RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

Geobia, tree decision and hierarchical classification for mapping gully erosion

Geobia, árvore de decisão e classificação hierárquica para mapeamento de voçoroca

Andrea Tedesco [*], **Alzir Felippe Buffara Antunes**[**]

[*] Departamento de Geociências, Universidade Estadual de Ponta Grossa – UEPG, tedescodeda@hotmail.com
[**] Departamento de Geomática, Universidade Federal do Paraná- UFPR, felipe@ufpr.br

DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

Abstract

The gullies provoke environmental, social and financial damages. The application of corrective and preventive measures needs gullies mapping and monitoring. In this scope, this study proposes a methodology for gullies delimitation using object-oriented image analysis. For such, there were used high spatial resolution imagery and ALS data applied for two study areas, one in Uberlandia-Minas Gerais (Brazil) and another one in Queensland (Australia). The objects were generated by multiresolution segmentation. The most important attributes on the delimitation of the gullies were selected using decision tree induction algorithms, being them: spectral, altimetric and texture. Classifications by decision trees and hierarchical were carried out. The use of decision tree allowed the selection of attributes and the establishment of preliminary decision rules. However, since this procedure did not use fuzzy logic, mixtures between classes could not be evidenced in the rule base. Moreover, the classification was performed by a factor of scale only, which did not allow the identification of all the constituent features of the gully. In hierarchical classification, the procedure is performed on different scales, allowing the use of fuzzy logic to describe different degrees of membership in each class, which makes it a very attractive method for cases such as this study, where there is mixing of classes. The classification obtained with hierarchical classification it was more reliable with the field truth, by allowing the use of different scales, uncertainty insert and integration of knowledge, compared to the automatic classification by decision tree.
KEYWORDS: ALS Data; High Resolution Imagery; Multirresolution Segmentation.

Resumo

As voçorocas provocam danos ambientais, sociais e financeiros. A aplicação de medidas corretivas e preventivas necessita do mapeamento e monitoramento das voçorocas. Nesse escopo, este estudo propõe uma metodologia para delimitação de voçorocas usando análise de imagens orientada a objetos. Para tanto, foram utilizados imagens de alta resolução espacial e dados ALS, para duas áreas de estudo, uma em Uberlândia- Minas Gerais (Brasil) e outra em Queensland (Austrália. Os objetos foram gerados por segmentação multirresolução. Os atributos mais importantes na delimitação das voçorocas foram selecionados utilizando algoritmos de indução de árvores de decisão, sendo eles: espectral, altimétrico e de textura. Foram realizadas classificações

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

por árvores de decisão e hierárquicas. A árvore de decisão permitiu selecionar atributos e estabelecer regras preliminares de decisão. No entanto, como esse procedimento não utilizou lógica fuzzy, misturas entre classes não puderam ser evidenciadas na base de regras. Além disso, a classificação foi realizada apenas com um fator de escala, o que não permitiu a identificação de todas as características constituintes da voçoroca. Na classificação hierárquica, o procedimento é realizado em diferentes escalas, permitindo que o uso da lógica fuzzy descreva diferentes graus de associação em cada classe, o que o torna um método muito atraente para casos em que há mistura de classes. A classificação obtida com classificação hierárquica foi mais similar à realidade de campo, ao permitir o uso de diferentes escalas, inserção de incerteza e integração do conhecimento, em relação à classificação automática por árvore de decisão.
Palavras-Chave: Dados ALS, Imagens de Alta Resolução, Segmentação Multirresolução.

I. INTRODUCTION

The gullies are the biggest erosive processes and, consequently, responsible for environmental, social and financial damages. Corrective and preventive measures need mapping and monitoring, which can be made by local measurements or by remote sensing. Local measurements can be done by staking (HESSEL and VAN ASCH, 2003; MORGAN, 2005), by topographic surveys, by GNSS receivers, or using TLS (Terrestrial LASER Scanning) (PERROY et al., 2010). However, these methods needs equipment installation on edges and inside the gullies, which can aggravate erosive processes and it can be a risk for surveyors.

Monitoring by remote sensing has been carried out by using aerophotos (MARZOLFF and POESEN, 2009; WANG et. al, 2014), multiespectral images (KING et al., 2005; VRIELING, RODRIGUES and STERK, 2005), DTM (Digital Terrain Model) (MARTÍNEZ-CASASNOVAS, RAMOS and POESEN, 2004) and ALS - Airborne LASER Scanning data (JAMES, WATSON and HANSE, 2007). Recently, UAVs (Unmanned Aerial Vehicles) have been used to map gullies (WANG et al., 2016). GIS (Geographical Information Systems) are been used to generate reliable gully erosion susceptibility maps (ARABAMERI et al., 2018; ZABIHI et al., 2018). GIS combined with fractal analysis were used to determine the stability of gullies (REAL et al., 2020). Researches have used GEOBIA (Geographic-Object-Based Image Analysis) for detection, mapping, monitoring, volume calculation and predictive models of erosion risk (EUSTACE, PRINGLE and WITTE, 2009; SHRUTI et al., 2011 and 2015). Using Tree Decision it is possible to determine the susceptibility to landslides and mass movements as a function of topographic and geological variables, from DTM (ALKHASAWNEH et al, 2014). Other artificial intelligence techniques have been used to find the most important factors in gully modelling (BUI et al., 2019).

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

In relation to the remote sensing, the gully erosion presents spectral heterogeneity (soil, vegetation, shade and water mix), spatial heterogeneity (existence of features as head and canals with irregular forms and variable dimensions) and altimetry variation (with big slope on the edges). Due to spectral heterogeneity, the use of only spectral data is not enough, which makes it necessary to use auxiliary data, such as altimetry and texture data. Using auxiliary data is recommended apply data mining because of the large amount of data.

In this context, this study proposes a methodology for delimitation of gullies on image classification procedures based on GEOBIA, identifying attributes to establish a decision rule base. For such, there were used an Ikonos image, orthophotos and ALS altimetry and intensity data of an area located in Uberlandia - Minas Gerais – Brazil and other area located in Queensland - Australia. The objects were generated by multiresolution segmentation (FNEA-Fractal Net Evolution Approach method). The most important attributes in the gullies mapping were selected by decision tree, being these attributes spectral, altimetry and texture, and a classification by tree decision was carried out. The hierarchical classification was carried out and presented satisfactory results by allowing the use of different scale factors, uncertainty insert (by fuzzy logic) and integration of knowledge (the rule base) compared to the automatic classification by decision tree. As there was no available near infrared band for Australian ortophotos, it was not possible to generate NDVI. However, it is proposed a new index, obtained combining intensity image from ALS data and the red band. The IBCI (Intensity-Based Contrast Index), instead of NDVI, enhances the soil.

II. MATERIAL AND METHODS

STUDY AREAS

The experiment were performed for two study areas which have gully erosions: one located in Uberlandia, Minas Gerais – Brazil (figure 1) and another located in Queensland, Australia (figure 2).

DATA

For the Brazilian study area, there were used a 1 meter spatial resolution and 11 bits radiometric resolution Ikonos II image and ALS data from ALTM 2025 Optech (1 meter spatial resolution rasterized).

For the Australian study area, there were used 0,5 meter spatial resolution and 8 bits radiometric resolution orthophotos and ALS data from Riegl LMS-Q560 (0,5 meter spatial resolution rasterized).

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

The procedures were carried out by using ENVI 4.7 (Exelis Visual Information Solutions, Boulder, CO, USA), ALDPAT (Airborne LiDAR Data Processing and Analysis Tools) and eCognition™ Developer 8.8.

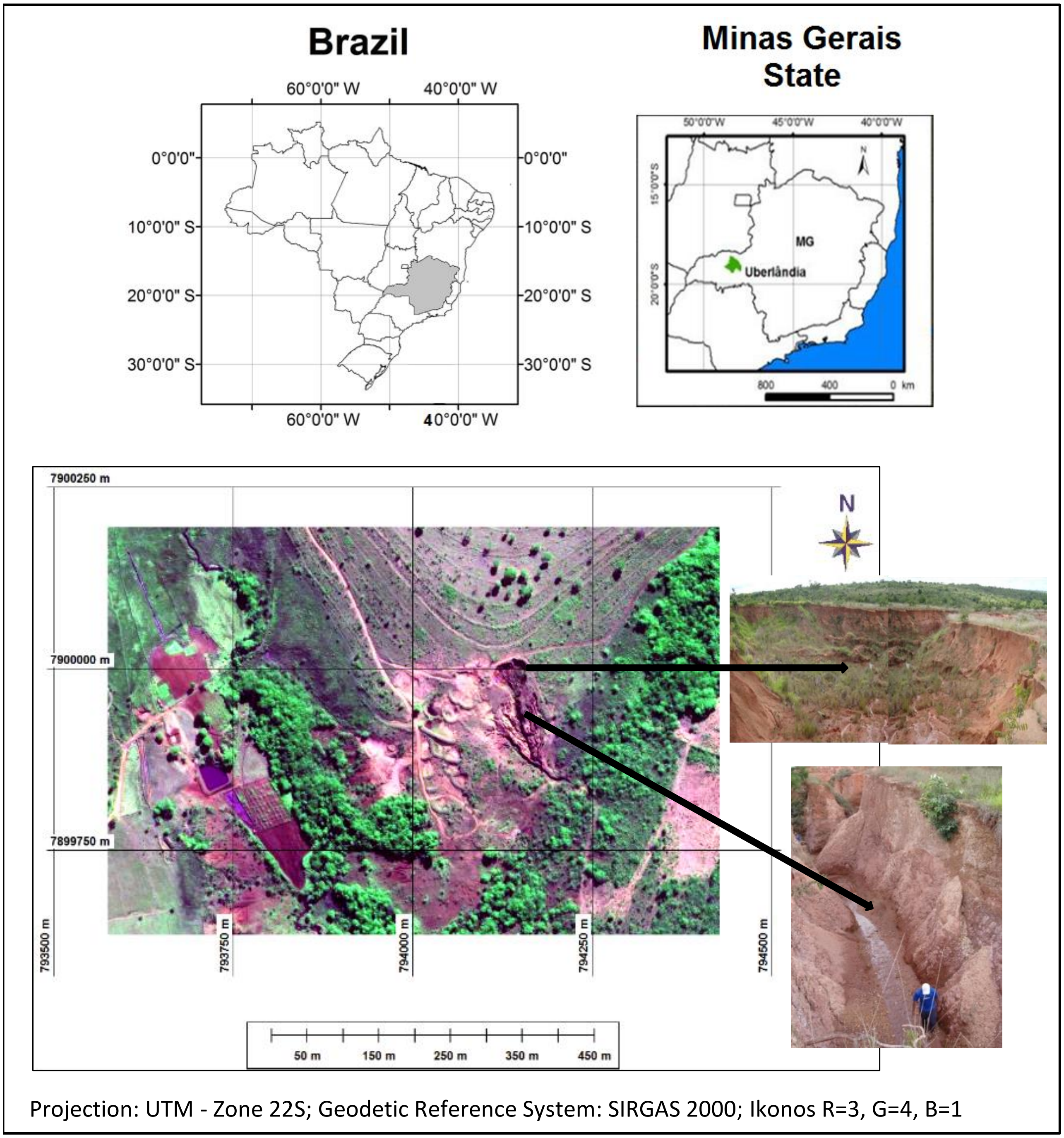
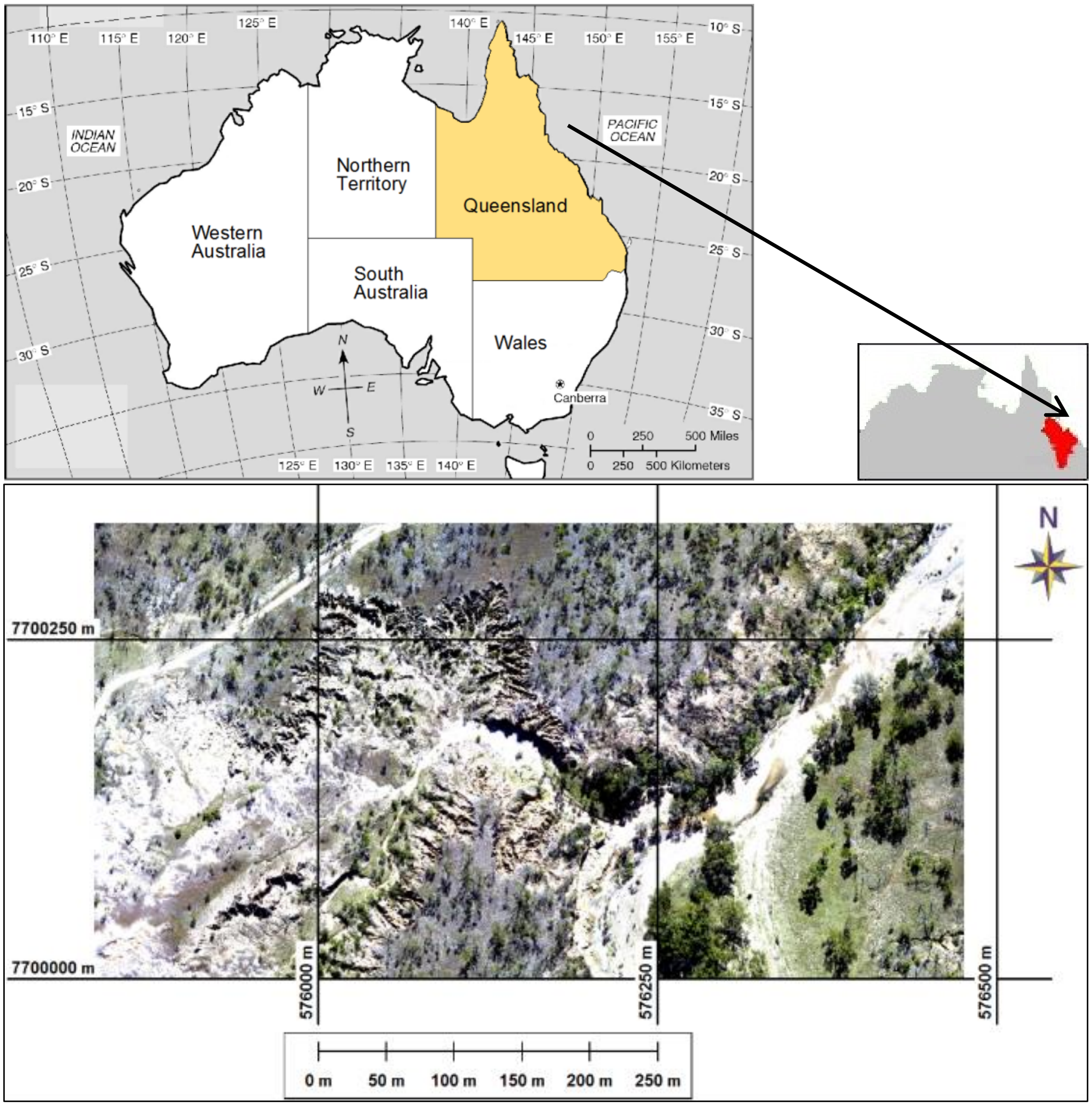


