ASSOCIATION OF LANDSLIDES IN THE CITY OF BLUMENAU WITH ACCUMULATED RAINFALL IN THE REGION IN 2013 USING LOGISTIC REGRESSION

C.L.N. Seefelder[1&2], C.H.R. Lima[3], L. N. L. Gomes[3] and S. Koide[3]

1. *The National Department of Transport Infrastructure*
2. *Graduate Student in the Environmental Technology and Water Resources Program (PTARH), University of Brasilia*
3. *Department of Civil and Environmental Engineering, University of Brasília.*

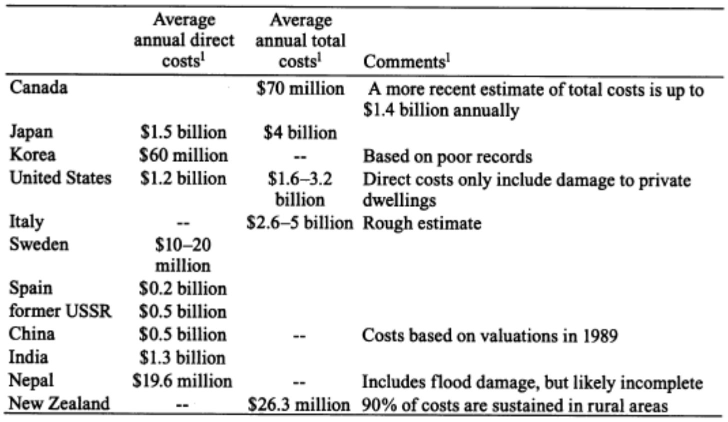
ABSTRACT: Landslide is a relevant natural disaster and the cause of significant economic and life losses worldwide. The understanding of the physical mechanisms responsible for a landslide event has thus practical importance to establish the most vulnerable regions and to develop control measures such as alert systems. Qualitatively, literature has shown that cumulative rainfall plays a significant role on landslide events, with heavy rainfalls associated with a higher chance of slope instability. This study investigates and develops a model to associate the likelihood of landslides in the city of Blumenau with the accumulated rainfall in the region. A logistic regression is proposed to establish a model to the probability of a landslide event occurring as a function of the previous rainfall. The results reveal that the accumulated rainfall for the previous 24 hours tends to be the most relevant variable to predict the probability of a landslide event in Blumenau, with a threshold value of roughly 50 mm leading to probabilities above 50%.

Key Words: rainfall, landslides, logistic regression.

1. INTRODUCTION

Problems with terrain instability are common worldwide. These events are responsible for a large number of life and economic losses. Table 1 shows the annual costs associated with impacts of landslides in different countries based on data from Sidle and Ochiai (2006).

Table 1: Estimated average annual costs (in US$) of landslides in various nations (source: Sidle e Ochiai, 2006).



	Average annual direct costs[1]	Average annual total costs[1]	Comments[1]
Canada		$70 million	A more recent estimate of total costs is up to $1.4 billion annually
Japan	$1.5 billion	$4 billion	
Korea	$60 million	--	Based on poor records
United States	$1.2 billion	$1.6–3.2 billion	Direct costs only include damage to private dwellings
Italy	--	$2.6–5 billion	Rough estimate
Sweden	$10–20 million		
Spain	$0.2 billion		
former USSR	$0.5 billion		
China	$0.5 billion	--	Costs based on valuations in 1989
India	$1.3 billion		
Nepal	$19.6 million	--	Includes flood damage, but likely incomplete
New Zealand	--	$26.3 million	90% of costs are sustained in rural areas

Impacts of land movements in Brazil have also been significant in the last years. For instance, the damage costs for the heavy rainfall that took place in January 2011 at Rio de Janeiro's mountainous region, whose main impact was due to landslides, were estimated around US$ 2.16 billion, of which approximately 66% was due to public sector and 34% due to private sector (World Bank, 2012).

Among the natural causes described in literature, water is regarded as one of the natural mechanisms that trigger land instability (Highland and Bobrowsky, 2008; Meisina and Scarabelli, 2007). According to Bressani and Bertuol (2012), in Brazil, intense rainfall is strongest trigger of mass movements, often catastrophic.

The importance of water for terrain stability can be supported by indications in researches (Coutinho and Bandeira, 2012; Futai et al., 2012; Highland and Bobrowsky, 2008) regarding drainage structure as a form of avoiding instability issues, pointing to direct relationship between mass movement and water in loco.

Soares and Marton (2006), Riekmann et al. (2005), Rio de Janeiro (2013), Coutinho (2002), and other authors have indicated the relationship between rainfall and landslides in the southeastern region. Coutinho (2002) verified rainfall data on Arthur hill, in Blumenau, and recognized that both accumulated and daily rainfall are important to analyze landslides.

Regarding the analysis on rainfall intensity and duration as well as the previous rainfall influencing landslides events, Simões (1991) points out some important issues concerning the Reconcavo Sedimentary Basin landslides. The author states that the previous rain influences the instability of slopes because of increasing soil moisture and the subsequent advance of the wetting front during heavy rains.

