

SUSCEPTIBILITY MAPPING OF LINEAR EROSION PROCESSES USING OBJECT-BASED ANALYSIS OF VHR IMAGES

D. P. Passo^a, E. S. Bias^a, R. S. Brites^a, G. A. O. P. Costa^b, R. R. Antunes^a

^a Federal University of Brasilia, Brazil – geodenilson@gmail.com, edbias@gmail.com, brites.ricardo@gmail.com and rodrigorantunes@hotmail.com

^b Rio de Janeiro State University – gilson.costa@ime.uerj.br

ABSTRACT:

Linear erosion is a natural phenomenon. However, inadequate occupation of the environment or the implementation of engineering works, without the due care, accelerates this process, which has been acknowledged as the main cause of land degradation worldwide. The use of high-resolution satellite imaging to map risk areas for this process, may contribute to devising prevention strategies. Linear erosion is a process dependent on thresholds controlled by many variables. This study has used only topographic variables (altimetry, slope, curvature profile, curvature plan, slope orientation, accumulation flow, humidity index, sediment transport capacity, potential flow and drainage network) and a vegetation index, which were selected due to their influence on linear erosion processes. The study was developed in two 6,000 x 4,500 meter areas, located in the eastern part of the Federal District – Brazil. The classification model building was done using open source software packages, namely InterIMAGE and WEKA. The aim of this study was to develop a routine for automatic mapping of areas susceptible to linear erosion. The accuracy rate achieved by the model was 87.5%, as 21 of 24 linear erosion processes were identified. The percentage of the mapped area in relation to the total study area also showed that the classification was not overestimated.

KEYWORDS: RapidEye, Segmentation, Linear Erosion, Digital Elevation Model.

1. INTRODUCTION

The erosion process represents a major environmental impact caused by inappropriate occupation of rural and urban areas. This process results from the disintegration and natural movement of the particles of the soil; and it can be accelerated by anthropic action. It may occur when, for example, the original vegetation is removed and the uncovered soil becomes more vulnerable. It also happens when the appropriate conservation techniques are not used to manage the soil, causing imbalances in the interaction between soil, climate, relief and vegetation, and also instability in the system as a whole.

Studies regarding erosion processes have been developed by many researchers for a long time, usually based on empirical equations regarding the loss of soil layers. According to Lanza (2011), there are many approaches to this analysis, such as: the creation of cartography of areas at risk, field observation and experiment activities, spatial incidence of centers of erosion features, and also estimation of soil loss rates.

There is, however, a gap in erosion studies concerning automatic mapping of linear erosion processes. Measurements need to be taken *on site* to assess and monitor these processes. For the measurements to be taken on site, a great number of people and a large amount of equipment is necessary on the edges and in the interior of the process, which may worsen it, and puts the researchers on the field at risk. When it comes to locating, assessing and monitoring these processes, a viable alternative is to use remote sensing based methods, such as automatic classification of high-resolution spatial imagery.

Some of the features that make linear erosion a different process with respect to other classes of erosion are: spectral heterogeneity (mixed types of soil, vegetation, and water), complexity of the target, caused by its size and varied shapes

(irregular shapes, variable dimensions, asymmetry, and variable width/length proportion), and altimetry variation (which can range from centimeters to meters, with steep declivities on the edges). So it is necessary to devise an approach that take these specific features into account.

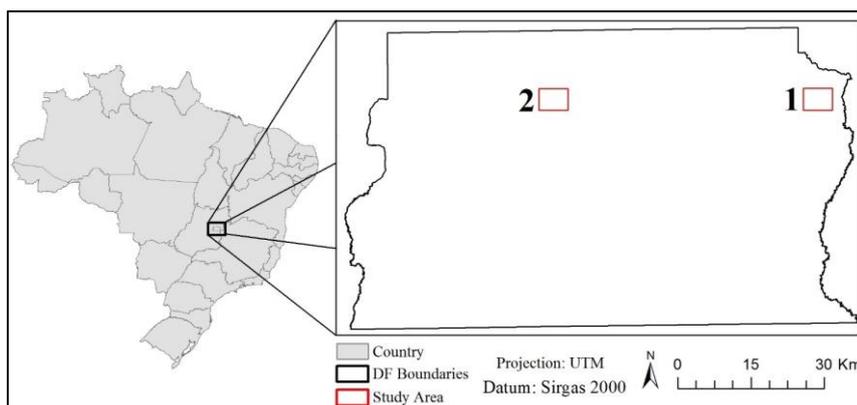
Among high-resolution remote sensing imagery classification approaches, Geographic Object-Based Image Analysis GEOBIA (Hay and Castilla, 2008) has gained relevance in the last years. This approach represents an improvement in respect to conventional automatic remote sensing analysis approaches, as it makes it possible to model human knowledge regarding specific characteristics of the target classes of objects, with the aid of features like brightness, texture, shape, proximity relations, among others, as well as being able to be integrate with other techniques, such as mining, *Fuzzy* logic, and fractal geometry. GEOBIA focuses on the classification of image segments, which is different from traditional classification approaches, which perform pixel-wise classification.