Figure 1: study area in Minas Gerais, brazil

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega



Projection: MGA1994 – Zone 55; Geodetic Reference System: GDA94; Ortophoto R=3, G=2, B=1

Figure 2: study area in Queensland, Australia

METHODOLOGY STEPS

The experiments were conducted as shown in diagram illustrated in figure 3.

Preprocessing

From the ALS data, there were generated the DSM (Digital Surface Model) and intensity image for both study areas, rasterizing point clouds by linear interpolation. From the DSM, there were generated the DTM using

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

the ALS last return for Australian data and using ATIN (Adaptive TIN Filter) filter for Brazilian data, the NDSM (Normalized Digital Surface Model) subtracting the DTM from the DSM using band math and the slope map.

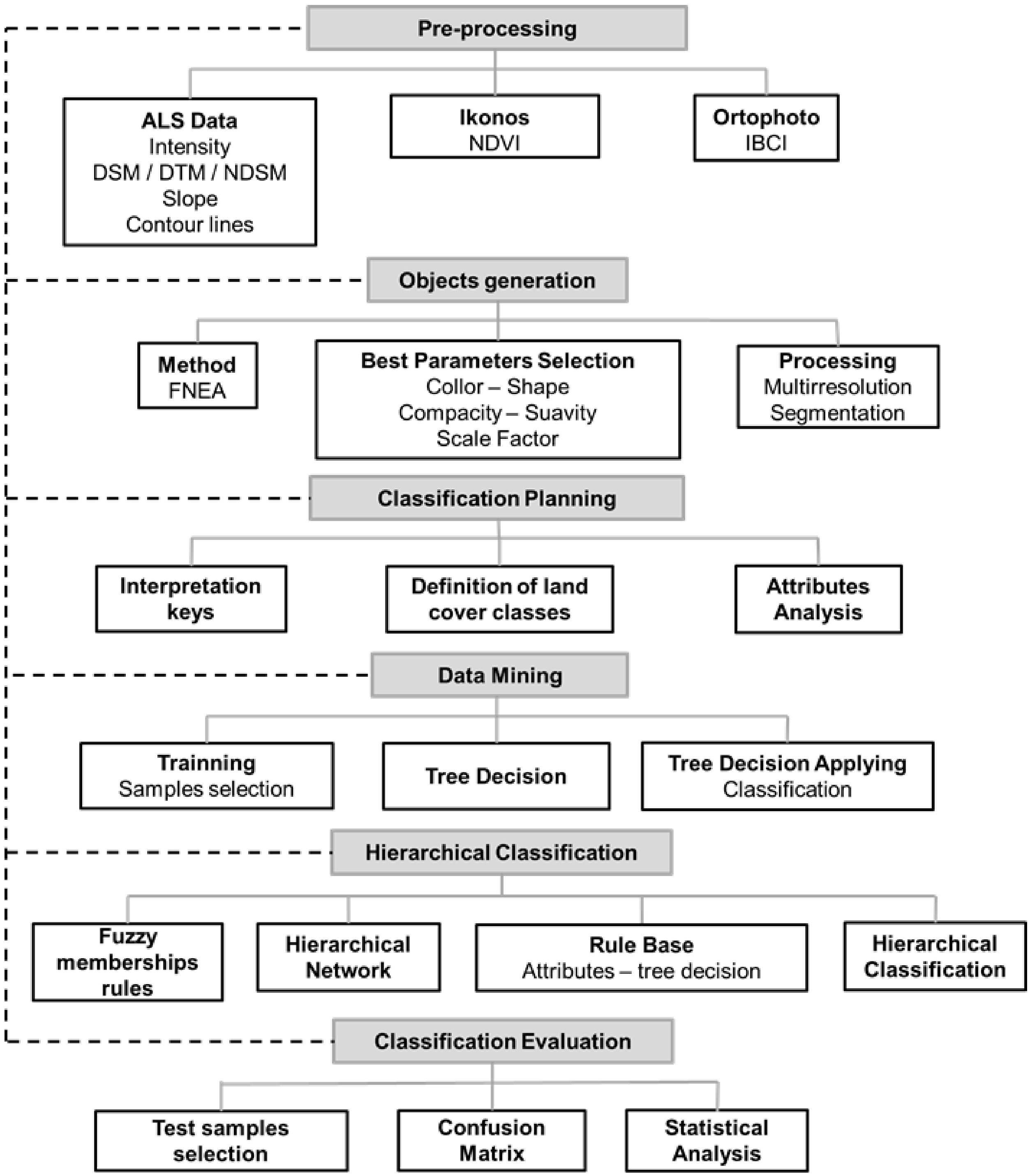


Figure 3: Methodology steps

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

The Figure 4 shows the DTM and the Figure 5 shows the slope map obtained for Brazilian study area. The figures 6 and 7 show the same for the Australian study area. It is possible to verify that altimetry data is an important auxiliary data in gully erosion mapping. In the slope map the gully system edge is evidenced because of its big slope.

