

# MODELAGEM DE ÁREAS SUSCETÍVEIS A MOVIMENTOS DE MASSA COM REDES NEURAIS ARTIFICIAIS

Renata Pacheco Quevedo<sup>1</sup>, Guilherme Garcia de Oliveira<sup>2</sup>, Samuel Gameiro<sup>3</sup>, Luís Fernando Chimelo Ruiz<sup>4</sup>, Laurindo Antonio Guasselli<sup>5</sup>

<sup>1</sup> Mestranda no PPGSR<sup>a</sup>, renata.quevedo@ufrgs.br; <sup>2</sup> Professor do DI<sup>b</sup> e do PPGSR<sup>a</sup>, g.g.oliveira10@gmail.com; <sup>3</sup> Mestrando no PPGSR<sup>a</sup>, samuel.gameiro@ufrgs.br; <sup>4</sup> Doutorando do PPGSR<sup>a</sup>, ruiz.ch@gmail.com; <sup>5</sup> Professor do DG<sup>c</sup>, do POSGea<sup>c</sup> e do PPGSR<sup>a</sup>, laurindo.guasselli@ufrgs.br.

<sup>a</sup>Pós-Graduação em Sensoriamento Remoto (PPGSR), Centro Estadual de Pesquisas em Sensoriamento Remoto e Meteorologia, Universidade Federal do Rio Grande do Sul – UFRGS, Av. Bento Gonçalves, 9500, Porto Alegre/RS.

<sup>b</sup>Departamento Interdisciplinar (DI), UFRGS, RS-030, km 92, 11.700, Tramandaí – RS.

<sup>c</sup>Departamento de Geografia (DG) e Programa de Pós-Graduação em Geografia (POSGea), Instituto de Geociências, UFRGS, Av. Bento Gonçalves, 9500, Porto Alegre/RS.

## RESUMO

As RNAs se destacam na modelagem e mapeamento de suscetibilidade por apresentarem elevada acurácia dos mapas gerados, capacidade de aprendizagem e generalização dos resultados. Assim, este estudo teve como objetivo o mapeamento de áreas suscetíveis a movimentos de massa na Bacia Hidrográfica do Rio Rolante, a partir de um modelo de RNA. Para tal, foi elaborado um inventário de cicatrizes, extraídos parâmetros morfométricos do MDE ALOS PALSAR, organizado um conjunto amostral e realizados treinamento e validação dos modelos. Foram testadas 8 configurações de modelo, com variação do número de parâmetros de entradas e de neurônios. Como resultados, dentre os atributos extraídos, ELV, DCL, LS e PRV apresentaram maior correlação com a ocorrência de movimentos de massa, enquanto que ASP, ACF, TWI e DRD apresentaram correlação baixa. Concluiu-se que as RNAs se mostraram capazes de modelar as áreas suscetíveis, com acurácia no mapeamento superior a 0,9.

**Palavras-chave** — Modelo Digital de Elevação, ALOS-PALSAR, Bacia Hidrográfica do Rio Rolante.

## ABSTRACT

*The RNAs stand out in the modeling and mapping of susceptibility because they present high accuracy of generated maps, learning capacity and generalization of the results. Thus, this study had the objective of mapping areas susceptible to mass movements in the Rolante River Basin, from an RNA model. For that, an inventory of scars was drawn, morphometric parameters of the ALDE PALSAR MDE were drawn up, a sample set was organized and training and validation of the models were carried out. Eight model configurations were tested, varying the number of inputs and neurons parameters. As a result, among the attributes extracted, ELV, DCL, LS and PRV presented a higher correlation with the occurrence of mass movements, whereas ASP, ACF, TWI and DRD presented a low*

*correlation. It was concluded that the RNAs were able to model the susceptible areas, with accuracy in the mapping higher than 0.9.*

**Key words** — Digital Elevation Model, ALOS-PALSAR, Rolante River Basin.

## 1. INTRODUÇÃO

Os movimentos de massa podem ser definidos como deslocamentos descendentes de materiais (solos, fragmentos de rochas, etc.), sendo normalmente deflagrados por precipitações intensas, que promovem a saturação do solo. Na maioria dos casos, esses movimentos são o resultado de processos naturais que integram a evolução do relevo e paisagem ao longo do tempo geológico.

Esses fenômenos são um dos principais tipos de desastres registrados no Brasil, responsáveis pelo segundo maior número de óbitos entre os anos de 1991 e 2012 [1]. Isso se dá, entre outros motivos, pelo processo de urbanização rápido e desordenado, ocorrido principalmente a partir da década de 1960. Nesse sentido o mapeamento de áreas suscetíveis a movimentos de massa surge como uma ferramenta para auxiliar no ordenamento do território [2].