Rahardjo et al. (2008) compared soil properties on the surface layer of different slopes suggesting that relatively finer residual soil slopes and lower permeability require higher rates for the previous rainfall than less fine residual soil slopes with higher permeability to produce the highest (worst) pore-water pressure profiles in the slope. In other words, the previous rainfall is more significant to finer residual soils with lower permeability than to less fine residual soils with higher permeability when producing the highest (worst) pore–water pressure profile in a slope.

However, in situations where the instability mechanism occur through a network flow formation parallel to the contact surface for materials with different permeability, Simões (1991) explains that heavy rains are able to increase the percolation network thickness, having rainfall history minor due to high drainage capacity and water evaporation through these materials.

Regarding rocky slopes, the author explains that the instability is due to rainwater infiltration, in intense and long-lasting events that cause rapid elevation of water level within the mass to fill the fractures, generating hydrostatic pressures on the wedges, especially when the volume of water infiltration is greater than the capacity of draining solid.

By analyzing researches results, Gostelow (1991) suggests two separate processes affect shallow soil stability. Firstly, loss of suction at the unsaturated zones, i.e. a comparatively large change in pore pressure followed by a loss of shear strength. Secondly, increased water level after a delay resulting in a slight change in pore-pressure (comparatively) and strength.

The indications above regarding water as an important factor to the instability process reveal that it is important to study the relationship between rainfall and precipitation as a form of indicating vulnerable area and rain threshold triggers landslides, supporting managers in the implementation of public policies.

Coelho Neto et al. (2012) observed the distribution of monthly average rainfall (records on 14 years), the total monthly for the previous six months on critical events in February 1988 and 1996 and the daily rainfall distribution in these two months verifying that the landslides events had a great variety in hydrologic perform.

Coutinho (2002) studied the relationship between rainfall and mass movement in *Morro do Artur*, Blumenau and concluded that the mass movement there is not associated only with high precipitation records, but also with extensive rainfall of low intensity leading to saturated soil.

Azevedo (2011) used recorded data on accumulated rainfall for an hour and 24 hours to calculate the probability of rupture using Gaussian curve for each studied station. In general, the curves presented increasing probability of rupture in relation to increasing accumulated rainfall, however, for some stations, after reaching a maximum value; the curves began to present decreased probabilities in relation to increased accumulated rainfall. Fact the authors attributed to increased stability in this area provided as most accidents were established according to lower rates of accumulated rainfall leading to an equilibrium state. Some curves of rupture probability present difference when compared with data on accumulated rainfall for an hour and 24 hours for the same station, which, according to the author, indicates that each region has their own characteristics leading to a form of supporting what rainfall requires.

All these studies have indicated a causal relationship between accumulated rainfall and landslides events, suggesting also that providing landslides probability in an area could be an important tool to mitigate and eventually avoid the impacts of landslide disasters.

In this context, this study focus on modeling the probability of landslides for the city of Blumenau using empirical data of rainfall and landslides occurrence in 2013 in a logistic regression framework. Such curve provides public managers with a useful instrument when adopting measures to create acceptable rainfall thresholds to settle indicative measures regarding each risk.

This manuscript is organized as follows. In the next section we present the region of study and the methodology adopted here. The results are presented in section 3, followed by conclusions.

2. REGION OF STUDY AND METHODOLOGY

The main goal of this study is the analysis and modeling of the relationship between rainfall intensity and landslide occurrence in the city of Blumenau for the period from March 01, 2013 through September 22, 2013. The city of Blumenau is located in Santa Catarina state, along the in *Itajaí Valley*, Brazil (see Figure 1). The city has frequently experienced water related disasters, mainly floods and landslides.

Data on dates and location of landslide occurrence across the city are not made available for any institution, so we use press information available on the web to identify the dates of landslide occurrences. The analyzed period revealed seven landslides on 03/20/2013, 06/21/2013, 07/20-21-22/2013, 09/17/2013 and 09/21/2013.

Rainfall data is provided by the National Institute of Meteorology (INMET) for the *Indaial* station (Figure 2).

We model the occurrence of a landslide as a Binomial random variable Y, assuming the values 1 and 0 (no landslide event). The expected value E(Y), i.e., the probability of a landslide event, is non-homogeneous and model as a function of a covariate x, which in our case is related to the previous rainfall (see Table 2). A logistic regression framework is then used to estimate the model parameters, where a monotonically increasing (or decreasing) function in S-shaped (or inverted S-shaped) is used to connect (link function) the expected value of the response variable Y to the covariate x. Here we use the logit response (see equation 1):

$$E(Y) = \frac{exp(\beta_0 + \beta_1 x)}{1 + exp(\beta_0 + \beta_1 x)} \qquad [1]$$

where β0 and β1 are unknown parameters estimated by the maximum likelihood method. The expected E (Y) value gives the average probability of a sliding event as a function of the value of the covariate x. The model is fitting using the observed values of Y, which is equal to 1 in the days there was a landslide event and 0 otherwise, and the values of x for the respective days.