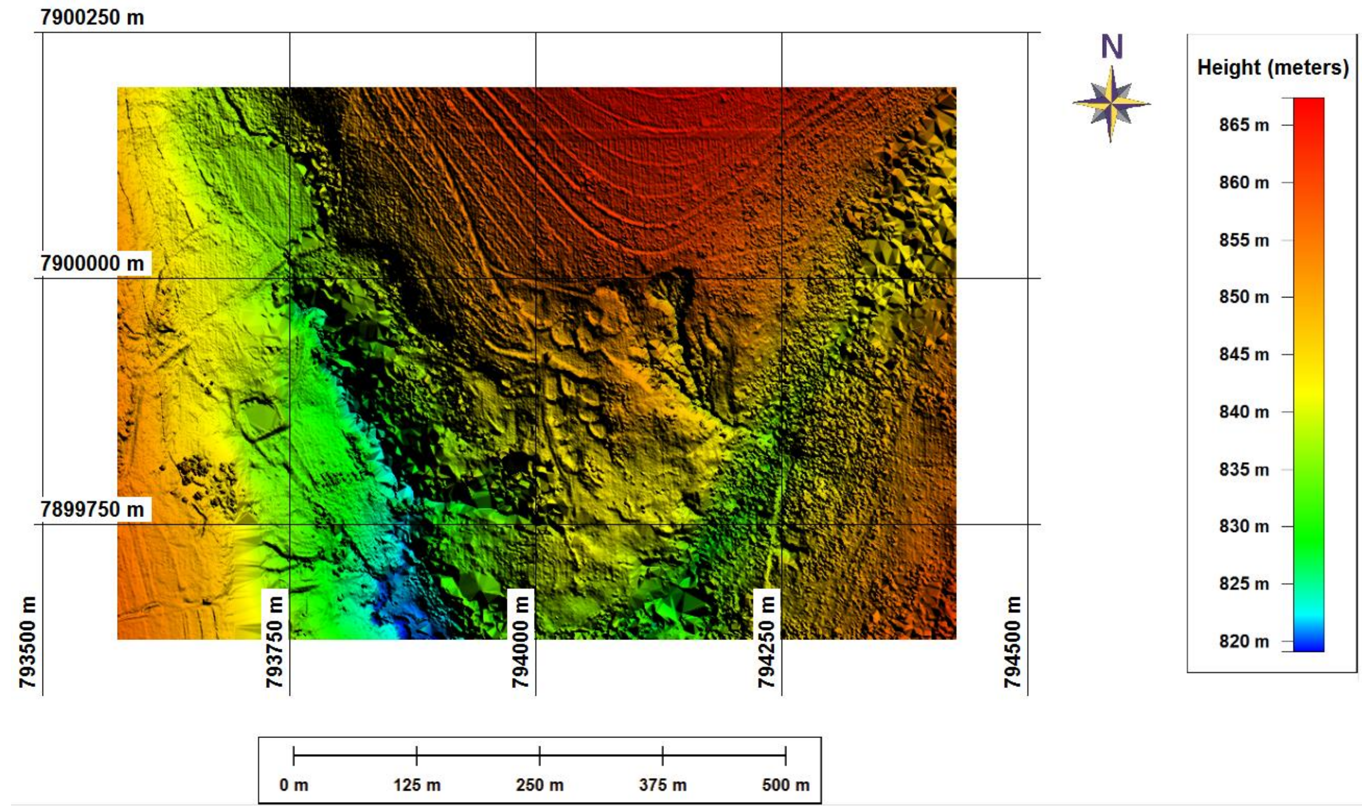


Figure 4: DTM - Brazilian study area

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

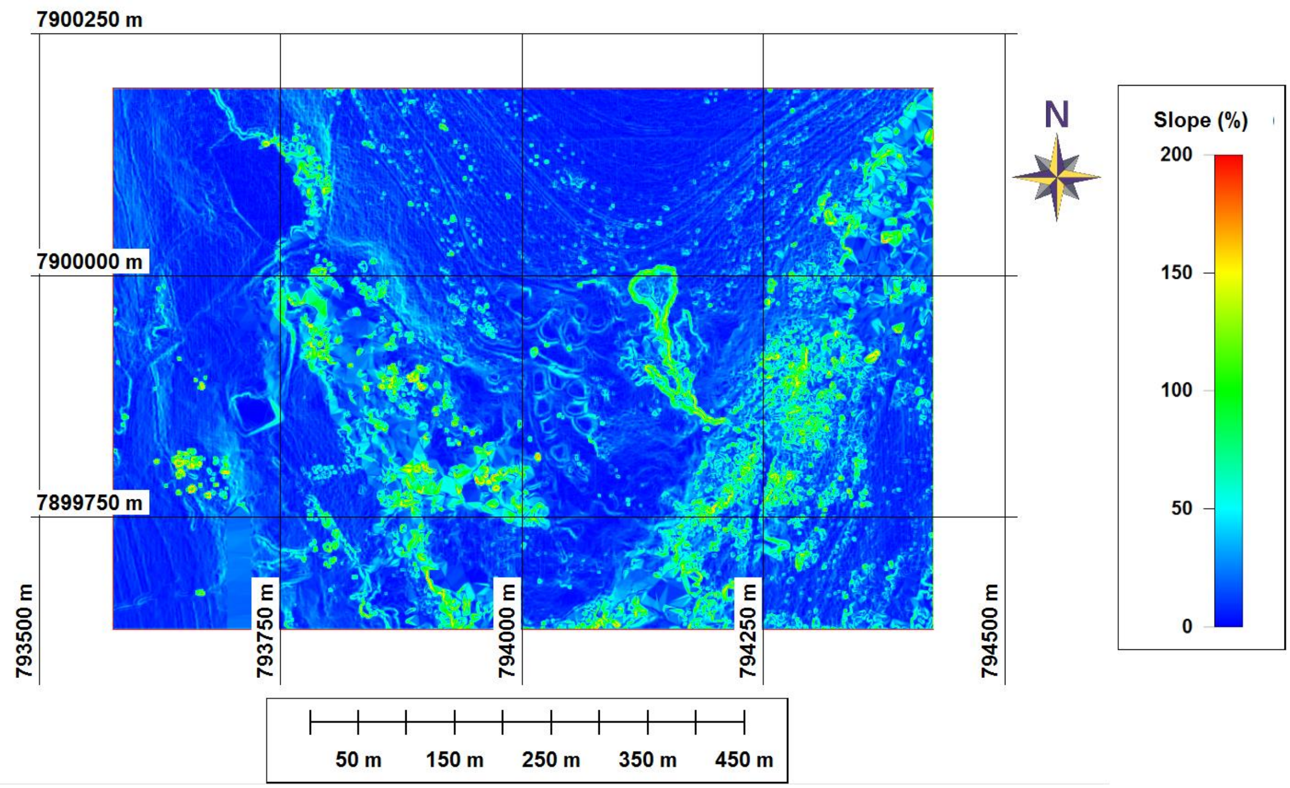


Figure 5: slope map - Brazilian study area

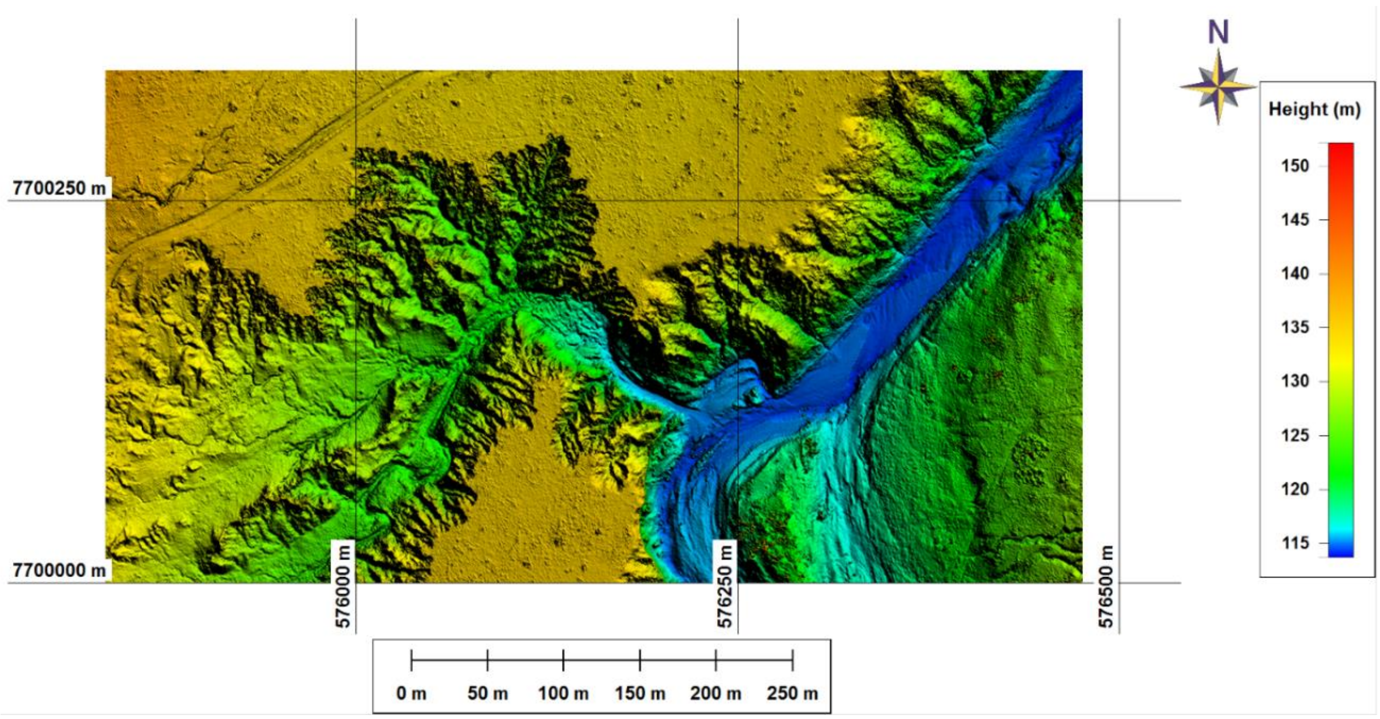


Figure 6: DTM - Australian study area

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

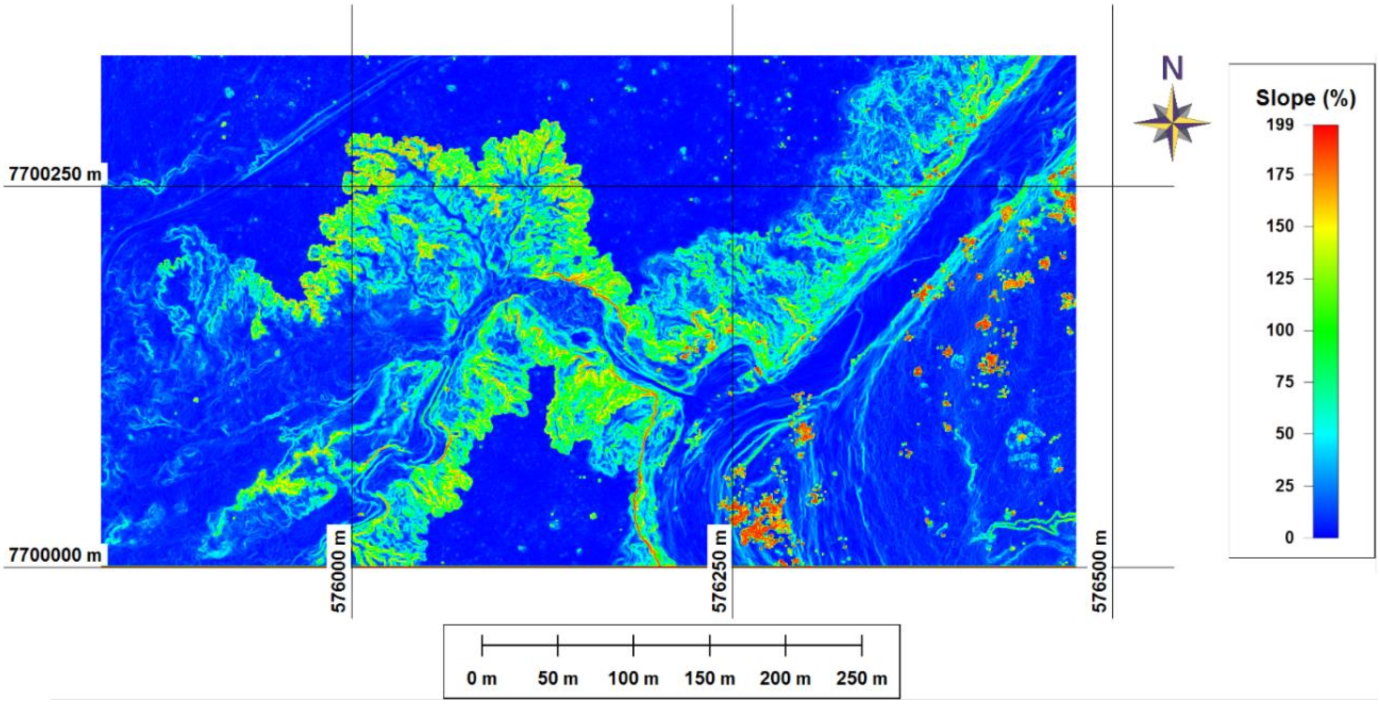


Figure 7: slope map - Australian study area

The NDVI was generated by Ikonos Image for Brazilian area, as shown in figure 8. The vegetated and not vegetated areas can be discriminated, as well tree vegetation and ground vegetation areas. For the Australian study area, as the infrared image is not available, it was proposed a new index based on intensity image. The Intensity-Based Contrast Index – IBCI, can be obtained from:

$$IBCI = (Red - Intensity) / (Red + Intensity) \qquad (1)$$

Unlike NDVI, the IBCI highlights the ground but not the vegetation. It is clear that spectral and altimetry data should be acquired at the same time and should be georreferenced. The figure 9 shows IBCI image for Australian data. The areas with soil are evidenced. The Figure 10 shows the IBCI image for the Brazilian study area. Comparing to the figure 8, it is evident the featured in the soil areas.