Information obtained with the object-based method proposed in this work provide an important support to environmental and territorial planning and managing. Such information helps in the efforts of locating the distribution and expansion of linear erosion processes, as well as making it possible to monitor changes in short intervals of time.

The objective of this study is to employ object-based classification on high-resolution images, using topographic variables to automatically map areas susceptible to linear erosion.

2. MATERIALS AND METHODS

The area under study is in Distrito Federal, one of 27 Brazilian states, located in the Midwestern region of the country, on the Central Plateau, with an area of 5,814 square kilometers.



Picture 1 - Location of test-areas.

The proposed methodology was applied in two test-areas (Picture 1). Each test-area is 6,000 X 4,500 meters in size: that is, 27,000 m². The delimitation of these areas took into account the existence of several ongoing erosion processes in the areas under study.

Test-area 1 was used for selecting samples to train the data mining procedure, so as to build a decision tree that aids mapping of ravines and gullies. Test-area 2 was chosen to test if the decision tree could be applied to an environment that was not influenced by the presence of training samples.

The following products were used to execute this study:

- RapidEye image, with 5 meters' spatial resolution, with a spectral range between 440 and 850 nm, and a 12-bit radiometric resolution;
- vector data: hydrography, height level curves, and critical points;
- Digital elevation model (DEM) generated by interpolation of topographic height level curves and critical points;
- DEM-generated data: altimetry, declivity, profile curvature, plan curvature, aspect, flow accumulation, humidity index, sediment transport volume, and drainage potential.

The following software packages were used to execute this study:

- ArcGIS v10.3 (ESRI, 2008);
- ENVI v5.3 (ITT, 2009);
- InterIMAGE v1.42 (Costa et al. 2010);
- WEKA v3.7 (Witten and Frank, 2005).

The proposed methodology starts by selecting and acquiring a high-resolution remote sensing system. The RapidEye imaging system was chosen because of its spatial resolution and because MMA – Ministry of the Environment of Brazil provides it at no cost for research projects.

TERRACAP – Land Development Company of Brasília provided vector data (level curves, critical points, and drainage network) that made possible the construction of the DEM. This data was gathered during a cartographic mapping of the Federal District, at 1:10,000 scale, undertaken by TOPOCART in 2010.

After generating the following classification image planes from the DEM: altimetry, declivity, profile curvature, plan curvature, aspect, flow accumulation, humidity index, sediment transport volume and drainage potential; those planes were cropped so they would fit the limits of the test-areas and could, then, be inserted on InterIMAGE as bands of a single, synthetic image, which was subjected to segmentation using the multi resolution segmentation algorithm proposed in (Baatz and Schäpe 2000).

Once the image segments were produced, segments representing linear erosion processes were selected as training samples for the data mining procedure. This selection was based on a mapping produced by Oliveira (2011) through visual interpretation of aerial and satellite imagery, that aimed to identify and analyze factors related to linear erosion processes in the Federal District.

After creating a single synthetic image by stacking the input image planes, segmentation and feature extraction was carried out with InterIMAGE. Then, a table was created, each line corresponding to a segment and each column, to a computed feature. The data thus obtained was transferred to the data mining software, through which the features were analyzed, and the most relevant features were selected to take part of in the classification model.

The J48 decision tree algorithm in the WEKA data mining software package was used to define a decision tree based on the grouping of specific segment features. The generated decision tree was then translated into decision rules in InterIMAGE, which was responsible for the final classification.

In order to assess classification quality, a verification based on a generic accuracy metric (*non-site specific accuracy*) was conducted. In this verification only the total area of the mapped category is analyzed, without taking the location of the category into account. In other words, it is a comparison (ratio) between the area of the category on the map (linear erosion susceptible areas in this case), which is generated by comparing the classification and the reference data (ground truth) (Congalton and Green, 1999).

