural products. In relation to passenger transport, EFVM transported almost one million people in 2014. (Vale, 2017).

The present article proposes to identify clusters in the slopes with scars in a specific section of the Estrada de Ferro Vitória Minas; this clustering allows to

interpret the sectorization of the railroad and prioritization of intervention on slopes with landslides. The prioritization will be based on slope height, inclination, scar area and volume variables analyzed by cluster analysis. The cluster's definition helps the decision-making associated to containment measures, mapping and study of landslides for the Estrada de Ferro Vitória-Minas. In this context this work can contribute to studies about landslides in Brazil, through the proposition of empirical relationships for landslide control estimates.

2. Materials and methods

The database used in this article was compiled from the data of Gomes' work (2014) in conjunction with image analysis of the Google Earth Pro

software and field work in the studied stretches of the railway.

The variables collected on the slopes are slope's height (HT) and

inclination (IT). These variables are qualitative, classified by Gomes (2014) in bands presented in Tables 1 and 2.

Name	Definition in database	Bands	Classification
Slope's height	HT	HT < 6 meters	Low
		6 meters < HT < 10 meters	Average
		HT > 10 meters	High

Table 1
Slope's height variable definition.

Name	Definition in database	Bands	Classification
Slope's inclination	IT	IT < 60°	Low
		IT > 60°	High

Table 2
Slope's inclination variable definition.

The scar's area and scar volume variables are quantitative and were estimated in fieldwork in conjunction

with the use of Google Earth Pro software images and tools. Figure 1 shows the 493+600 stretch with a slide

scar, field image and Google Earth Pro image.



Figure 1
Partial view of landslide's scar in stretch 493+600 of EFVM (left). Image of the year 2016 from Google Earth Pro of stretch 493 + 600 with polygon marking the scar (right).

The processing of database and cluster analysis, along with the final script was performed in software R, R Core Time (2016), version 1.0.136.

Ward's method presented better

suitability for the database. According to Ward (1963) Ward's method is to use the individuals themselves to construct the overall measure of group heterogeneity (W), that is, the method assumes that

each individual forms a group. W represents the sum of the squared Euclidean distances between each individual, i, and the mean of their corresponding group, g, see Equation 1.

$$W = \sum_g \sum_{i \in g} (X_{ig} - \bar{X}_g)^t (X_{ig} - \bar{X}_g) \quad (1)$$

The database is composed of quantitative and qualitative variables, mixed variables, and therefore

for the distance measurement of similarity, used was the Gower distance. Gower's distance was defined

by Gower (1971) and is presented in Equation 2.

$$G_{jk} = \left(\frac{\sum_{i=1}^n W_{ijk} S_{ijk}}{\sum_{i=1}^n W_{ijk}} \right) \quad (2)$$

G_{jk} - Gower's distance between the individuals j and k

W_{ijk} - Variable weight i between the individuals j and k

S_{ijk} - Partial similarity of variable i between the individuals j and k

3. Results and discussions

In order to analyze the relationship between the variables to obtain the

parametric relationship, a regression was performed between the two variables:

scar's area and volume. Figure 2 presents the result.