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

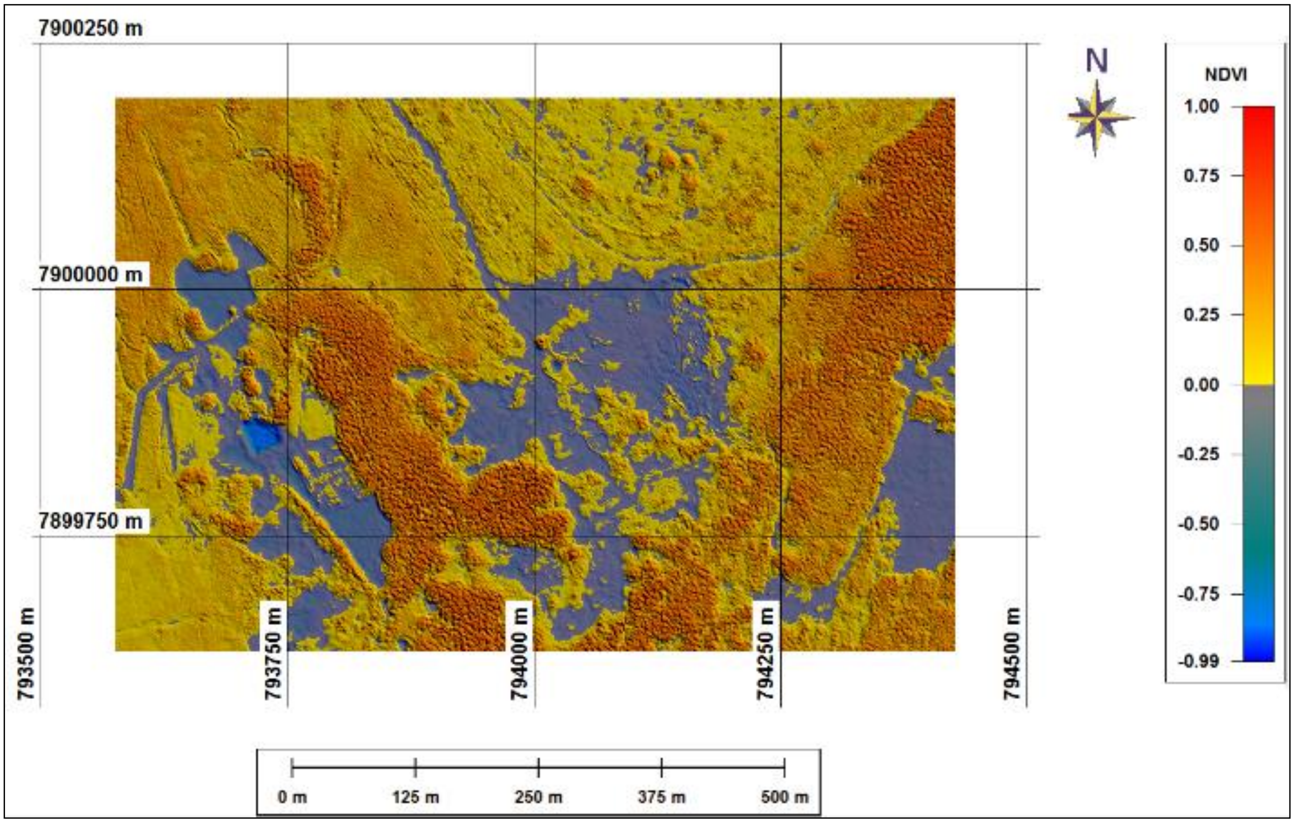


Figure 8: NDVI image - Brazilian study area

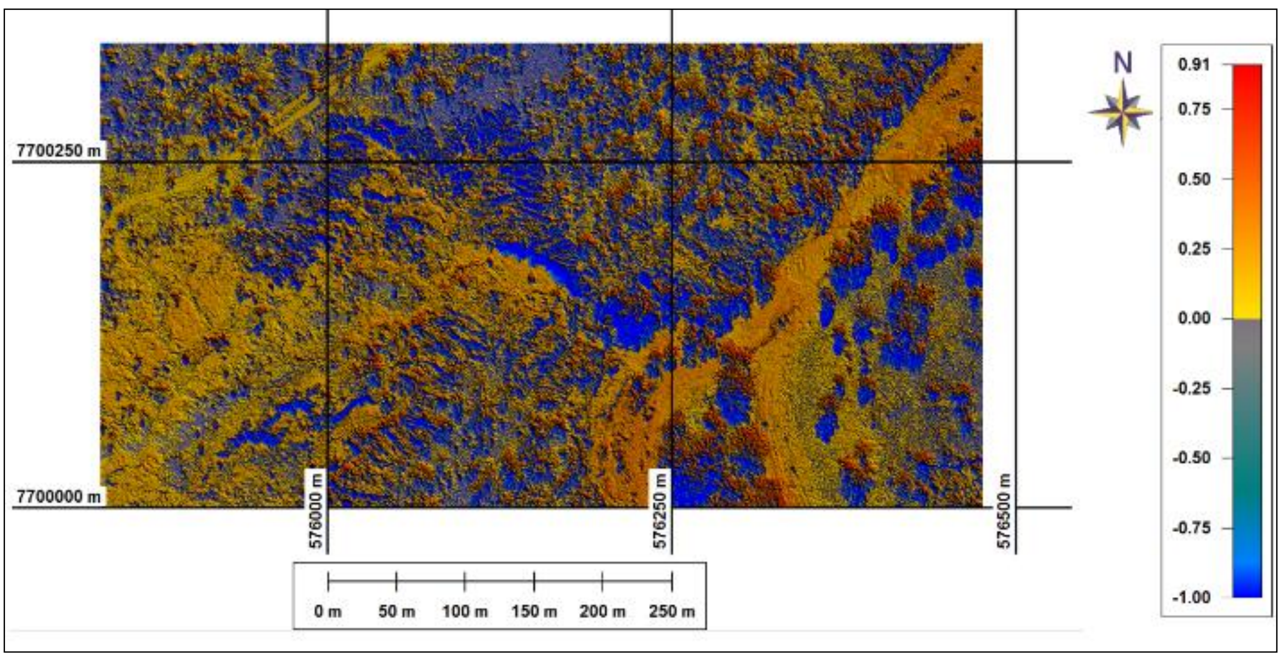


Figure 9: IBCI Image - Australian study area

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

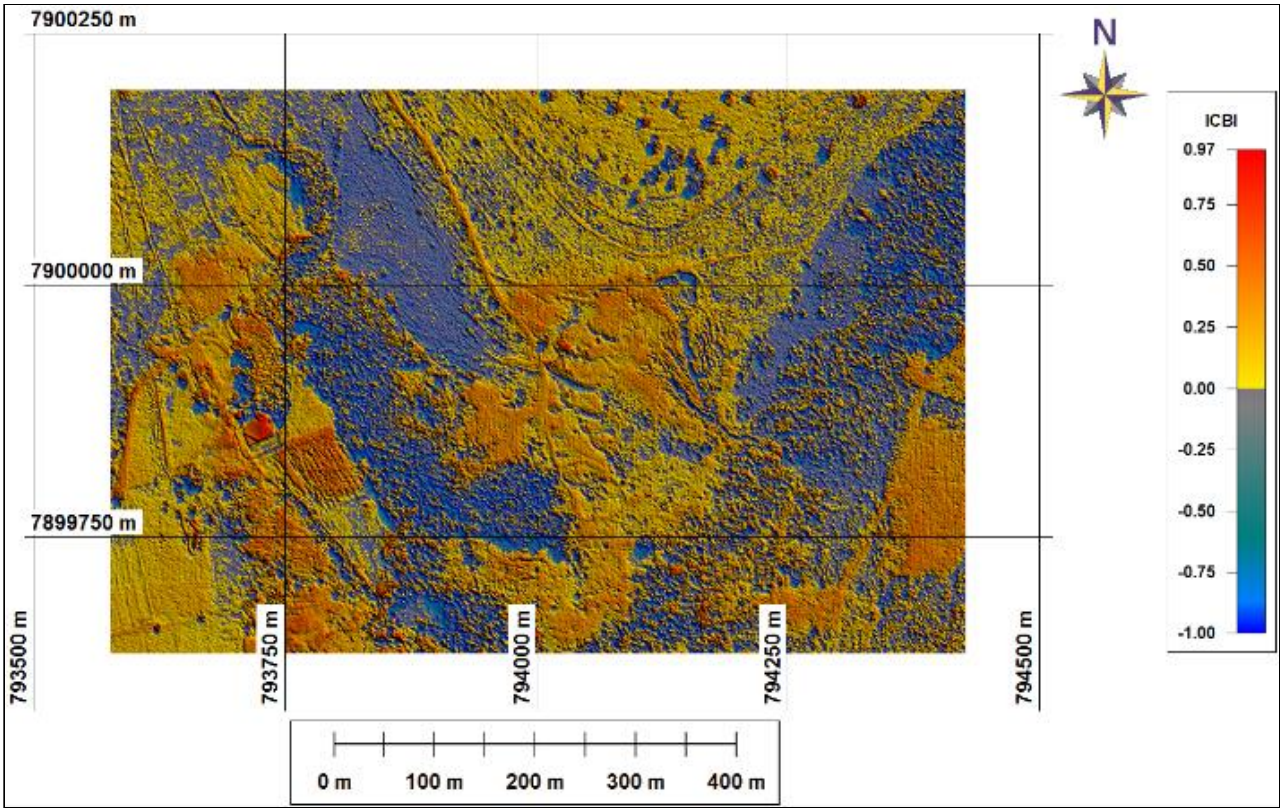


Figure 10: IBCI Image - Brazilian study area

OBJECTS GENERATION

The objects were obtained by using multiresolution segmentation, based on FNEA - Fractal Net Evolution Approach (BAATZ and SCHÄPE, 1999), applied to spectral data (because there are not a perfect coincidence between edges in the image and in the ALS data), ranging the scale factor between 5 and 100, with multiple range of 5. The composition of homogeneity criterion was shape = 0.1 and compactness = 0.5.

CLASSIFICATION PLANNING

The classification planning involved four steps: choosing land cover classes, identifying interpretation keys, defining hierarchical net and selecting attributes.

For the Brazilian study area there were identified these land use classes: tree vegetation, ground vegetation, shadow, water, bare soil and gully erosion. For the Australian study area there were identified these land use classes: vegetation, shadow, bare soil and gully erosion. The figure 11 shows the interpretation keys for Brazilian study area and figure 12 for Australian study area.

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

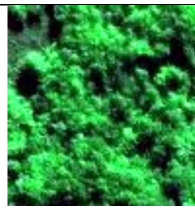
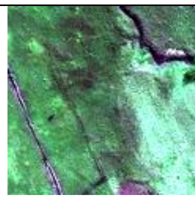
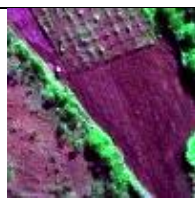
Class	Sample	Collor	Shape, size, texture, height variation
Tree Vegetation		Green (ranging from dark to medium tones)	Irregular shapes and varied sizes due to the presence of both isolated and clustered trees; rough texture; variation of height.
Ground Vegetation		Green (ranging from dark to light tones)	Regular shapes almost rectangular; medium to large sizes; smooth texture; without variation of height.
Dry Bare Soil		White to pink	Irregular forms; sizes varying from medium to large; smooth texture; without variation of height.
Wet Bare Soil		Purple	
Water body		Blue	Regular shapes; medium size; smooth texture; without variation of height
Gully erosion		Mix of soil / vegetation and shadow (pink, purple, green, black)	Irregular forms; sizes ranging from small to medium, due to the presence of different features; rough texture; variation of height.

Figure 11: Interpretation keys for Brazilian study area

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

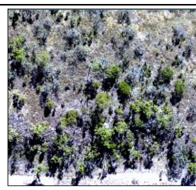
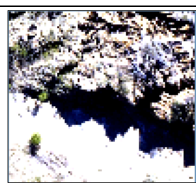
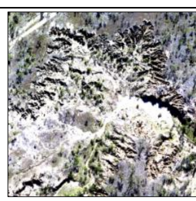
Class	Sample	Collor	**Shape, size, texture, height variation**
Vegetation	  	Green (ranging from dark to medium tones)	Irregular shapes and varied sizes due to the presence of both isolated and clustered trees; rough texture; variation of heights.
Bare soil		Purple, pink, white, light brown and dark	Irregular shapes and varied sizes. Smooth texture.
Shadow		Black	Irregular shapes and varied sizes, since they are in the surroundings of agglomerated and isolated trees, and of the gully
Gully		Mix of soil / vegetation and shadow (white, green and black)	Irregular shapes; sizes varying from small to medium, due to the presence of different features, at different stages of evolution; rough texture; variation of height, high slope at the edges.