Diversos autores têm apresentado o uso de modelos para simular a ocorrência de movimentos de massa e produzir, assim, mapeamento de suscetibilidade [3] [4]. Dentre os modelos utilizados, as Redes Neurais Artificiais (RNAs) têm se destacado pela acurácia, pela capacidade de aprendizagem e de generalização dos resultados [5].

Nesse contexto, o presente estudo teve como objetivo o mapeamento das áreas suscetíveis a movimentos de massa a partir de um modelo de RNAs, tendo como variáveis de entrada, os atributos do terreno extraídos de um Modelo Digital de Elevação (MDE), e o conjunto amostral obtido em função de cicatrizes mapeadas sobre imagens de satélite.

## 2. MATERIAIS E MÉTODOS

### 2.1) Área de estudo

A Bacia Hidrográfica do Rio Rolante está localizada na porção nordeste do estado do Rio Grande do Sul (Figura 1), na Região Hidrográfica do Guaíba. A região compreende um relevo declivoso, com altitudes variando entre 152 m e 992 m e apresenta um clima subtropical úmido, sendo o regime de chuvas bem distribuído ao longo do ano, influenciado principalmente pelo ingresso de sistemas frontais de precipitação, mas também com influência da orografia e de sistemas convectivos de mesoescala.

Em 05 de janeiro de 2017, ocorreu um evento extremo de precipitação, caracterizado por sua concentração espacial na região da cabeceira da bacia do rio Rolante, na sub-bacia do rio Mascarada. Os dados oficiais de chuva deste evento subestimaram o valor de precipitação (50 mm), devido à falta de medidores nos locais de ocorrência. De acordo com pluviômetros particulares, instalados nas áreas rurais, a precipitação que deflagrou os movimentos de massa teria sido maior (valores entre 90 e 272 mm) [6].

A duração estimada do evento de precipitação foi de 4 h, a qual resultou em diversos movimentos de massa. O material deslocado para o fundo do vale formou barreiras naturais ao longo do leito do rio Mascarada, que, após rompimento, gerou uma enxurrada de grandes proporções a jusante, atingindo o município de Rolante [6].

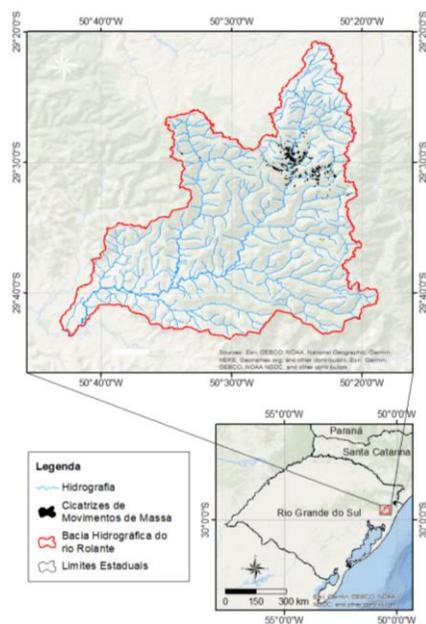


Figura 1. Localização da área de estudo.

## 2.2) Métodos

A metodologia foi dividida em cinco etapas: i) mapeamento de cicatrizes de movimentos de massa; ii) extração de parâmetros morfométricos de MDE; iii) pré-processamento e organização do conjunto amostral; iv) treinamento das RNAs; v) validação dos modelos.

O inventário de cicatrizes de movimentos de massa foi obtido a partir da interpretação de imagens de junho e agosto de 2017, disponíveis no *software* Google Earth, com conseguinte vetorização dessas áreas. Os polígonos foram exportados para o formato *shapefile*, visando o processamento em um Sistema de Informação Geográfica.

Na segunda etapa, foi obtido o MDE do satélite ALOS (*Advanced Land Observing Satellite*), sensor PALSAR (*Phased Array type L-band Synthetic Aperture Radar*), a partir do *website* Alaska Satellite Facility. Este MDE apresenta uma resolução espacial de 12,5 m. Utilizando os *softwares* SAGA-GIS e ArcGIS, foram extraídos 11 atributos morfométricos do MDE: elevação (ELV); declividade (DCL); orientação do terreno (ORT); curvatura horizontal (CVH); curvatura vertical (CVV); relevo sombreado (RLS); fluxo acumulado (FAC); fator LS (LS); índice de saturação do solo (TWI); profundidade do vale (PRV); distância vertical à hidrografia (DVH).