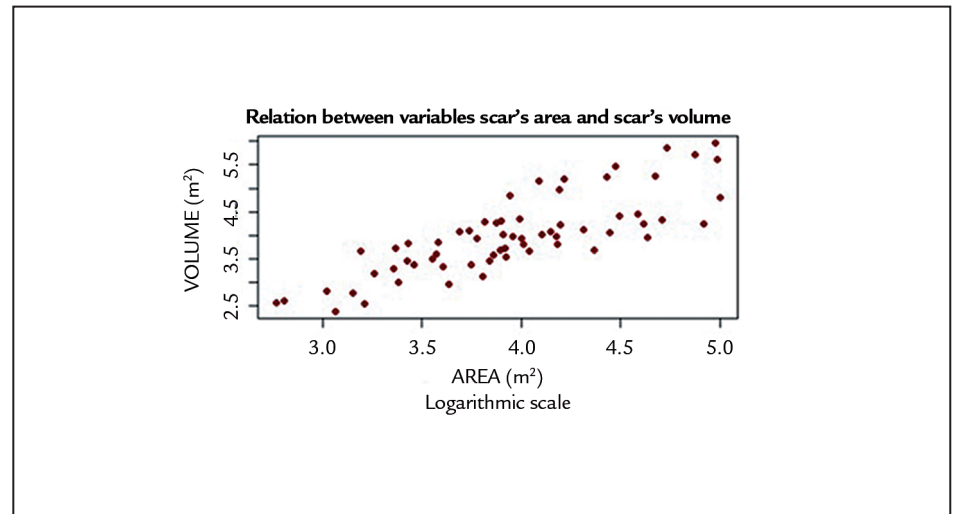


Figure 2
Relation between
variables: scar's area and volume.

The correlation coefficient between the scar's area and volume variables

was 0.66. Equations 3 and 4 present the relationship between the two variables,

which was obtained by means of the graph presented in Figure 2.

$$\log(V)=0.7338 \log (A)^{1.2249} \quad (3)$$

$$V=0.1416(A)^{1.2237} \quad (4)$$

Equations 3 and 4 can be compared with the relationships proposed by Simonett (1967), Rice *et al.* (1969), Innes (1985), Guthrie and Evans (2004), Korup (2005), Imaizumi and Sidle

(2007), Guzzetti *et al.* (2008) and Imaizumi *et al.* (2008). The main differences between the relationships may be due to the local geology, the data collection period, and the number of slopes stud-

ied. Table 3 presents the relationships cited and the relationship found in this research, and it is possible to conclude that the relationship is in agreement with literature.

Author	Relationship	Cases	Areas bands	Unit
Present work	$V=0.1416 \times A^{1.2237}$	64	$5.8 \times 10^2 < A < 1.0 \times 10^5$	m ²
Simonett (1967)	$V=0.2049 \times A^{1.3680}$	207	$2.5 \times 10^1 \leq A \leq 2.0 \times 10^6$	ft ²
Rice <i>et al.</i> (1969)	$V=0.2340 \times A^{1.1100}$	29	$2.1 \times 10^0 \leq A \leq 2.0 \times 10^2$	m ²
Innes (1985)	$V=0.0329 \times A^{1.3852}$	30	$3.0 \times 10^1 \leq A \leq 5.0 \times 10^2$	m ²
Guthrie and Evans (2004)	$V=0.1549 \times A^{1.0905}$	124	$7.0 \times 10^2 < A < 1.2 \times 10^5$	m ²
Korup (2005)	$V=0.0200 \times A^{1.9500}$	23	$A > 1.2 \times 10^6$	m ²
Imaizumi and Sidle (2007)	$V=0.3900 \times A^{1.3100}$	51	$1.0 \times 10^1 < A < 3.0 \times 10^3$	m ²
Guzzetti <i>et al.</i> (2008)	$V=0.0844 \times A^{1.4324}$	539	$1.0 \times 10^1 < A < 1.0 \times 10^9$	m ²
Imaizumi <i>et al.</i> (2008)	$V=0.1900 \times A^{1.1900}$	11	$5.0 \times 10^1 < A < 4.0 \times 10^3$	m ²

Table 3
Empirical relationships
between volume and area of landslide scars.

In order to calculate the similarities between the slopes, the Gower's distance was used and to create the grouping, the Ward's method was

used. The cophenetic correlation in Ward's method was 0.68, presenting an adequate value for the research. The choice of number groups was

helped by the analysis of distance chart in groups and by the dendrogram analysis presented in Figure 3.