Figure 12: Interpretation keys for Australian study area

For the Brazilian area the definition occurred as a function of field recognition at the site and data analysis. For Australian area, due to the impossibility of on-site recognition, the semantic network was defined

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

only based on the observation of available data. The keys of interpretation were then analyzed, regarding the characteristics of color, shape, size, texture and height variation of the representative elements of each class.

The available attributes in eCognition software were evaluated for objects generated with a scale factor (SF) = 50 and by visual analysis to allow comparison with the automatic selection using the CART - Classification And Regression Trees algorithm (BREIMAN et al., 1984). The spectral, geometric, texture and context attributes were evaluated. As the study areas present distinct characteristics, both in terms of the classes in the scene and in relation to the spectral response, to the ground cover and to the available bands for processing, the attributes selected for each area are also different. For the Brazilian area were selected attributes related to the spectral / altimetric response of each band, geometric, texture and pixel based. For Australian area only the spectral and texture attributes were relevant for class discrimination.

III. RESULTS

DATA MINING AND CLASSIFICATION BY TREE DECISION

A decision tree is a classifier expressed from a partition or recursive division of the sample space (MAIMON and ROKACH, 2010). The decision tree induction algorithms correspond to a supervised learning method, since they allow the extraction of knowledge from input samples in the form of training samples (TSO and MATHER, 2009).

For this study it was necessary to select a level of segmentation in which the procedure was performed. On the next step, training samples were collected at the selected level. For the two study areas the selected segmentation level was SF = 50. The selection was based on the analysis of the most representative objects in the region of the gully erosion. For the training, in order to avoid that it becomes over-adapted to a certain class, it is necessary that samples are given in similar quantities for each class. The attributes to be considered were indicated by the authors and evaluated by using CART algorithm, implemented in eCognition. From the samples, the decision tree training took place. On the next step, the decision tree was generated based on the most relevant attributes in order to generate the smallest possible tree. The decision tree was then applied to the image for automatic classification. The figures 13 and 14 shows the tree decisions obtained for the study areas.

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

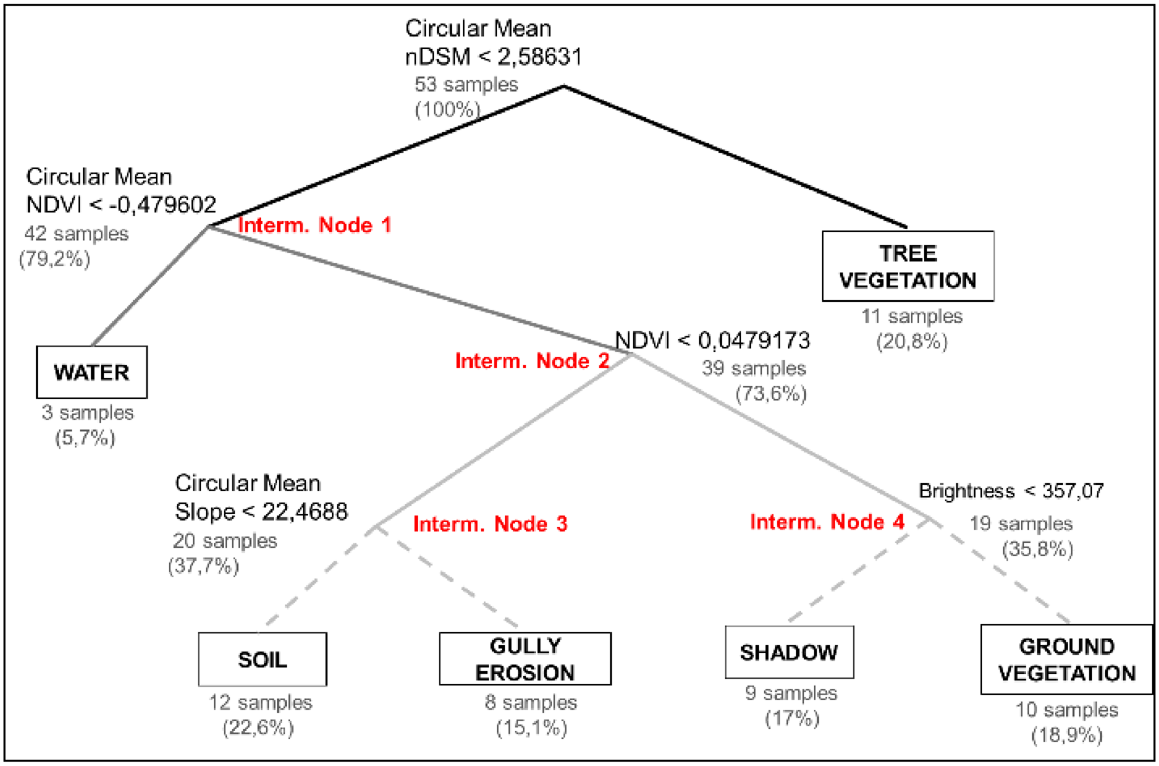


Figure 13: Decision tree - Brazilian study area

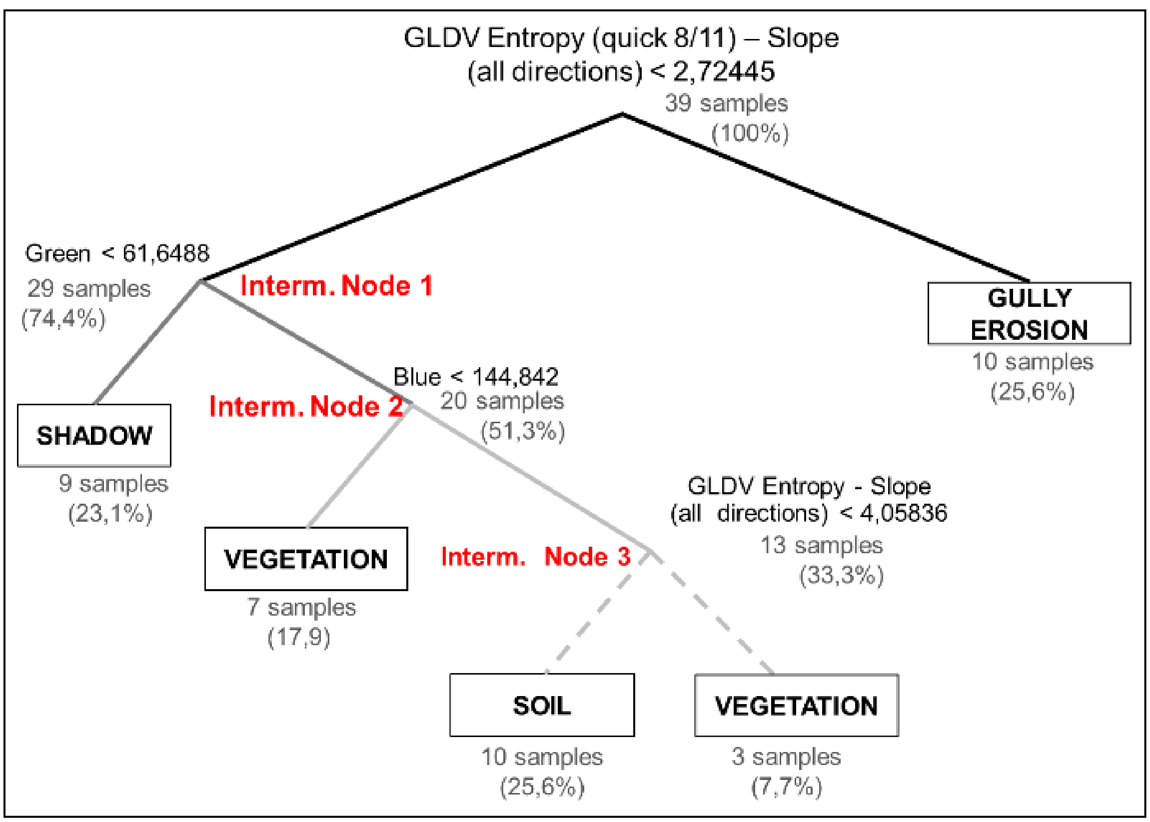


Figure 14: Decision tree - Australian study area

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

For the Brazilian study area, in a first node, using circular mean applied to NDSM, the samples were divided on an intermediate node 1 and on a leaf of the Tree Vegetation class. In the node 1, using circular mean applied to NDVI, the samples were divided on a leaf of the Water class and in the intermediate node 2. In the node 2, using NDVI, the samples were divided on intermediate nodes 3 and 4. In node 3, using circular mean applied to slope map, the samples were classified on Soil class or in Gully erosion class. In node 4, using Brightness, the samples were classified on Shadow or Ground Vegetation classes. For this data set, the major attribute for gully classification was the slope.

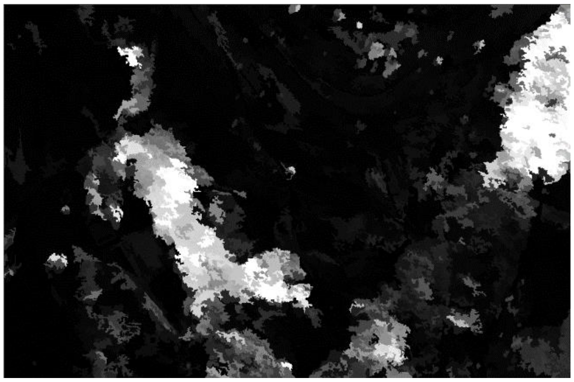
For the Australian study area, in a first node, using Texture after Haralick - GLDV entropy (HARALICK, SHANMUGAM and DINSTEIN, 1973), applied to slope map, the samples were divided on an intermediate node 1 and on a leaf of the Gully Erosion class. In the node 1, using Green band, the samples were divided on a leaf of the Shadow class and in the intermediate node 2. In the node 2, using Blue band, the samples were divided on a leaf of the Vegetation class and in the intermediate node 3. In the node 3, using Texture after Haralick (GLDV entropy) applied to slope map, the samples were classified in the Soil class or in the Vegetation class.

The figure 15 shows, as example, the images related to the attributes selected by the decision tree for Brazilian area. It is possible to verify the effectiveness of the discrimination between the classes suggested by the tree in each tree decision node, according to the selected attribute.

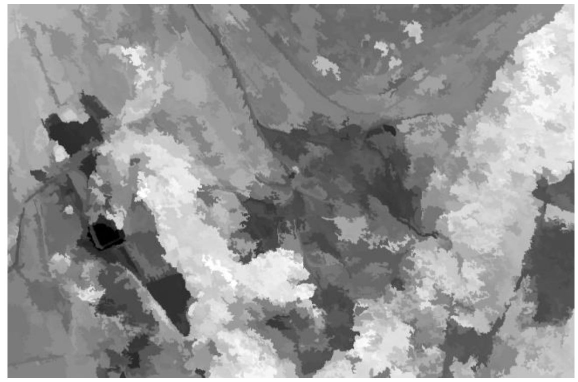
The figures 16 and 17 show the classifications performed by decision tree. In figure 16, for Brazilian area, it is possible verify that occurred confusion between gully erosion and vegetation classes, in the edges of the tree vegetation due to the big slope in these areas (height difference between ground level and tree tops). Even with the similarity between soil and gully erosion classes there was few confusion areas due to the use of the slope attribute to separate these classes. The gully system neighborhood was incorporated to the gully class because it was used the circular mean of the slope map attribute which expanded gully erosion area to the soil area. There was confusion between tree and ground vegetation. The shadow areas were evidenced.