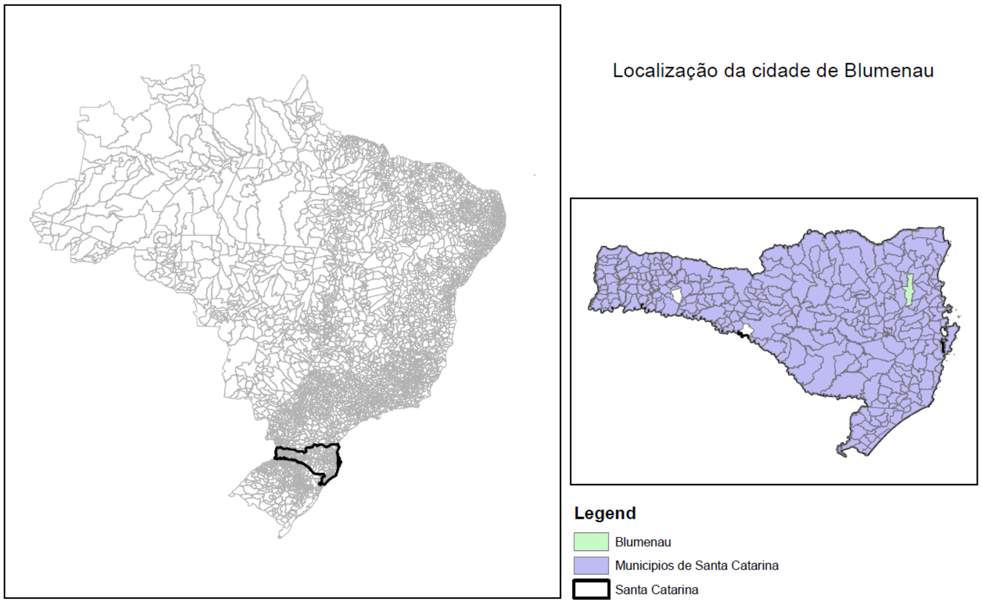


Figure 1: Location of Blumenau city.

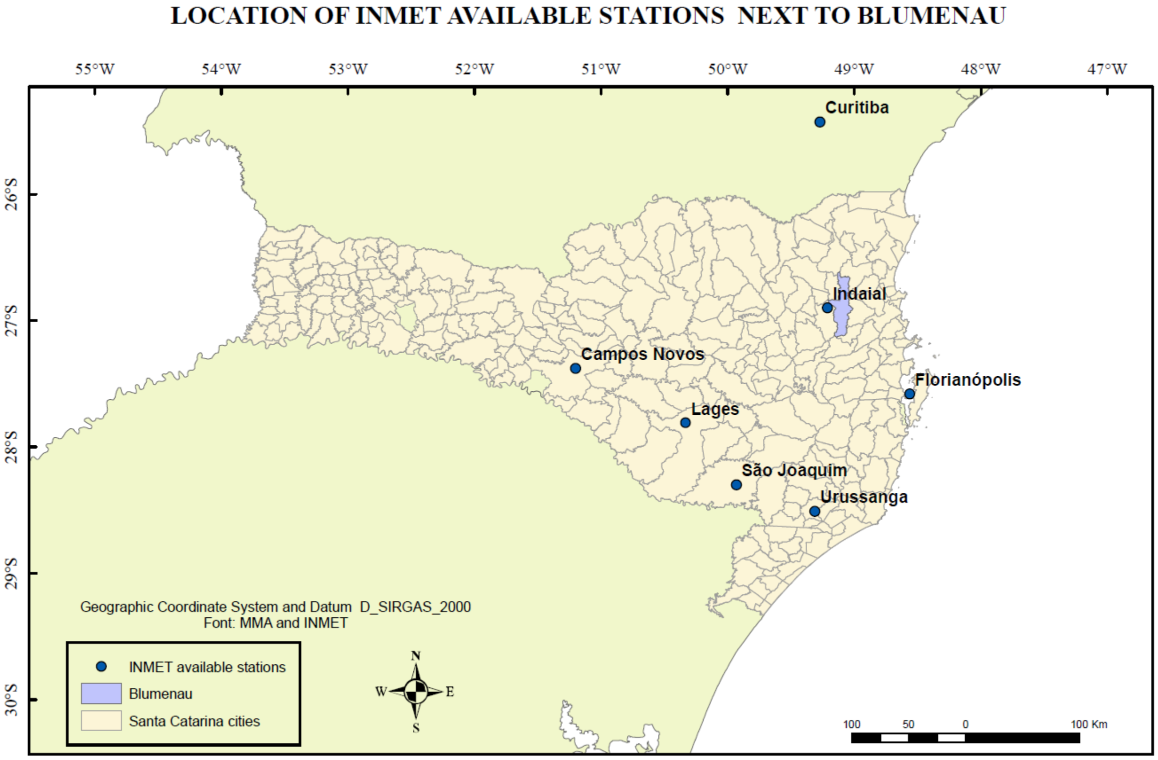


Figure 2: INMET station location map nearest to Blumenau-SC city.