Na terceira etapa, o inventário de cicatrizes foi convertido de polígonos para o formato de pontos, sobrepondo cada pixel das imagens de atributos do terreno. Em paralelo, foi gerado um polígono envolvendo o inventário, e obtidos pontos aleatórios externos às cicatrizes, os quais foram considerados como amostras de não ocorrência de movimentos de massa. Desta forma, o conjunto amostral foi estabelecido por pontos localizados no interior das cicatrizes (Classe=1) e pontos externos (Classe=0).

O modelo de RNA foi escrito no *software* MATLAB. O treinamento da rede foi realizado pelo método retropropagativo com múltiplas camadas [7], usando a Regra Delta para atualização dos pesos sinápticos da rede [8]. Os valores dos atributos de entrada foram escalonados a partir de transformações lineares. A função de ativação utilizada nas camadas do modelo foi a sigmoideal. Uma série de validação cruzada paralela à série de treinamento foi utilizada para evitar o superajustamento do modelo.

Foram testadas oito configurações de modelo, variando o conjunto de atributos de entrada, desde um modelo mais completo, com todos os atributos, até modelos mais compactos com duas, três ou quatro variáveis de entrada. Para cada configuração de modelo, foram testados diferentes números de neurônios na camada oculta da rede. Uma vez que a inicialização dos pesos é aleatória, foram executadas 20 iterações para cada configuração do modelo. Em cada iteração, foi definido um número máximo de 8.000 ciclos de aprendizagem para interrupção, caso o modelo não atinja o limiar de convergência.

Para validação dos resultados, o índice de acurácia foi calculado por meio da *Area Under the Receiver Operating Characteristic Curve* (ROC). A curva ROC é uma representação bidimensional do desempenho do modelo. Uma vez que o modelo esteve ajustado, as amostras de teste foram preditas no intervalo entre 0 e 1. Posteriormente, as amostras foram divididas em vários grupos de duas classes por meio de limiares e para cada divisão foram calculadas a

Taxa de Verdadeiros Positivos e a Taxa de Falsos Positivos, formando assim, a curva ROC [9].

### 3. RESULTADOS E DISCUSSÕES

O primeiro resultado se refere ao mapeamento das cicatrizes, no qual foram identificadas 308 cicatrizes do evento ocorrido em janeiro de 2017. A área total dessas cicatrizes foi de 70,4 ha, sendo que a maior delas apresentou 3,6 ha de área. A maior parte das cicatrizes apresentou largura inferior a 15 m e comprimento na ordem de 200 m.

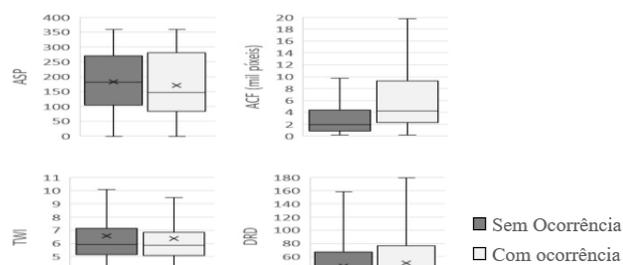
O conjunto amostral, obtido a partir das cicatrizes e dos pontos aleatórios externos a elas, resultou em 28.659 amostras centralizadas em pixels do MDE. Destes, 19.659 se referem a áreas de ocorrência (1) e 9.000 se referem a locais onde não ocorreram (0).

Para determinação das combinações de entradas no modelo de RNA, foram calculados os coeficientes de correlação linear entre os atributos do terreno e as classes de ocorrência e não ocorrência de movimentos de massa (Quadro 1). A maior correlação foi obtida para os atributos relativos a declividade (0,57), fator LS (0,57), profundidade dos vales (0,52) e elevação (-0,49).

**Quadro 1. Taxa de correlação de atributo e área de ocorrência de movimentos de massa.**

	ELV	DCL	ORT	CVH	CVV	RLS	FAC	LS	TWI	PRV	DVH
CLASSE	-,49	,57	-,05	-,20	,25	,23	,04	,57	-,03	,52	,04

Os quatro parâmetros que apresentaram menor taxa de correlação foram analisados por meio de gráficos do tipo *boxplot*, nos quais é possível verificar a distribuição dos valores de cada atributo tanto em áreas de ocorrência de movimentos de massa, quanto naquelas de não ocorrência (Figura 2). Ao analisar os gráficos, fica evidente que os atributos ORT, TWI e DVH apresentam distribuição muito semelhante para as duas classes de amostras, indicando que essas variáveis não contribuem estatisticamente para separar as áreas de movimentos de massa das demais amostras.