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega

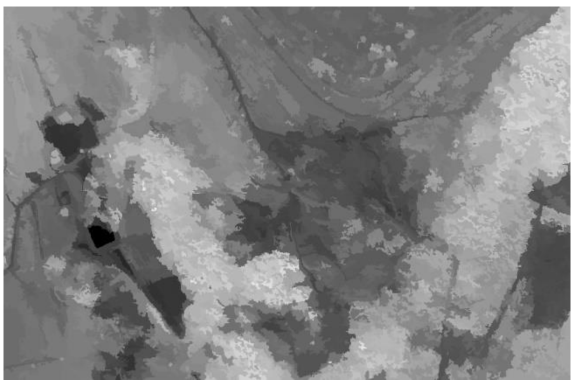
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842



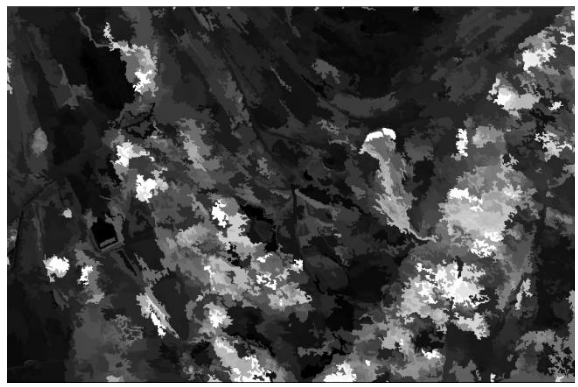
Circular Mean – NDSM
Tree Vegetation X Node 1



Circular Mean– NDVI
Water X Node 2



NDVI
Node 2 X Node 3



Circular Mean - slope
Soil X Gully Erosion

Brightness
Shadow X Ground Vegetation

Figure 15: Attributes selected by data mining - Brazilian study area

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

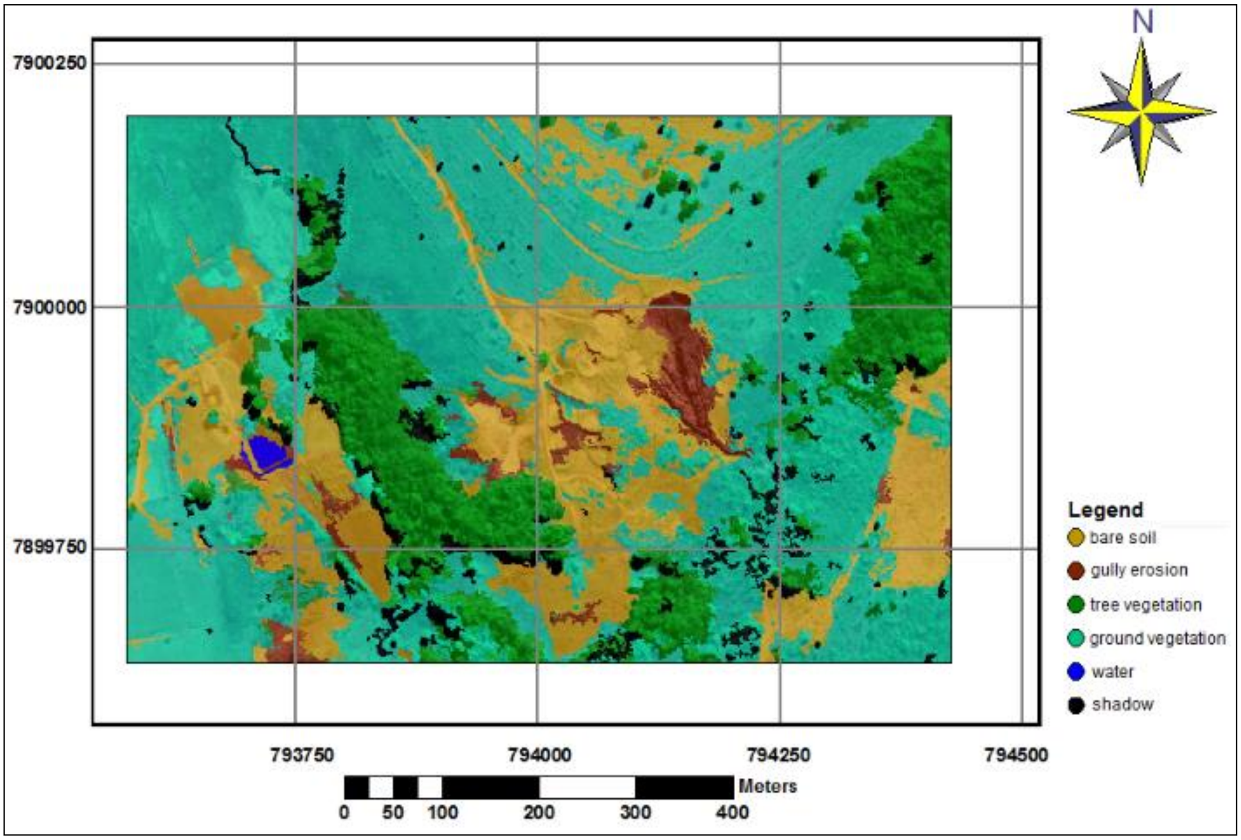


Figure 16: Classification by decision tree - Brazilian study area

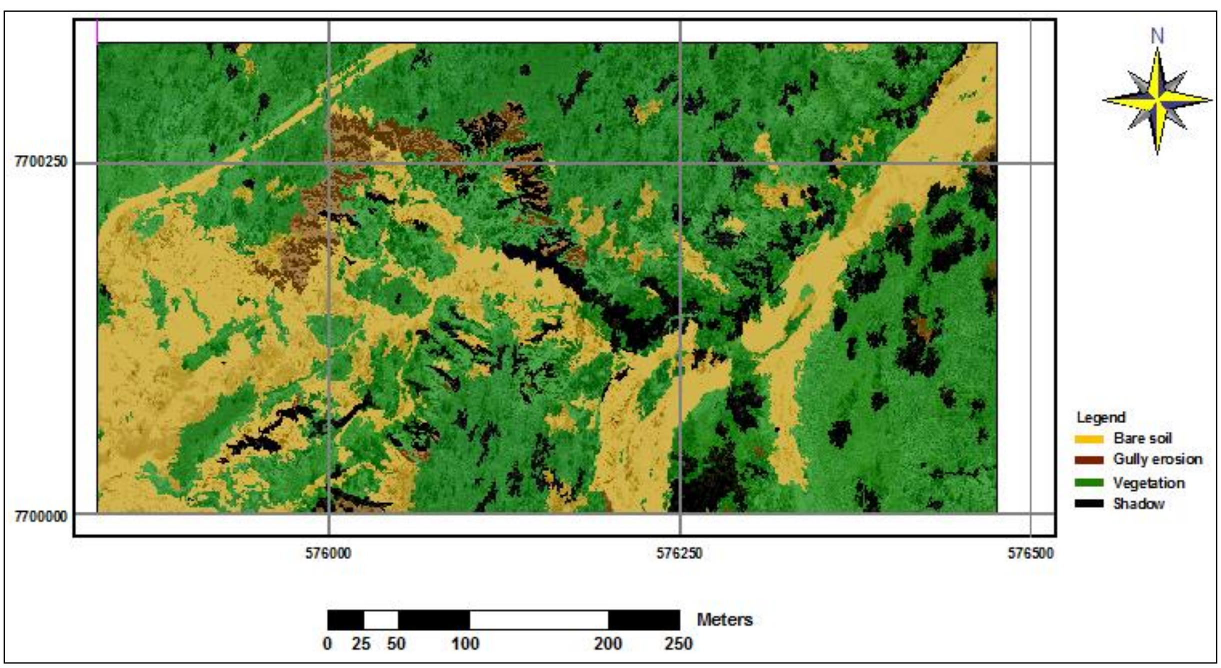


Figure 17: Classification by Decision tree - Australian study area

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

In the figure 17, that shows the Australian study area classification by decision tree, it is possible to verify that there was confusion between gully erosion and vegetation classes, due to the big slope between ground level and tree tops. The shadow areas were evidenced. Even with the high similarity between the bare soil and gully erosion classes, the gully could be discriminated using attributes related to slope, as well as in Brazilian area. This reinforces the effectiveness of the incorporation of altimetric data in the classification of gullies, as verified also by JOHANSEN et al. (2012).

Another automatic classification was performed using SVM - Support Vector Machine (TSO and MATHER, 2009), for the Brazilian area, just to compare with the results achieved with Tree Decision, due to both are automatic classifiers. The attributes selected for the proceeding were based in the selection performed by visual analysis, which were:

-Mean of the layers: Ikonos bands (blue, green, red and near infrared), brightness, intensity image ALS, MDSN, Max. Difference, NDVI, slope map.

-Standard deviation of slope map.

-Skewness of the layers: slope, MDS and Ikonos bands.

-Pixel-based: ratio, minimum pixel value, maximum pixel value, mean of inner border, mean of outer border, border contrast and circular mean

-Geometry: extent → length / width

-Shape: density, rectangular fit

-Texture after Haralick

The figure 18 shows the result. It is possible to map gully, but there is a lot of confusion between gully erosion and bare soil classes. Besides, many areas were classified as shadow.

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

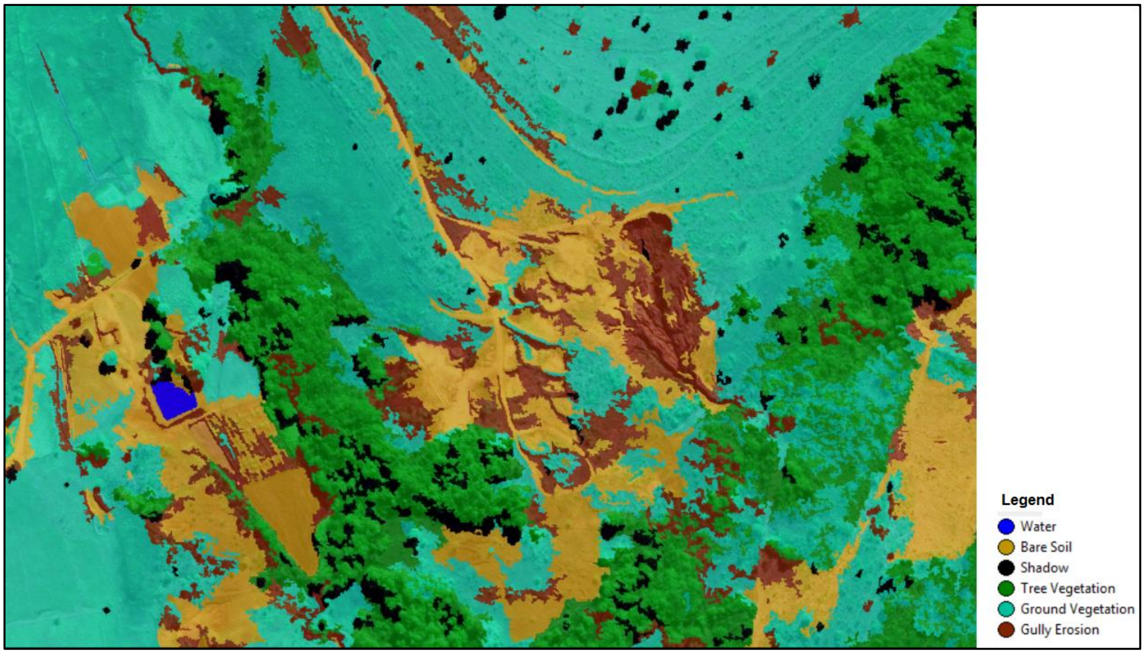


Figure 18: Classification by SVM - Brazilian study area

HIERARCHICAL CLASSIFICATION

The hierarchical classification, integral to the FNEA, is based on fuzzy logic (ZADEH, 1965 and 2008) and it is performed based on the different levels from the multiresolution segmentation, being, therefore, an object oriented method. Features are classified according to the level at which they are best segmented. The classification is performed based on the established hierarchical network, the samples selected for training the algorithms and the attributes selected as descriptors for each class (BAATZ and SCHÄPE, 1999). The interactive edition of fuzzy membership functions allows the introduction of knowledge.