Here we consider different set of predictors x for E(Y). Table 2 shows the different models analyzed here. Model 1 and 2 consider the observed rainfall at days t and t-1, respectively. Model 3 is the accumulated rainfall over days t and t-1 and model 4 considers the accumulated rainfall in the last 15 days. Finally, a weighted average of the rainfall at day t and the 48-hour accumulated rainfall is used as predictors. The weights are the first loading obtained from principal component analysis (see Andrade *et al.* (2003), Costa (2008) and Bernardi *et al.* (2009) for more details) on these two variables. Note that the values of Y (0 or 1) refer to the occurrence (or not) of a landslide event at day t

Table 2: Covariates used in the logistic regression model.

Model	Rainfall Variables	Remakers
1	x_t	Analyzed day
2	x_{t-1}	Accumulated rainfall for the previous day
3	$x_t + x_{t-1}$	48-hour accumulated rain (Analyzed day + Accumulated rainfall for the previous day)
4	$\sum_{i=0}^{14} x_{t-i}$	15-days accumulated rainfall (14 previous days + analysed day)
5	$0.51x_t + 0.86(x_t + x_{t-1})$	A combined rain of 24 and 48-hour accumulated rain

Finally, the model fit is evaluated through the Akaike and Bayesian information criteria and the Nagelkerke R ^{2}.

3. RESULTS AND DISCUSSION

The first analysis conducted on available records concerned rainfall data normality verified no regular behavior. These analyses were carried out for the 24-hour accumulated rainfall (slip event day), accumulated rain the day before the event, the 48-hour accumulated precipitation (slip event day + day before the event) and accumulated precipitation for 15 days.

Figures 3 to 7 display the fitting results for the models indicated in table 2. The model skills are evaluated based on the p-value of β_1 at the $\alpha = 5\%$ significance level, on the Akaike information criterion (AIC), on the Bayesian information criterion (BIC) and on the R ^{2}Nagelkerke.

The p-value for β_1 represents the probability of obtaining a test statistics equal to or more extreme than the one observed in the sample, under the null hypothesis (which in this case is $\beta_1 = 0$, i.e., no statistical significance for β_1). If the p-value is less than α, then we should reject the null hypothesis for β_1, so it is considered statistically different from zero. The results show that all p-values but the model 5 one are less than α, indicating that only the 15-day accumulated rainfall predictor is not statistically significant.

The sort of soil in the examined region is relatively permeable (Aumond et al., 2009) therefore the accumulated precipitation extended periods are not significant in the stability analysis and the lack of significance of the 15-day accumulated rainfall is possibly associated with the low water retention in the soil for this region.

Vieira and Furtado (2005) analyzed rainfall records regarding the occurrence of slips from 1997 to 2001 in the city of Blumenau observing that the instability events are associated with the accumulated precipitation for a period of three to four days, having occurred cases of slips even in the absence of rain in 24 hours.

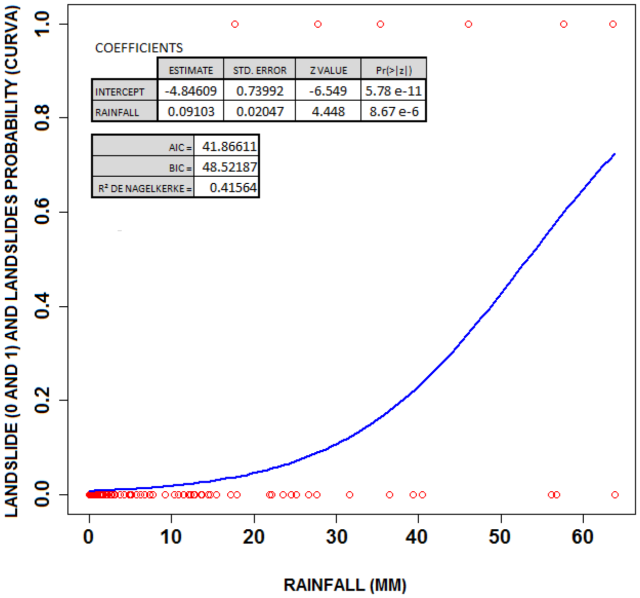
Such data can infer that the soil should influence the actuation of slips, since this variable affects the amount of water stored underground. By analogy, one can also infer that the variables of occupation, land use and vegetation cover should influence the instability events as well.

The observations by Vieira and Furtado (2005) are in accordance with the results found in the statistical analysis performed in this study revealing an increasing influence on the likelihood of slip when having the 24-hour and the 48-hour accumulated rainfall analyzed and no influence when having the 15-day accumulated precipitation analyzed.