**Figura 2. Gráficos de caixa (boxplots) para os atributos morfométricos com menor correlação linear em relação às amostras de ocorrência e não ocorrência.**

Na Tabela 1 são apresentados os índices de acurácia (AUC) calculados para as séries de treinamento, validação cruzada e teste das RNAs, considerando a melhor configuração obtida para cada combinação de dados de entrada. Observa-se que, independente do conjunto de atributos usados e da complexidade da RNA, o AUC foi sempre superior a 0,9, indicando um excelente grau de acerto no mapeamento da suscetibilidade a movimentos de massa.

A RNA 01, configuração mais complexa, apresentou a maior acurácia na série de testes e verificação do modelo (0,948), sendo composta por 11 atributos de entrada, 21 neurônios na camada oculta e 274 conexões ou pesos sinápticos. Outras configurações destacadas na Tabela 1 se referem às RNAs 02, 03 e 05, todas com AUC igual ou superior a 0,92, tendo um número reduzido de variáveis de entrada e, respectivamente, 208, 193 e 145 pesos sinápticos. Essas configurações são mais parcimoniosas, uma vez que preservam um bom desempenho no mapeamento dos movimentos de massa, com um número reduzido de parâmetros da RNA.

**Tabela 1. Acurácia dos modelos de RNA utilizados.**

RNA	Input	Hidden	Con.	AUC Tr.	AUC Val.	AUC Tes.
01	11	21	274	0,959	0,951	<b>0,948</b>
02	07	23	208	0,942	0,933	<b>0,933</b>
03	06	24	193	0,941	0,932	<b>0,932</b>
04	05	20	141	0,924	0,919	0,918
05	04	24	145	0,923	0,916	<b>0,920</b>
06	03	12	61	0,916	0,912	0,916
07	03	05	26	0,905	0,907	0,906
08	02	14	57	0,900	0,904	0,903

**Em que RNA se refere a configuração da modelagem; Input, número de variáveis de entrada; Hidden, número de neurônios na camada oculta; Con., número de conexões ou pesos sinápticos, que reflete a complexidade da RNA; AUC, acurácia nas séries de treinamento (Tr.), de validação cruzada (Val.) e de teste (Tes.).**

As quatro melhores configurações de RNA foram selecionadas para análise visual dos mapas de saída, a partir da espacialização dos dados (Figura 3). Os mapas resultantes foram classificados em cinco graus de suscetibilidade a movimentos de massa, considerando o intervalo de 0 a 1.

Observa-se que a RNA 01, a qual apresentou maior acurácia, delimitou como área altamente suscetível a movimentos de massa (de 0,9 a 1) inclusive áreas de planícies e terraços fluviais, algo equivocado em relação a esse tipo de evento extremo. Esse modelo, apesar de apresentar o melhor desempenho no conjunto de amostras de teste, não apresenta uma boa capacidade de generalização e/ou extrapolação, tendo dificuldade para modelar áreas mais distantes das cicatrizes usadas no treinamento da rede.

Por outro lado, as outras três configurações de RNAs (Figura 3) não apresentaram o padrão incoerente observado na saída do modelo RNA 01. São configurações que consideram apenas as variáveis mais correlacionadas com as

classes de ocorrência e não ocorrência. A RNA 02 desprezou os atributos ORT, TWI, FAC e DVH. A RNA 03 é uma simplificação da anterior, excluindo a variável LS, uma vez que essa apresenta elevado grau de correlação com o atributo de declividades (DCL). Por sua vez, a RNA 05 é uma configuração que considerou apenas os quatro atributos mais relevantes: ELV, DCL, LS e PRV.

Os resultados obtidos neste estudo reforçam as conclusões apresentadas em [10], de que é importante realizar uma pré-seleção dos atributos de entrada dos modelos de RNAs, de modo a eliminar variáveis redundantes ou com baixa correlação com a saída do modelo. A maior complexidade da RNA nem sempre resulta modelos com boa capacidade de generalização, sendo necessária uma análise visual dos resultados para escolher aqueles que resultam no melhor mapeamento de áreas suscetíveis a movimentos de massa.