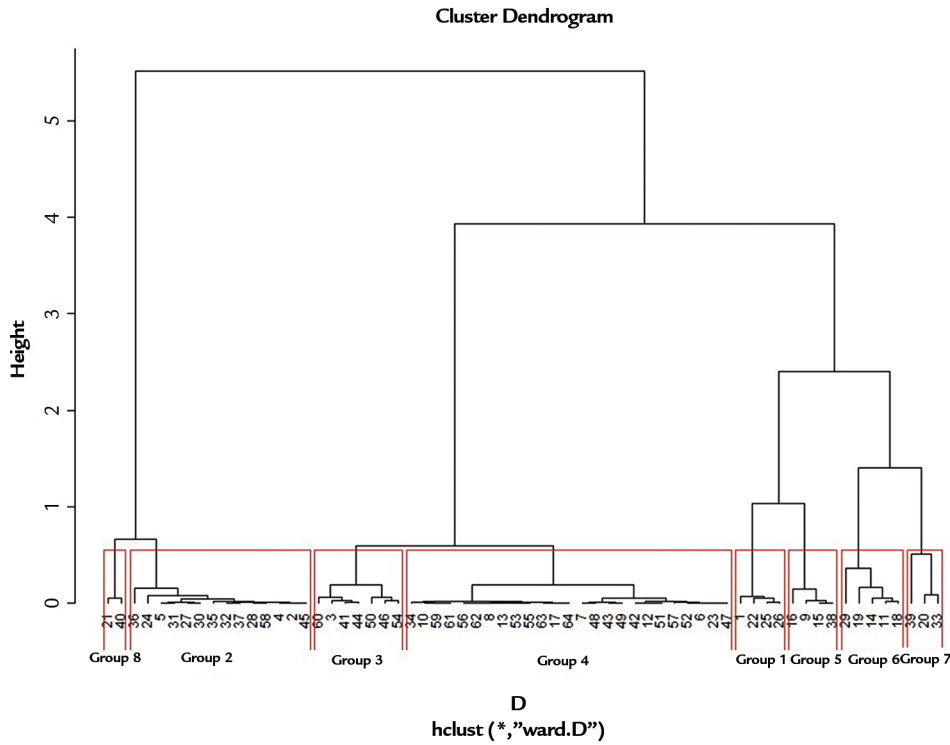


Figure 3
Cluster dendrogram, Ward's method.

For analysis of the dendrogram, presented in Figure 4, with the characteristics of the slope's scars, eight groups were selected for the studied stretches of the EFVM.

For the results of the cluster analysis, in relation to the quantitative variables, scar's area and scar's volume, the cluster technique was able to separate the area and volume bands

characterizing each group. Figures 4 and 5 show the boxplot of the variables with the bands comprising each group for the area and volume variables, respectively.