Statistical tests were carried out to assess how the available models are close to reality. This analysis applied the measures AIC, BIC and R2 Nagelkerke. The first two measures are based on the maximum likelihood function and have their values estimated by comparing two models (the analysis with "real" or "true"). Both models presenting lower values of AIC / BIC were more satisfactory.

The analyzed data revealed that the lowest values of AIC and BIC were obtained for the 48-hour accumulated precipitation, indicating that these data have more satisfactory results.

The R ² Nagelkerke is a variation of the R ² of Cox and Snell (also based on the likelihood function) which expresses the power of the model prediction. The values vary from zero to one, where one is the perfectly satisfactory and zero reveals the lack of correlation between the independent and dependent variables.



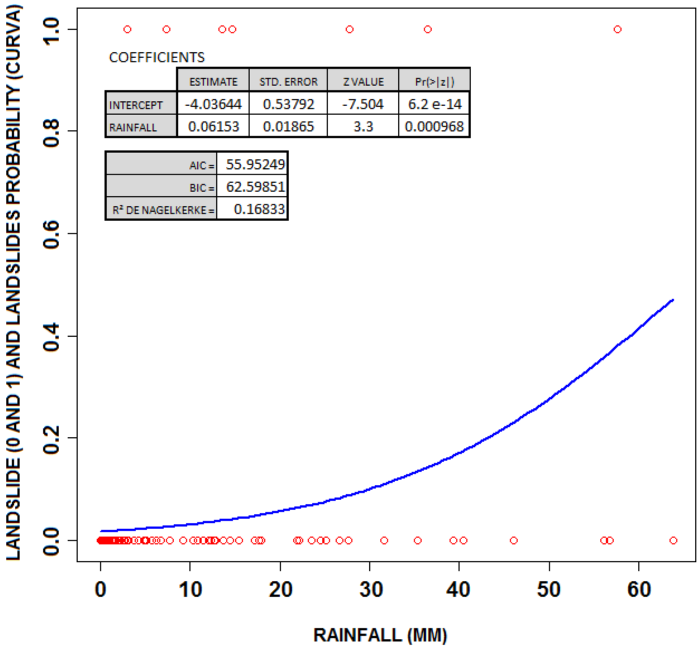
	ESTIMATE	STD. ERROR	Z VALUE	Pr(>|z|)
INTERCEPT	-4.84609	0.73992	-6.549	5.78 e-11
RAINFALL	0.09103	0.02047	4.448	8.67 e-6

AIC =	41.86611
BIC =	48.52187
R² DE NAGELKERKE =	0.41564

Figure 3: Fitting results for model 1. The red circles show the occurrences (1) and absences (0) of landslide events for the period analyzed. The blue curve shows the probability of a landslide event as a function of the predictor x based on the model fitting.The boxes within the figure show the parameter estimates and the AIC, BIC and R^2 criteria to measure the model quality.

By analyzing the R ² Nagelkerke coefficient, we can state that model 3, which has the 48-hour accumulated rainfall curve as predictor, has the best fit, with a value of 47%, while the analysis on the 24-hour accumulated precipitation has a value of 41%, the previous days submitted 17% and the 15-day accumulated rainfall value was only 5.4%.

This result indicates that the rainfall for the previous days influenced the slip phenomenon. When jointly analyzing these data, it is possible to reach a more satisfactory prediction on the probability of instability events occurring. Considering the 48-hour accumulated precipitation, it seems that an approximate value of 75 mm provides a 50% probability of landslide occurrence growing to 90% when rainfall reaches 106 mm. Data on time variation of rainfall can be useful to managers when defining levels of risk and warnings on preventive actions to avoid landslides-related losses.

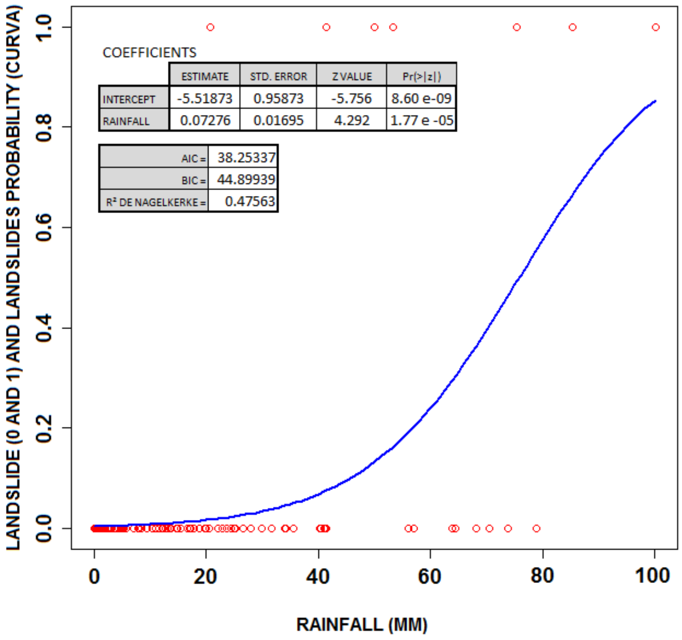


COEFFICIENTS

	ESTIMATE	STD. ERROR	Z VALUE	Pr(>|z|)
INTERCEPT	-4.03644	0.53792	-7.504	6.2 e-14
RAINFALL	0.06153	0.01865	3.3	0.000968

AIC =	55.95249
BIC =	62.59851
R² DE NAGELKERKE =	0.16833

Figure 4: As in Fig. 3, but for model 2.