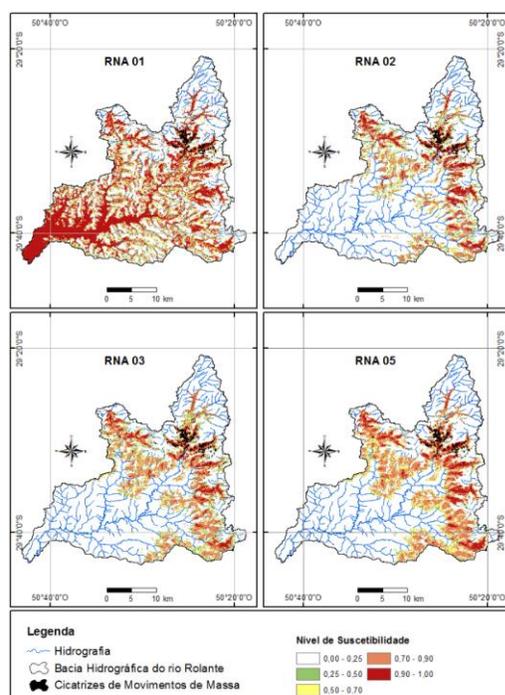


Figura 3. Mapas de suscetibilidade a movimentos de massa de acordo com o modelo de maior AUC de cada RNA.

#### 4. CONCLUSÕES

Com base nos objetivos propostos, na metodologia aplicada e nos resultados obtidos, as principais conclusões são:

i) Dos atributos extraídos do MDE ALOS-PALSAR, a ELV, a DCL, o fator LS e a PRV foram os que apresentaram maior correlação com as áreas de ocorrência de movimentos de massa;

ii) Alguns atributos amplamente usados na literatura científica, como o ASP, ACF, TWI e DRD, não

apresentaram um padrão bem definido nos movimentos de massa ocorridos em 2017 na bacia do rio Rolante;

iii) As RNAs se mostraram capazes de modelar as áreas suscetíveis, com acurácia no mapeamento superior a 0,9;

iv) A análise visual é de suma importância, uma vez que modelos com elevada acurácia podem não apresentar boa capacidade de generalização.

#### 5. AGRADECIMENTOS

À FAPERGS, pelo apoio financeiro, processo 17/2551-0000894-4, Edital 01/2017 e à CAPES, pelas bolsas de mestrado e doutorado.

#### 6. REFERÊNCIAS

[1] CEPED UFSC. “Atlas Brasileiro de Desastres Naturais – 1991 a 2012. Volume Brasil”. Florianópolis: [s. n.], 2013.

[2] Riffel, E. S.; Ruiz, L. F. C. e Guasselli, L. A. “Mapeamento de suscetibilidade a deslizamentos a partir de Mineração de Dados e do modelo SHALSTAB”, *Revista Brasileira de Cartografia*, n. Ed. 68/9 – Especial Movimentos de Massa e Processos Erosivos, 2016.

[3] Tsangaratos, P.; Ilia, I; Hong, H.; Chen, W. e Xu, C. “Applying Information Theory and GIS-based quantitative methods to produce landslide susceptibility maps in Nancheng County, China”. *Landslides*, v.14, pp. 1091 – 1111, 2017.

[4] Zêzere, J. L.; Pereira, S.; Melo, R.; Oliveira, S. C.; e R. A. C., Garcia. “Mapping landslide susceptibility using data-driven methods”. *Science of the Total Environment*, v. 589, pp. 250-267, 2017.

[5] Kawabata, D. e Bandibas, J. “Landslide susceptibility mapping using geological data, a DEM from ASTER images and na Artificial Neural Network (ANN)”. *Geomorphology*, 113, pp. 97-109, 2009.

[6] SEMA. “Diagnóstico Preliminar: Descritivo dos eventos ocorridos no dia 5 de janeiro de 2017 entre as regiões dos municípios de São Francisco de Paula e Rolante/RS”. Secretaria do Ambiente e Desenvolvimento Sustentável. Porto Alegre, pp. 26. 2017.

[7] Rumelhart, D. E.; Hinton, G. E. e Williams, R. J. “Learning representations by back-propagating errors”. *Nature*, 323, pp. 533-536, 1986.

[8] Widrow, B. e Hoff, M. E. “Adaptive Switching Circuits”. 1960 IRE WESCON Convention Record, New York: IRE Part, pp. 96-104, 1960.

[9] Fawcett, T. “An introduction to ROC analysis”. *Pattern Recognition Letters*, 27, 8, pp. 861-874, 2006.

[10] Oliveira, G. G.; Pedrollo, O. C. e Castro, N. M. “Simplifying artificial neural network models of river basin behaviour by an automated procedure for input variable selection”. *Engineering Applications of Artificial Intelligence*. 40, pp. 47-61, 2015. *Intelligence*, 40, 47-61, 2015.