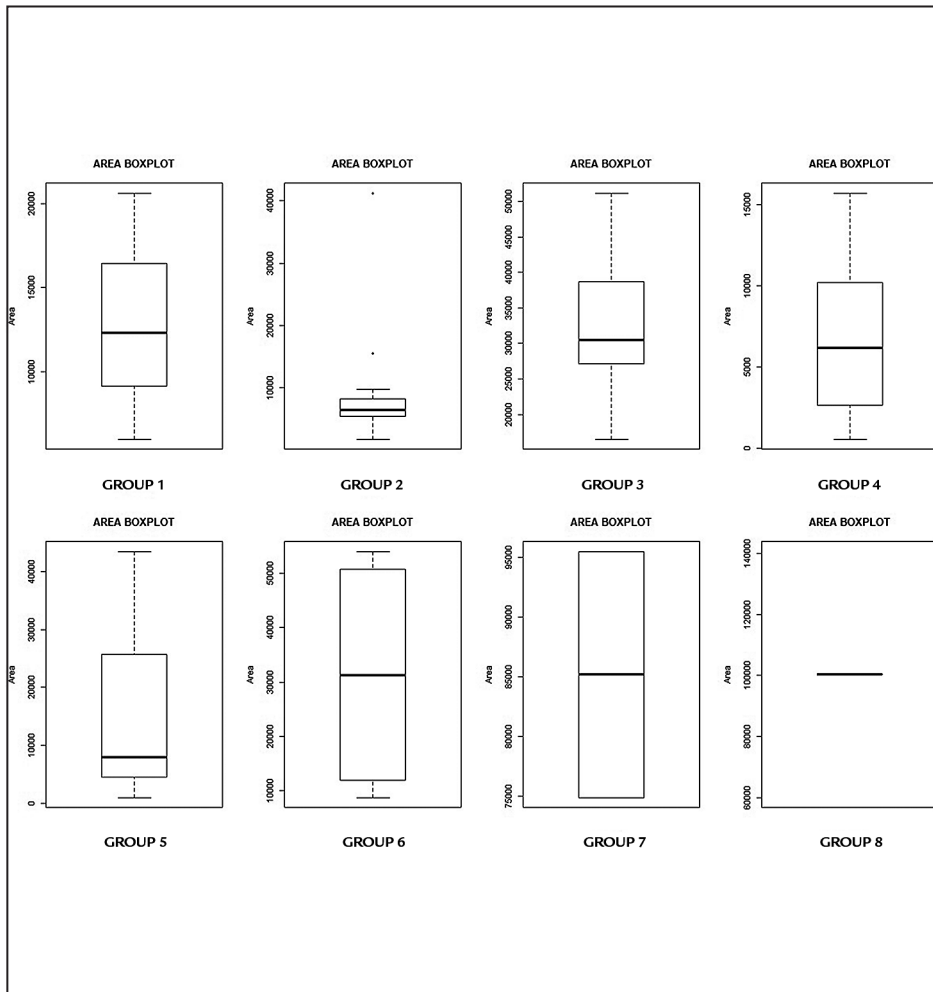


Figure 4
Area boxplot.

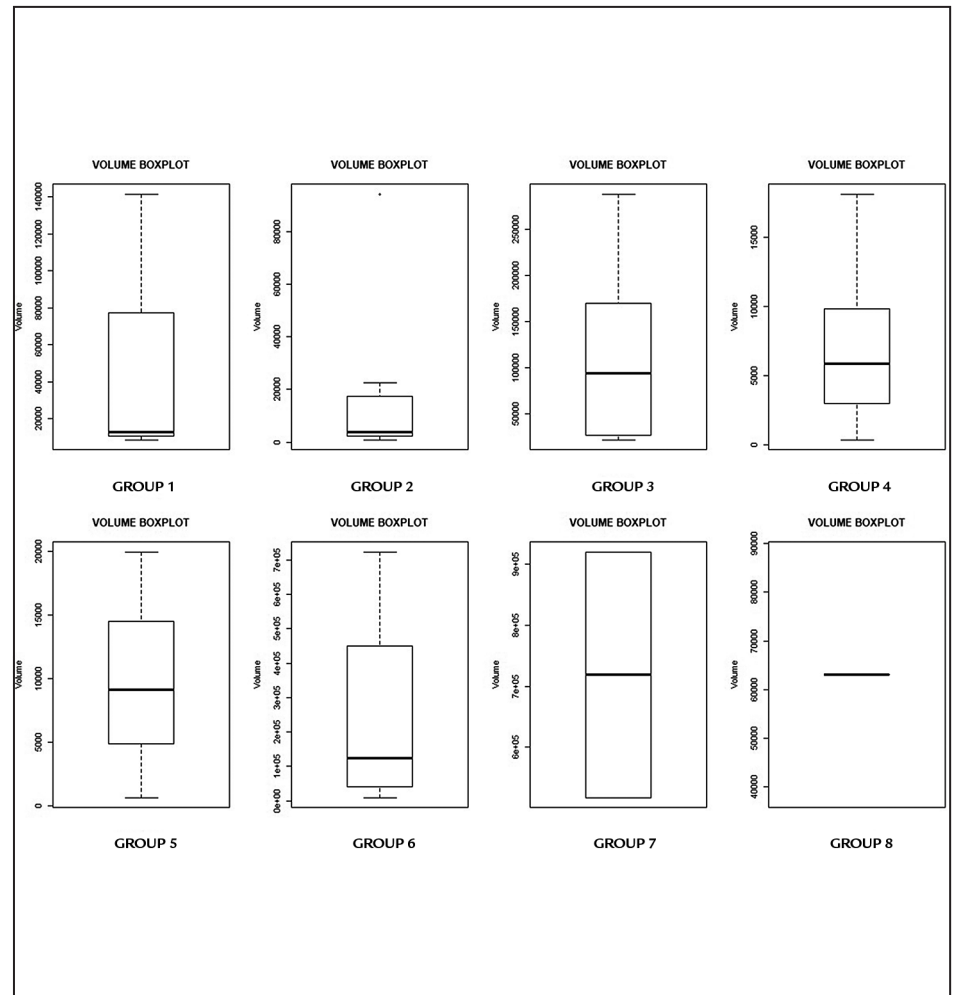


Figure 5
Volume boxplot.