From the decision trees obtained, the most relevant attributes in the class discrimination were selected. The decision tree also resulted in the explanation of the rules for classification, suggesting a hierarchy, through the division in nodes of the trees. Based on the hierarchy suggested by the tree and the expert knowledge of the authors, a hierarchical network was proposed, selecting the levels of segmentation and the classes to be discriminated in each level. In addition to the attributes indicated by decision trees, other attributes and fuzzy memberships functions were incorporated, according to the expert knowledge, in the discrimination of each class. The hierarchical classification was then performed for the two areas.

The figures 19 and 20 shows the hierarchical rule bases, with fuzzy memberships functions, for Brazilian and Australian study areas respectively.

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

* note: NN = nearest neighbor algorithm; SF = scale factor

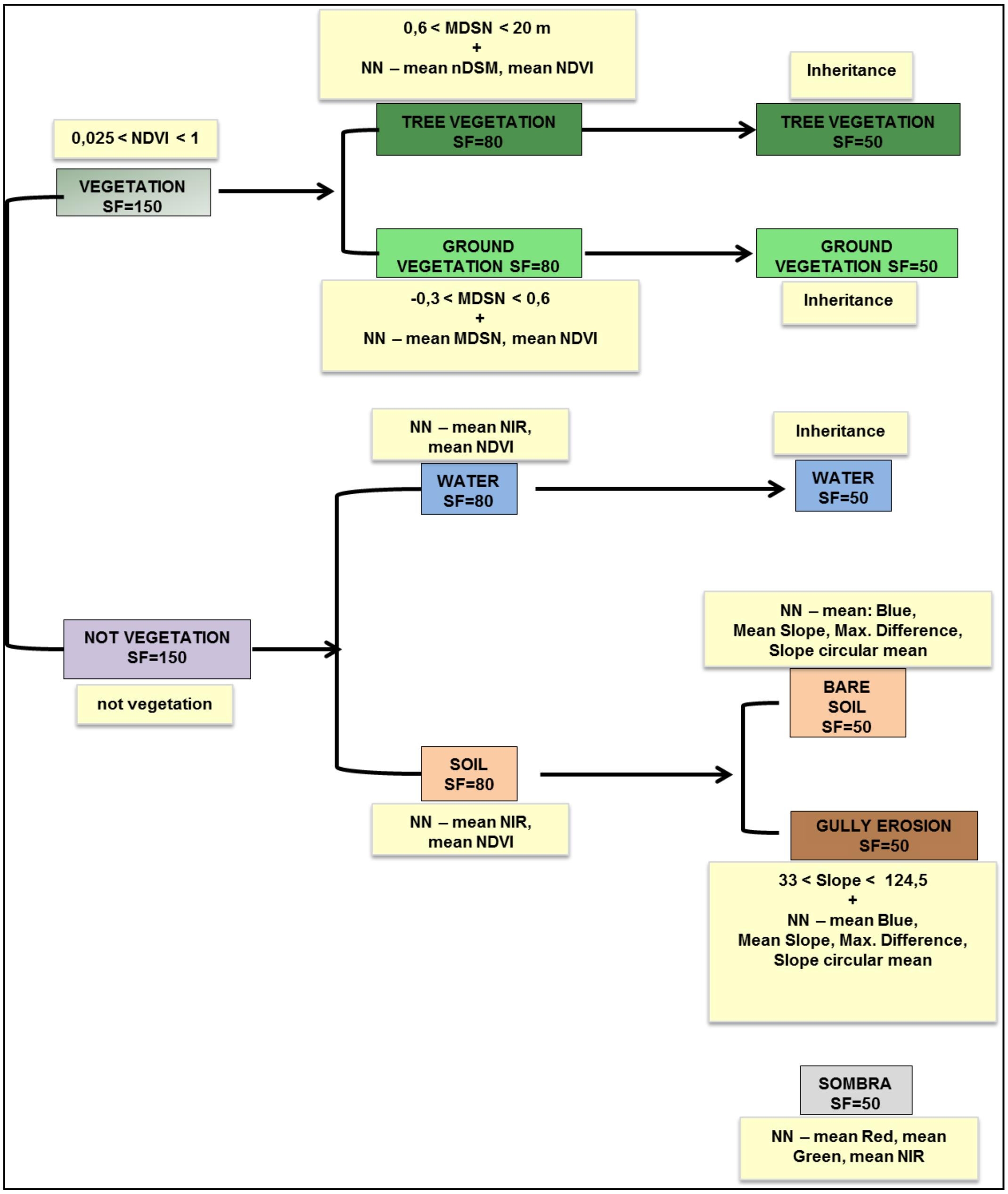


Figure 19: Hierarchical rule base - Brazilian study area

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

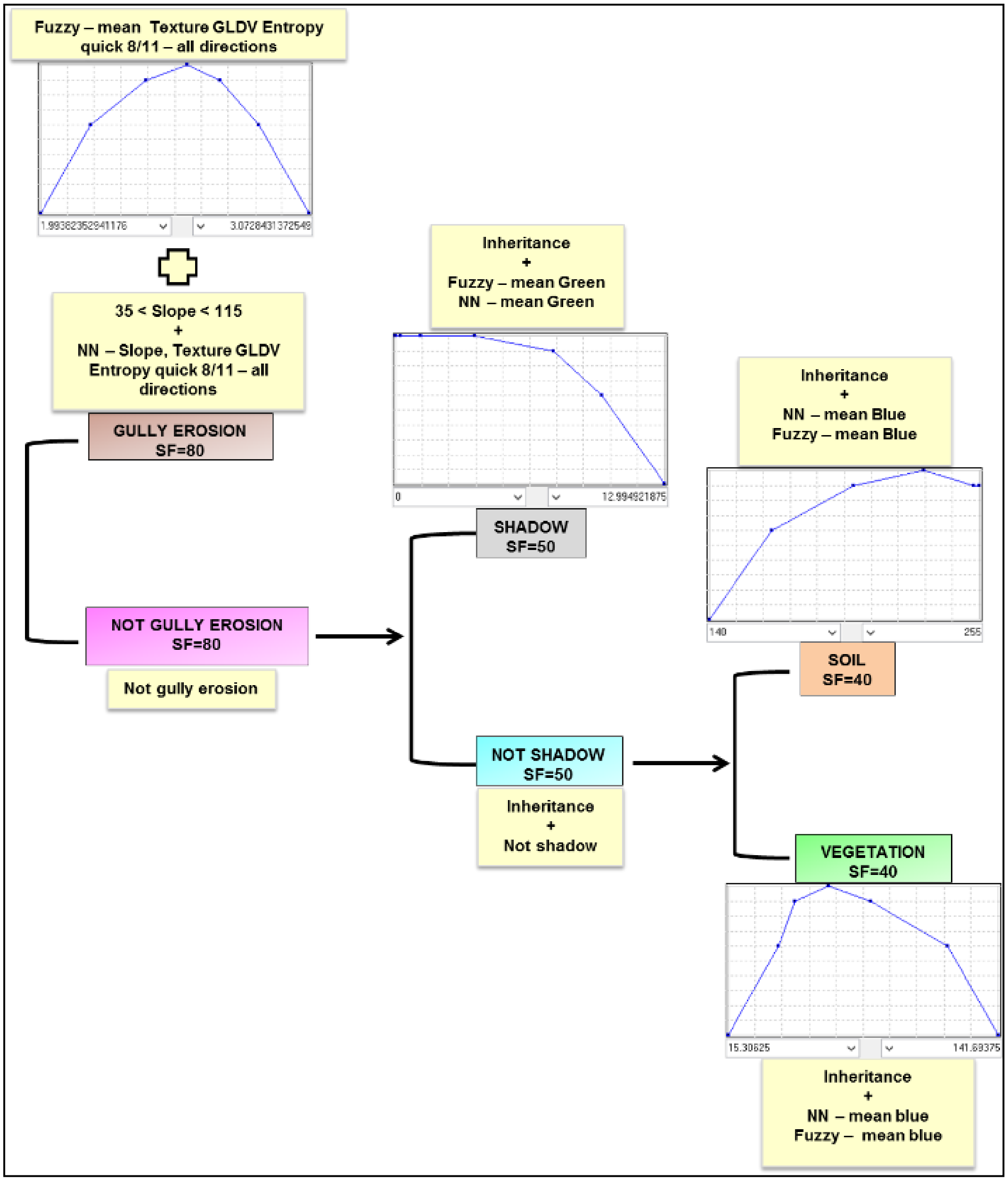


Figure 20: Hierarchical rule base - Australian study area

The figure 21 shows the Brazilian study area hierarchical classification. The gully system edge was better mapped than by tree decision classification. The inside areas of the gully system were classified in Soil class due to the use of slope attribute (inside the gully, the slope is small). The water body and vegetated areas were better mapped in this classification.

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

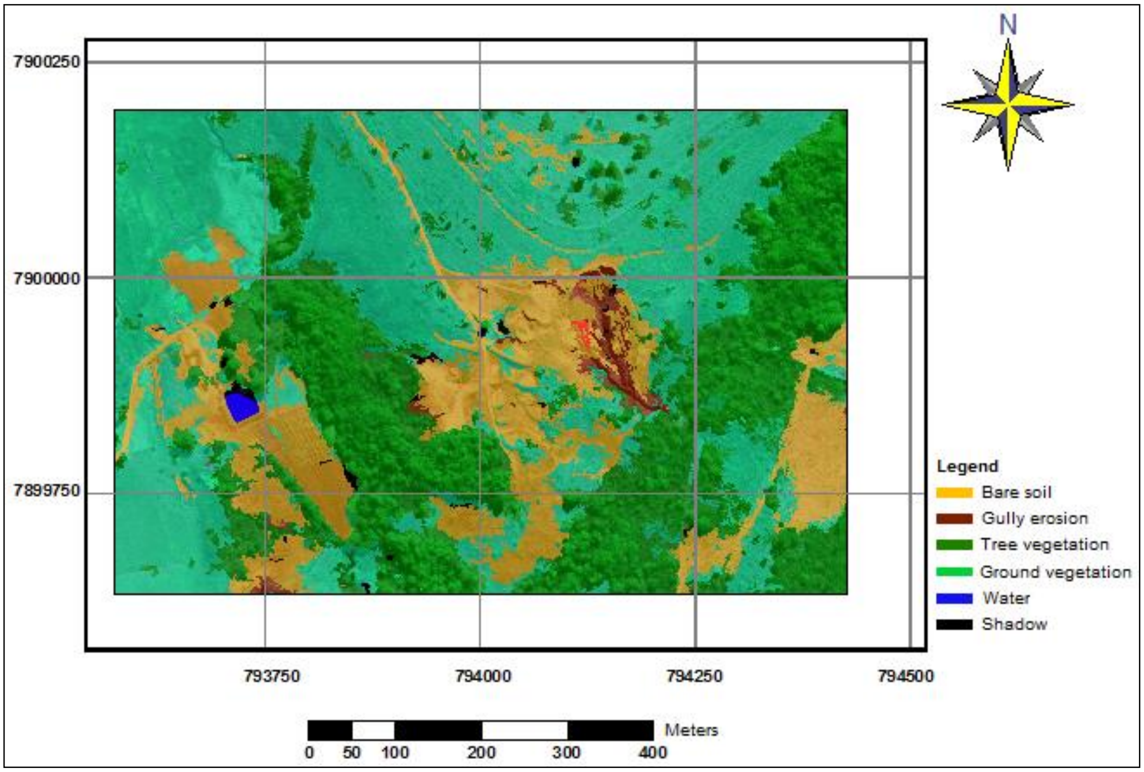


Figure 21: Hierarchical classification – Brazilian study area

The figure 22 shows the Australian study area hierarchical classification. The gully system was better classified than on the classification by tree decision. The main attributes selected for gully erosion classification were slope and texture. As vegetated areas have big slope and texture too, there was confusion between gully an vegetation classes.