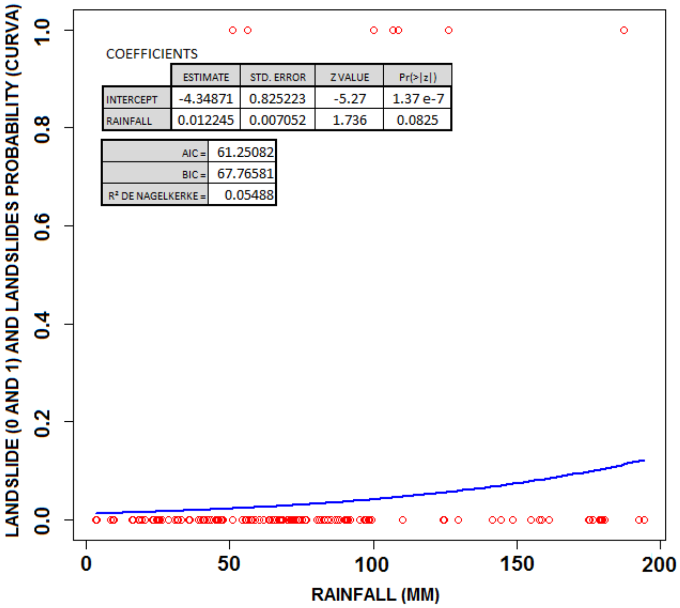
COEFFICIENTS

	ESTIMATE	STD. ERROR	Z VALUE	Pr(>|z|)
INTERCEPT	-5.51873	0.95873	-5.756	8.60 e-09
RAINFALL	0.07276	0.01695	4.292	1.77 e -05

AIC =	38.25337
BIC =	44.89939
R² DE NAGELKERKE =	0.47563

Figura 5: As in Fig. 3, but for model 3.

After these analyses, a new situation was observed after combining the 24-hour with the 48-hour accumulated rainfall and the using the method of principal component analysis. This new proposal aims at checking if there is a gain in predictability when jointly analyzing rainfall. The rate obtained for the combination is presented in Table 2.



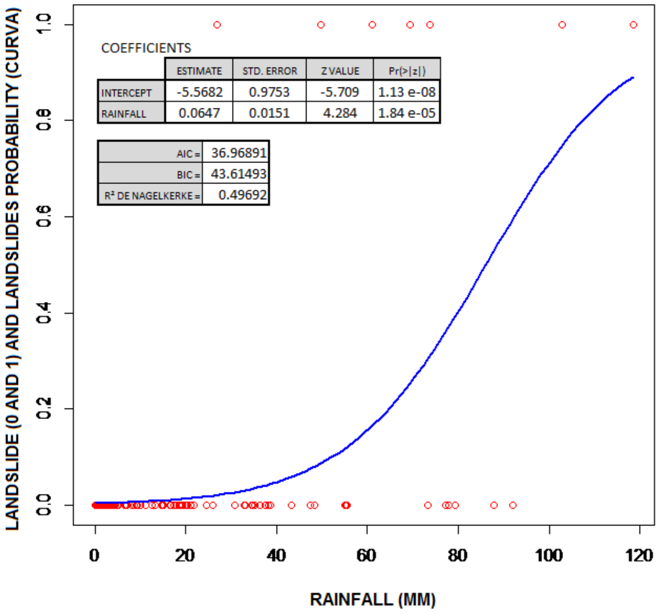
COEFFICIENTS

	ESTIMATE	STD. ERROR	Z VALUE	Pr(>|z|)
INTERCEPT	-4.34871	0.825223	-5.27	1.37 e-7
RAINFALL	0.012245	0.007052	1.736	0.0825

AIC =	61.25082
BIC =	67.76581
R² DE NAGELKERKE =	0.05488

Figura 6: As in Fig. 3, but for model 4.

Figure 7 presents the results obtained for the statistical analysis on the combined rain above, using the same methods and the same analyses conducted described for the previous data.



COEFFICIENTS

	ESTIMATE	STD. ERROR	Z VALUE	Pr(>|z|)
INTERCEPT	-5.5682	0.9753	-5.709	1.13 e-08
RAINFALL	0.0647	0.0151	4.284	1.84 e-05

AIC =	36.96891
BIC =	43.61493
R² DE NAGELKERKE =	0.49692

Figure 7: As in Fig. 3, but for model 5.