3. RESULTS AND DISCUSSION

The segmentation process preceded that of classification. Segmentation parameter values were input on the operator TA_Baatz_Segmenter (Baatz and Schäpe 2000), which is capable of multi-resolution segmentation, by accepting values for the relative weights of the synthetic image bands, and for compactness, color, and scale. The “weight” parameter defines each image band’s relevance in the segmentation. In the experiments the same weight was defined for all bands. The “compactness” parameter represents the way the pixels in each segment are grouped, with a higher value meaning more compact objects. The “color” parameter refers to the relative importance of spectral information, i.e., pixel intensity values, in relation to the morphological characteristics of the segments, i.e., shape information. The “scale” parameter value roughly defines the size of the final image segments.

The J48 decision tree algorithm in the WEKA package can be tuned to constrain the model to be generated. Therefore, the parameter “MinNumObj” was set. This parameter controls the size and the complexity of the tree to be generated, and was set so that a minimum of 4 objects (segments) per leaf of the decision tree should be selected. This means that decisions that did not classify at least 4 segments were automatically eliminated from the model.

To build the decision tree (Table 1), the J48 algorithm selected eight features as the most relevant for the mapping of areas susceptible to linear erosion process. These features were altimetry, NDVI, profile curvature, plan curvature, topographic humidity index, sediment transport capacity index, proximity to drainage network, and aspect.

```

NDVI <= 0.26
|Altimetry <= 950.59
||Profile curvature <= -0.095
||| Profile curvature <= -0.204
||| Profile curvature > -0.204
||||Altimetry > 894.85
||| Profile curvature > -0.095
||| Prox. to drainage network <= 245.81
|||| Altimetry <= 918.79
|||||NDVI <= -0.005
|||||Altimetry <= 887.048
|||||Altimetry > 887.048
|||||Plan curvature > 0.156
|||||NDVI > -0.005
||||| Prox. to drainage network <= 55.20
||||| Plan curvature > 0.223
|||||Altimetry > 918.799
||||| Profile curvature > 0.13
| Altimetry > 950.59
|NDVI <= 0.017
|||Topographic humidity index > 4.399
||| Prox. to drainage network <= 16.126
||| Sediment transport capacity > 2.023
||| Altimetry <= 1208.53
|||Aspect > 129.75
|||Altimetry > 992.37
|||Altimetry > 1208.536
    
```

Table 1. Tree in textual form, generated by the J48 algorithm.

The altimetry feature was then used as an indirect measurement of spatial variations in temperature and/or precipitation, as well as to estimate spatial distribution of climatic variables, such as average seasonal and annual precipitation, maximum daily precipitation, monthly average and temperature (Moore et al. 1991). This justifies the feature being chosen by the data mining algorithm, because, according to Pruski (2006), rain is the agent responsible for providing the necessary energy for hydric erosion to occur, because of the direct impact of water drops on the surface of the soil and also because of its capability of producing superficial drainage..

The data mining algorithm selected NDVI as its first criterion to divide the areas (segments) susceptible to linear erosion processes, which confirms Lima (2003), who states that the presence of vegetation is a very important characteristic in the study of erosion, because it is a factor related to the direct protection of the soil against the impact of water drops. Vegetation also stops erosive action originated from superficial drainage.

The profile curvature and plan curvature are present because both are a characteristic of the terrain associated with hydrological, pedological, and transport of solid matter, as well as many other aspects, indirectly. The combination of both curvatures is represented by a concave-convergent form (maximum drainage concentration and accumulation) and by a convex-convergent form (maximum drainage dispersion), as described by Araújo (2006).

Even though the declivity feature is defined as inclination of slopes, which causes locations with a higher declivity to have more frequent linear erosion processes, the data mining algorithm did not select it directly. It was used indirectly, nevertheless, because it was part of the calculation to obtain other variables taken into account, such as: topographic humidity index and sediment transport capacity index.

The aspect variable was also selected by the data mining algorithm, which shows the importance of this feature in the mapping. Proof of this importance is the fact that this feature is used as local attribute in geological investigations, and may indirectly influence linear erosion processes by controlling exposure to different climatic conditions (time of exposure to sunlight, precipitation intensity, humidity retention, etc.) and vegetation (Dai et al., 2001; Çevik and Topal, 2003; Pulice et al., 2009; Conforti et al., 2010).

The values of the flow accumulation and declivity features were used to calculate other features taken into account, such as the humidity topographic index and the sediment transport capacity index.

The fact that the data mining algorithm selected the humidity index feature may be explained by the relationship between this index and saturated areas, as well as the generation of superficial drainage.

The feature concerning proximity to drainage network was also chosen because a relationship between this variable and linear erosion processes was inferred, which confirms Werlang (2004), who states that the relation between headwaters and linear erosion processes may be observed through the similarity between the concepts of incised river channels in valley headwaters and linear erosion processes. Incised river channels are the upper limit of flow and

sediment transport. Linear erosion processes are natural incised channels that originate from natural or anthropic imbalances.