When analyzing only the quantitative variables, it was observed that this technique was able to identify ranges for the scar's area and volume. In the case of the scar's area variable, the bands in some cases overlap; however, it does not happen

in the whole set values because each track presents unique values. For the volume, it is possible to observe that bands differ as to maximum values, not causing differences in the minimum values, except for Group 7.

Analyzing the final results of clustering in the whole database, qualitative and quantitative variables, it was possible to build Table 4, which presents the final characteristics of each group.

Group	Number of slopes	Slope's height	Slope's inclination	Maximum area (m ²)	Minimum área (m ²)	Maximum volume (m ³)	Minimum volume (m ³)
Group 1	4	low	low	2.06x10 ⁰⁴	1.62x10 ⁰³	1.42x10 ⁰⁵	3.40x10 ⁰²
Group 2	14	high	low	4.12x10 ⁰⁴	6.38x10 ⁰²	9.42x10 ⁰⁴	4.01x10 ⁰²
Group 3	7	high	high	5.11x10 ⁰⁴	1.65x10 ⁰⁴	2.89x10 ⁰⁵	1.17x10 ⁰⁴
Group 4	25	high	high	1.57x10 ⁰⁴	5.81x10 ⁰²	1.81x10 ⁰⁴	2.41x10 ⁰²
Group 5	4	low	high	4.34x10 ⁰⁴	1.05x10 ⁰³	1.99x10 ⁰⁴	5.91x10 ⁰²
Group 6	5	average	high	5.40x10 ⁰⁴	8.77x10 ⁰³	7.24x10 ⁰⁵	4.90x10 ⁰³
Group 7	3	high	high	9.75x10 ⁰⁴	7.49x10 ⁰⁴	9.20x10 ⁰⁵	4.08x10 ⁰⁵
Group 8	2	high	low	1.00x10 ⁰⁵	8.29x10 ⁰⁴	6.30x10 ⁰⁴	1.74x10 ⁰⁴

Table 4
Final characteristics of groups.

Among the grouped stretches, Group 7 stands out as the group that needs the most care in relation to the landslide development. In group 7, stretches 362 + 600, 093 + xxxE and 419 + 000 present high values of scar area and volume with high ranges of slope height and inclination, which can be interpreted as areas with predisposition for landslide occurrence. In cases of intervention measures, these would be the stretches that should be given priority.

Other stretches that need attention are in Group 3. Group 3 has seven slopes with high values for slope height and inclination, as well as ranges of values for the scar's area and volume. The stretches in Group 6 present average values for slope height with high slope inclinations, and when compared with the groups, stands out presenting high values for the scar area and volume variables.

The geographical aspect was analyzed, and allowed to sectorize the stretches in EFVM by cluster analysis. The groups formed are identified by characteristics of stretches in the group. Table 5 presents the stretches of EFVM with the results of cluster analysis.

Stretches	Group	Stretches	Group	Stretches	Group	Stretches	Group
011+000 D	1	094+330D	8	458+600	3	524+108	4
011+000 E	2	094+330E	1	465+042A	4	535+419	4
018+500	3	100+800	4	465+042B	4	535+420	4
019+500	2	103+100D	2	472+228	3	535+421	4
021+700	2	103+100E	1	472+307	2	093+xxxD	6
034+000	4	104+100D	1	477+225	3	093+xxxE	7
056+500	4	104+100E	2	478+500	4	419+000	7
056+xxx	4	108+150	2	482+299	4	442+400	8
062+xxxD	5	359+675A	6	482+930	4	521+000E	4
062+xxxE	4	359+675B	2	493+600	3	521+522	3
063+800D	6	359+675C	2	511+077	4	079+410D	4
063+800E	4	361+900	2	511+670	4	079+410E	6
064+xxx	4	362+600	7	511+899	4	404+000	2
072+500	6	369+200D	4	513+200	3	409+000	5
073+500	5	369+200E	2	515+300	4	519+900	4
078+900	5	395+850	2	518+790	4	521+000D	2

Table 5
Sectorization of stretches in EFVM.

4. Conclusions

The application of cluster analysis allowed the identification of eight groups with intrinsic characteristics. In the classified groups, there can be identified stretches of group 7, such as

362 + 600, 093 + xxxE and 419 + 000 with a tendency for landslides, due to the high values for slope height and inclination, along with considerable dimensions of scar area and volume.