By applying the same level of significance (α = 0.05) statistically significant results were obtained for the combination analyzed. The AIC and BIC values obtained for the combination had more satisfactory performance than those obtained in the previous analysis. Values of R^2 Nagelkerke corroborate this assessment when obtained by combining accounting for over 49% of the phenomenon explanation according to high levels of rainfall compared with 47% of the most satisfactory result previously obtained (accumulated rainfall and 48 hours). For this combination, values of 86 to 120 mm for the independent variable slip generated a probability of 50% and 90%, respectively. Figure 8 graphically illustrates the values of BCI and R^2 Nagelkerke for different models tested for graphical analysis as described in this statement.

The curves of slip probability that have the rain as variable of influence reveal that values of the 48-hour accumulated precipitation, individually assessed and combined with the rain on the day of the event, improve the predictive performance of landslides occurrence for the city of Blumenau. The correlation rates found indicate that rain is quite relevant when studying instability events , leading us to believe that more comprehensive studies on such influence can bring promising results to develop management tools for such events.

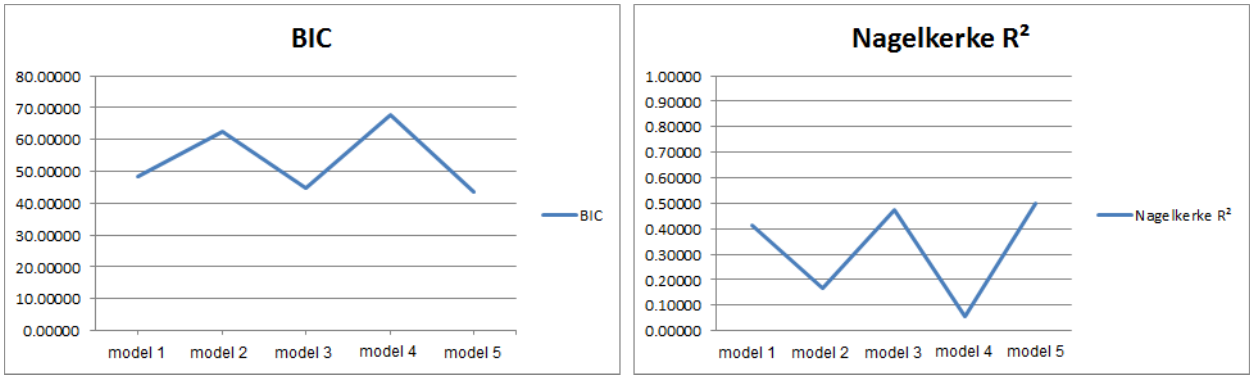


Figure 8: Obtained results for BCI and Nagelkerke R^2 for all tested models

4. CONCLUSIONS

Correlation data analysis between precipitation and landslide occurrences demonstrates the importance of the previous rainfall as a triggering factor of these events, indicating that rainfall is associated with the greatest probability of events occurrence, as demonstrated by logistic curves and the results of increasing occurrence of slip probability regarding increased rainfall values.

The results obtained here indicate that the rainfall on the day of the landslide event as well as the 48-hour accumulated rainfall play a significant role on the probability of landslide occurrences. For instance, a 75 mm accumulated rainfall in the last 48 hours leads to a 50% probability of a landslide event. This information is important for public agents (e.g. Civil Defense) to define rainfall thresholds needed to take emergency measures when crossed.

Available data on rainfall at shorter intervals would facilitate the measurement of previous rain and the discretization into smaller intervals in order to assess both the influence of time and the intensity of mass movements. Greater available rainfall data as rain gauge network expands as well as the precise location of landslides (inventory of landslides) and the recorded time of the event would enable greater accuracy of rainfall data and would reduce the influence of the spatial and temporal variability of rainfall events on the model fitting.

As future studies, we plan to extend the models developed here by including new variables that account for the exact location of the slip events as well as the slope shape, slope, catchment area, land cover and land use, which should improve our understanding on the factors that most affect the stability of slopes along the Blumenau city.

5. REFERENCES

Ana Luiza Coelho Netto, André de Souza Avelar, Anderson Mululo Sato, Maria Accioly Dias, Mônica Bahia Schlee, André Batista de Negreiros, 2012: “Vulnerabilidade em Geossistemas Montanhosos e Desastres Causados por Deslizamentos na Interface Florestal-Urbana: Controle Geológicos, Geomorfológicos e Geoecológicos”. In: Lacerda, Willy Alvarenga et. al. (Orgs). *Desastres Naturais: susceptibilidade e risco, mitigação e prevenção e gestão e ações emergenciais*. COPPE/UFRJ, Rio de Janeiro, RJ, 63-80.

Bernardi, Lacerda, Dórea, Landim, Gomes, Almeida, Manzatto e Bastos, 2009: “Aplicação da análise das componentes principais na ordenação dos parâmetros físico-químicos no alto Rio Madeira e afluentes, Amazônia Ocidental” *Geochimica Brasiliensis*, 23(1) 079-090.