For test-area 1 (Figure 1), where training samples were selected, the classification model achieved a 95.83% accuracy value for the mapping in relation to the amount of linear erosion processes identified. This means the mapping was able to find 46 of 48 existing linear erosion processes.

Test-area 2 was used for confirmation, and it was analyzed using the decision tree defined by the training samples obtained in test-area 1. The accuracy rate achieved by the model in Test-area 2 was 87.5%, that is, 21 of 24 linear erosion processes were identified (Figure 2). The percentage of the mapped erosion susceptible area in relation to the total area of the test-area was 17.5% which shows that the classification was not overestimated.

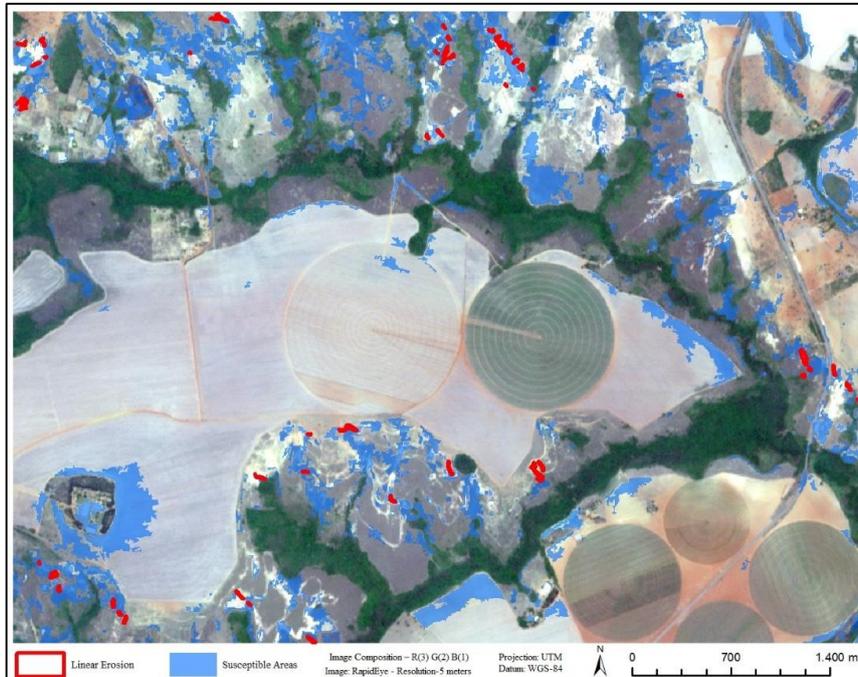


Figure 1. Map of susceptibility to linear erosion processes in test-area 1.

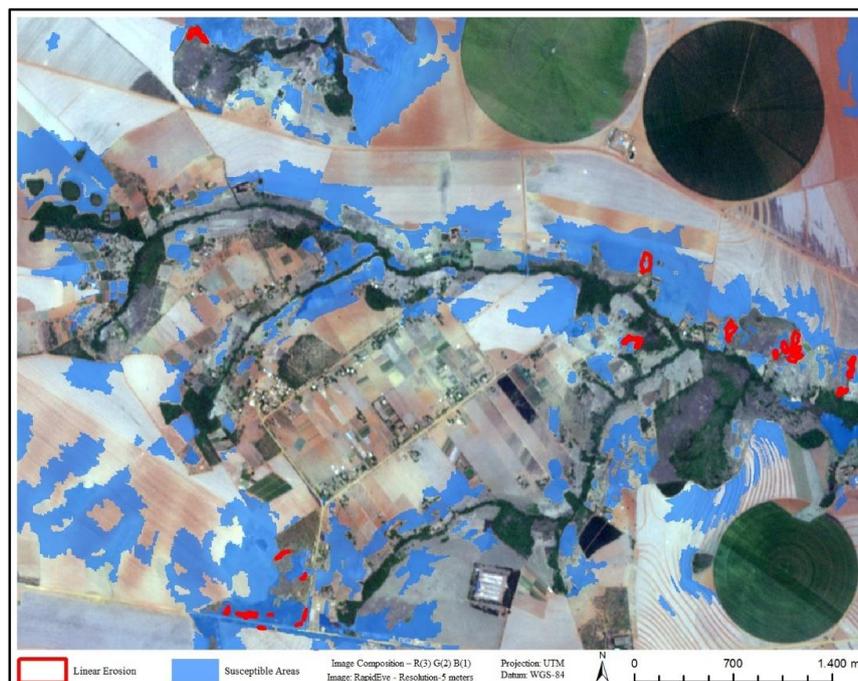


Figure 2. Map of susceptibility to linear erosion processes in test-area 2.