Other groups can be cited, groups 3 and 6, with slope height and inclination values and considerable scar area and volume values.

Acknowledgements

The authors are grateful to Universidade Federal de Ouro Preto (UFOP), Fundação de Desenvolvimento da Pes-

quisa de Minas Gerais (FAPEMIG) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

for financial support and to Núcleo de Geotecnia da Escola de Minas of the Universidade Federal de Ouro Preto.

References

- GOMES, R. C. *Setorização Geológica-Geotécnica da Via e Avaliação e Controle de Riscos Geotécnicos dos Taludes da EFVM: Relatório Técnico Final*. Ouro Preto: Universidade Federal de Ouro Preto, 2014.
- GOOGLE EARTH PRO, versão 7.1.5.1557. Disponível em: <<https://www.google.com.br/earth/download/gep/agree.html>>. Acesso em: 23 nov. 2016.
- GOWER J. C. A general coefficient of similarity and some of its properties. *Biometrics*, v. 27, p. 857-874, 1971.
- GUTHRIE, R.H., EVANS, S.G. Analysis of landslide frequencies and characteristics in a natural system, coastal British Columbia. *Earth Surface Processes and Landforms*, v. 29, p. 1321-1339, 2004.
- GUZZETTI, F., ARDIZZONE, F., CARDINALI, M., GALLI, M., REICHENBACH, P., ROSSI, M. Distribution of landslides in the Upper Tiber River basin, central Italy. *Geomorphology*, v. 96, p. 105-122, 2008.
- GUZZETTI, F., ARDIZZONE, F., CARDINALI, M., ROSSI, M., VALIGI, D. Landslide volumes and landslide mobilization rates in Umbria, central Italy. *Earth and Planetary Science Letters*, v. 279, p. 222-229, 2009.
- IMAIZUMI, F., SIDLE, R.C. Linkage of sediment supply and transport processes in Miyagawa Dam catchment, Japan. *Journal Geophysical Research*, v. 112, (F03012), 2007.
- IMAIZUMI, F., SIDLE, R.C., KAMEI, R. Effects of forest harvesting on the occurrence of landslides and debris flows in steep terrain of central Japan. *Earth Surface Processes and Landforms*, v. 33, p. 827-840, 2008.
- INNES, J.N. Magnitude-frequency relationships of debris flows in northwest Europe.

- Geographiska Annaler*, v. 67A (1–2), p. 23–32, 1985.
- KORUP, O. Distribution of landslides in southwest New Zealand. *Landslides*, v. 2, p. 43–51, 2005.
- POLANCO, L. S. E. *Correlações Empíricas para Fluxo de Detritos*. Rio de Janeiro, RJ: COPPE/UFRJ, 2010. Originalmente apresentada como dissertação de mestrado, Universidade Federal do Rio de Janeiro, 2010.
- R CORE TEAM. R: A language and environment for statistical computing
Foundation for Statistical Computing, Vienna, Austria. Disponível em: <<https://www.R-project.org>>. Acesso em: 01 dez. 2016.
- RICE, R.M., CORBETT, E.S., BAILEY, R.G. Soil slips related to vegetation, topography, and soil in Southern California. *Water Resources Research*, v. 5, n.3, p. 647–659, 1969.
- SIMONETT, D.S. Landslide distribution and earthquakes in the Bewani and Torricelli Mountains, New Guinea. In: Jennings, J.N., Mabbutt, J.A. (Eds.), *Landscape Studies from Australia and New Guinea*. Cambridge University Press, Cambridge, p. 64–84. 1967.
- VALE. Sobre a Vale. Notícias. Disponível em: <<http://www.vale.com/brasil/PT/aboutvale/news/Paginas/ferrovia-vitoria-minas-transportou-119-milhoes-toneladas-cerca-1-milhao-pessoas-2014.aspx>>. Acesso em: 27 de jan. 2017.
- WARD, J. H., JR. "Hierarchical Grouping to Optimize an Objective Function", *Journal of the American Statistical Association*, v. 58, p. 236–244, 1963.

Received: 5 December 2017 - Accepted: 19 February 2018.