Coutinho, 2002: “A precipitação e sua influência na movimentação de massa: estudo de um caso” *Health and Environment Journal*, v.3, n.2.

Eunice Maia de Andrade, Sidnéia Souza Silveira e Benito Moreira de Azevedo, 2003: “Investigação da Estrutura Multivariada da Evapotranspiração na Região Centro Sul do Ceará pela Análise de Componentes Principais” *Revista Brasileira de Recursos Hídricos*, V. 8n., p. 39–44.

Futai, Wolle, Bastos e Silvia Suzuki, 2009. “Utilização de túneis e galerias de drenagem para estabilização de encostas” 5ª Conferência Brasileira de Estabilidade de Encostas, São Paulo/SP, Brasil, Novembro 223-232, 2009.

George Fernandes Azevedo, 2011: *Análise da relação chuva-escorregamentos em regiões tropicais montanhosas urbanas, caso de estudo Rio de Janeiro, Brasil*. Universidade de Brasília, Brasília/BR.

Gostelow, 1991. “Rainfall & Landslides”. In: *Prevention and control of landslides and other mass movements*. CEC Report EUR 12918, Luxembourg, 37-53.

Juarês José Aumond, Lúcia Sevegnani, Mário Tachini e Lauro Eduardo Bacca, 2009: “Condições naturais que tornam o vale do Itajaí sujeito aos desastres”. In: Frank, B.; Sevegnani, L. (Orgs.). Desastre de 2008 no Vale do Itajaí. Blumenau: Fundação Agência De Água Do Vale Do Itajaí, 2009, 23-37p.

Luís Antônio Bressani e Fábio Bertuol, 2012: “Alguns Escorregamentos do RS e de SC e a Avaliação de Susceptibilidade e Risco de Encosta”. In: Lacerda, Willy Alvarenga *et al.* (Orgs). *Desastres Naturais: susceptibilidade e risco, mitigação e prevenção e gestão e ações emergenciais*. COPPE/UFRJ, Rio de Janeiro, RJ, 81-94.

Lynn M. Highland, Peter Bobrowsky, 2008: *The landslide handbook – A guide to understanding landslides*. U.S. Geological Survey Circular 1325, Reston, Virginia.

Meisina e Scarabelli, 2007: “A comparative analysis of terrain stability models for predicting shallow landslides in colluvial soils” *Geomorphology*, 87, 207-223.

Pedro Daniel Melo da Costa, 2008: *Dinâmica e Variabilidade do vapor de água na Península Ibérica*. Universidade de Aveiro.

Rafaela Vieira e Sandra Maria de Arruda Furtado, 2005: “Percepções frente ao risco de deslizamento” Revista Geosul, 20:40, 55-75.

Rahardjo, Leong e Rezaur, 2008: “Effect of antecedent rainfall on pore-water pressure distribution characteristics in residual soil slopes under tropical rainfall”. *Hydrological Processes*, 22, 506-523.

Riekmann, Pinheiro e Tachini, 2006: “Estudo da relação entre precipitações e deslizamentos em áreas de risco em Blumenau – SC” I Simpósio Sul-Sudeste de Recursos Hídricos, Curitiba/PR, Brasil, Agosto, 2006.

Rio de Janeiro, 2013: *Relatório de Correlação entre Chuvas e Escorregamentos para a Cidade do Rio de Janeiro no ano de 2012*. Relatório GEO-RIO/DEP/GPE - nº 5/2013. Disponível em: < http://www0.rio.rj.gov.br/alertario/upload/Relatorio_Escorregamentos_2012.pdf>. Acessado em 25/01/2014.

Roberto Quental Coutinho e Ana Patrícia Nunes Bandeira, 2012: “Gerenciamento de Áreas de Risco: Ações Estruturais e não Estruturais”. In: Lacerda, Willy Alvarenga et. al. (Orgs). *Desastres Naturais: susceptibilidade e risco, mitigação e prevenção e gestão e ações emergenciais*. COPPE/UFRJ, Rio de Janeiro, RJ, 163-178.

Sidle e Ochiai, 2006. Introduction and Overview of Landslide Problems. In: Sidle, R. C; Ochiai, H. Landslides - Processes, Prediction and Land Use. Whashington, DC: AGU Books Broad, 2006, p. 1-22.

Simões, 1991: *Mecanismos de instabilização dos taludes naturais das formações expansivas da bacia sedimentar do Recôncavo Baiano*. Universidade Federal do Rio de Janeiro. Rio de Janeiro/RJ.

Soares e Marton, 2006: “Relação entre precipitação e deslizamentos de encostas na região de Angra dos Reis” XIV Congresso Brasileiro de Meteorologia, Florianópolis/SC, Brasil, Agosto, 2006.

Worl Bank, 2012: Avaliação de Perdas e Danos: Inundações e Deslizamentos na Região Serrana do Rio de Janeiro – Janeiro de 2011. Brasília, p. 59.