4. FINAL CONSIDERATIONS

The association of object-based analysis and data mining has proven to be very effective for identifying areas susceptible to linear erosion processes.

The areas that were identified by the model coincided with a significant number of existing linear erosion processes, which led to high accuracy rates. As a result, using topographic attributes to map areas susceptible to linear erosion processes can be considered a coherent and effective method, and represents a useful tool for studies of this nature.

The study of topographic aspects has shown the strong relationship between physical elements of the terrain, hydrological processes, and, consequently, linear erosion processes. This relationship made it possible to devise an object-based methodology that delivered high accuracy rates. Due to the high accuracy obtained, the methodology has been found recommendable for the mapping of areas susceptible to linear erosion processes.

REFERENCES

- Araújo, E. P., 2006. Aplicação de dados SRTM à modelagem da erosão em microbacias por geoprocessamento. Master's Thesis, São José dos Campos: Instituto Nacional de Pesquisas Espaciais (INPE).
- Baatz, M., Schäpe, A., 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In: XII Angewandte Geographische Informationsverarbeitung, AGIT Symposium. Proceedings. Karlsruhe, Germany: Herbert Wichmann Verlag, Salzburg – Austria, pp. 12-23.
- Çevik, E., Topal, T., 2003. GIS-based landslide susceptibility mapping for a problematic segment of the natural gas pipeline, Hendek, Turkey. *Environ Geol*, n. 44, pp. 949-962.
- CODEPLAN. Companhia de Desenvolvimento do Planalto Central. (1971). Diagnóstico do Espaço Natural do Distrito Federal. Brasília: Edição CODEPLAN.
- Congalton, R. G., Green, K., 1999. Assessing the accuracy of remotely sensed data: principles and practices. New York: Lewis Publisher.
- Costa, G.A.O.P., Feitosa, R.Q., Fonseca, L.M., Oliveira, D.A.B., Ferreira, R.S., Castejon, E., 2010 Knowledge-Based Interpretation of Remote Sensing Data with the Interimage System: Major Characteristics and Recent Developments. In: Proceedings of the 3rd International Conference on Geographic Object-Based Image Analysis - GEOBIA 2010. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVII. ITC, Enshede. 2010.
- Dai, F. C. et al. 2001. Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. *Environ Geol*, n. 40, pp. 381-391.
- Environmental Systems Research Institute (ESRI), 2015. ArcGIS Professional GIS for the desktop.
- Hay, G. J., Castilla, G., 2008. Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline? In: Blaschke, T., Lang, S., Hay, G. J. (Eds.) Object-based image analysis spatial concepts for knowledge-driven remote sensing applications. Berlin: Springer, pp. 75-89.
- ITT Visual Information Solutions, 2009. ENVI 4.7, version 4.7. Boulder, Colorado, EUA.
- Lanza, D. S., 2011. Diagnóstico da erosão laminar na alta e média bacia do rio Paraopeba. Master's Thesis, Belo Horizonte: Universidade Federal de Minas Gerais.
- Lima, E. R. V., 2003. Erosão do solo: fatores condicionantes e modelagem matemática. Cadernos do LOGEPA. Série pesquisa. João Pessoa: Universidade Federal da Paraíba.
- Moore, I. D., Grayson, R. B., Ladson, A. R., 1991. Digital terrain modeling: A review of hydrological, geomorphological and biological applications. *Hydrological Processes*, v. 5, pp. 3-30.
- Oliveira, B. E. N., 2011. Mapeamento, identificação e análise dos fatores relacionados aos processos erosivos no Distrito Federal (DF) – Ênfase nas Voçorocas. Master's Thesis, Brasília: Universidade de Brasília.
- Pruski, F. F., 2006. Conservação do solo e da água: práticas mecânicas para o controle da erosão hídrica. Viçosa: UFV, pp. 131-171.
- Pulice, I. et al. 2009. Studio multidisciplinare di forme e processi denudazionali nell'area di Vrica (Calabria orientale). *Bollettino della Societa Geografica Italiana* 87(I-II), pp. 399-414.
- Werlang, M. K., 2004. Configuração da rede de drenagem e modelado do relevo: conformação da paisagem na zona de transição da bacia do Paraná na Depressão Central do Rio Grande do Sul. PhD Thesis, Santa Maria-RS: Universidade Federal de Santa Maria.
- Witten, I., Frank, E., 2005. Data Mining: Practical Machine Learning Tools and Techniques. 2. ed. San Francisco, CA: Morgan Kaufmann Publishers.