CLASSIFICATION EVALUATION

Brazilian Study Area

In the classification by decision tree, the areas of shadow were more evidenced, however, large areas of tree vegetation were erroneously classified as ground vegetation. This can be explained because that in the rule base the MDSN was not considered to separate these two classes and also they were discriminated in different branches of the tree. Some soil areas were classified as gully, even that the rule base have used the circular average of the slope image as discriminant attribute.

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

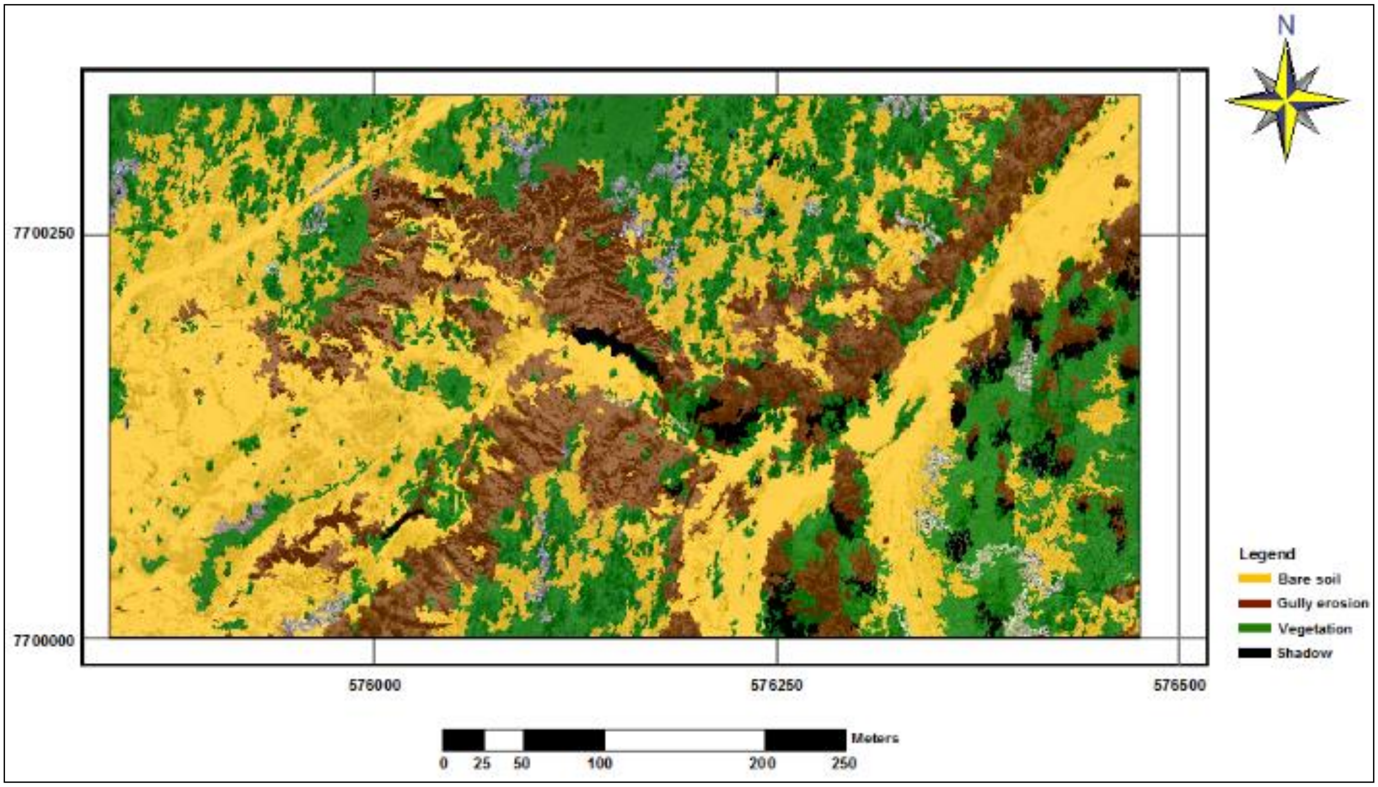


Figure 22: Hierarchical classification – Australian study area

The surrounding area of the gully erosion, which is bare soil, was mistakenly incorporated into the gully erosion class. In this separation of classes the tree decision used the circular mean attribute of slope. This attribute considers the behavior of neighboring objects within a ring, the inner and outer beams of that ring was defined by the user. Equal internal and external rays were defined, forming a circle of radius equal to 3. It can be seen in figure 15 (image of the mean circular attribute of the slope) that, when considering within the circle an object relative to one of the digits of the gully, the neighbors in a radius equal to 3 were incorporated with the same value of the central object. This way, the digits were connected to the main channel of the gully, influencing the classification.

In the hierarchical classification, the design of roads, planting plots and the body of water in the scene was obtained, as well as the identification of isolated trees. Few elements of bare soil were incorporated into the gully erosion class. Some regions around tree vegetation, which have high slope, were erroneously classified as gully. The interior of the head of the gully was classified as bare soil, since the discriminant attribute used in

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

this step of the processing was the slope and, inside the gully head, due to the intense transport of sediments, this region is planned.

The results show that the decision tree is useful in the selection of attributes and for a preliminary mapping. However, the possibility to explicit knowledge, the use of fuzzy rules and the classification made at different levels of segmentation, make hierarchical classification a more effective method in thematic mapping of soil coverages.

The table 1 shows the confusion matrix for this area, performed by selecting verifying samples. In the Brazilian study area image there are 2394 image objects and 217 samples were selected, with 95% confidence interval (error = 6.35%), yielding kappa index = 0.75 and overall accuracy = 82%. 3 of 14 gully samples were classified as bare soil and 1 as ground vegetation, 3 of 42 bare soil samples were classified as gully.

TABLE 1: CONFUSION MATRIX – BRAZILIAN STUDY AREA

Classif. / Ref. Class	ground veg.	tree veg.	bare soil	gully erosion.	water	shadow	Σ	Productor Accuracy
ground veg.	57	5	5	-	-	-	67	85.07
tree veg.	5	59	-	-	-	-	64	92.19
bare soil	13	2	42	3	-	1	61	68.85
gully erosion	1	-	3	14	-	-	18	77.78
water	-	-	-	-	3	-	3	100
shadow	1	1	-	-	-	2	4	50
Σ	77	67	50	17	3	3	217	
User Accuracy	74.03	88.06	84	82.35	100	66.67		

Australian Study Area

In the classification by decision tree, the areas of shadow were more evidenced, there were no unclassified objects and there was no confusion between the classes gully erosion and vegetation (mainly tree vegetation). However, few objects were classified as gully and regions of bare soil with purple tones were classified as vegetation (what would correspond to ground vegetation, in this case, since the semantic network showed only the vegetation class, not separating in tree and ground).

In the hierarchical classification, the gully objects were better identified, even the smallest and most isolated. However, many trees were incorporated into the gully erosion class. This is due to the fact that slope and texture had been used as determining attributes of this class. The trees present values, for these attributes, very similar to the gully objects, so that the proposed index IBCI was not sufficient for the separation of these

ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

classes. Evaluating only the interest class, it can be verified that the attributes and fuzzy rules selected were effective in the gully classification.

The heterogeneity of the scene, the existence of many isolated small objects (trees) and the existence of trees without foliage (identified only in function of the MDSN), made it difficult to separate the other classes.

The table 2 shows the confusion matrix for this area, in which there are 2265 image objects and 865 samples were selected, with 99% confidence interval (error = 5%), yielding kappa index = 0.46 and overall accuracy = 64.05%. 3 of 102 gully samples were classified as vegetation, 29 as bare soil and 1 as shadow. 61 of 127 vegetation samples were classified as gully, 27 of 309 of bare soil samples were classified as gully and 11 of 16 shadow samples were classified as gully.

TABLE 2: CONFUSION MATRIX – AUSTRALIAN STUDY AREA

Ref. Class \ Classif.	vegetation	bare soil	gully erosion	shadow	∑	Productor Accuracy
vegetation	127	47	61	-	235	54.04
bare soil	88	309	27	-	424	72.88
gully erosion	3	29	102	1	135	75.56
shadow	43	1	11	16	71	22.53
∑	261	386	201	17	865	
User Accuracy	48.66	80.05	50.75	94.12		

IV. CONCLUSIONS

As the gully systems are composed of features with highly variable shapes and sizes, they could be mapped from high resolution imagery, with auxiliary altimetry data and by GEOBIA. Using hierarchical classification, it is possible to select different scale factors appropriated to gully features with variable sizes. For both data sets the scale factor equals to 50 was enough to map gullies. A 1 meter resolution was enough to identify the gully features.

As the gully systems presents higher similarities when they correspond to the same stage of evolution and soil type, for example, there is no way to select attributes that are appropriate to the classification of all systems, requiring the investigation of discriminant attributes for each gully system, including on the basis of available data and existing land use classes in the scene.

Data mining by decision tree allowed rapid analysis and the best attributes selection, subsidizing the decision making process, replacing empiricism, providing a preliminary decision rule base, which can be adjusted

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

according to the expert knowledge to realization of hierarchical classification. A disadvantage of the classification by decision trees is that it can only do at one level of segmentation. As the same, data mining by SVM was performed only for one level of segmentation and the results were similar with the tree decision ones.

For data sets available and the specificities of the two study areas, the attributes that were more relevant to the discrimination of the gully class were the slope and texture. Regarding the data set, it is emphasized that the use of spectral data, high spatial resolution, coupled with the use of altimetry data allows the classification of gullies.

In general, as the gully is an object and not a land use class (resulting from the mixing of classes such as soil, vegetation, shade and water) her delimitation can not be obtained by pixel oriented classification procedures, only in object-oriented procedures. Due to the spectral mixture with the other classes, the incorporation of non-spectral data is necessary for classification, as is the case of altimetric data and texture attributes.

In the hierarchical classification, the procedure is performed in different scales, allowing the use of fuzzy logic to describe different degrees of relevance to each class, which makes the method quite attractive for cases like this one, where there is a mixture of classes. Thus, the classification obtained with the semiautomatic method of hierarchical classification proved to be more reliable to the field reality, since it allows the use of different scales, insertion of uncertainty (by fuzzy logic) and insertion of knowledge (by the established rule base), when compared to the automatic classification by decision tree or by SVM.

Note also that the index proposed IBCI allowed the enhancement of the soil, being an alternative in case of unavailability of the infrared band, but the availability of ALS intensity band.

ACKNOWLEDGEMENTS

The authors wish to thank the Remote Sensing Centre, Queensland Department of Environment and Resource Management (Mr. Dan Tindall) for giving ALS data and orthophoto of the Queensland study area. Also thank the Engesat Imagens de Satélite Ltda (Mrs. Jaqueline Cavalli) for giving Ikonos imagery and Esteio Engenharia e Aerolevantamentos S.A. (Mrs. Elizabete B. A. Peixoto) for giving ALS data of the Uberlandia study area.

RA'EGA
O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

V. REFERÊNCIAS

ALKHASAWNEH, M.; NGAH, U. K.; TAY, L. T.; ISA, N. A. M.; AL-BATAH, M. S. Modeling and Testing Landslide Hazard Using Decision Tree. **Journal of Applied Mathematics**, p. 1-9, 2014.

ARABAMERI, A.; REZAEI, K., POURGHASEMI, H. R.; LEE, S.; YAMANI, M. GIS-based gully erosion susceptibility mapping: a Comparison among three data-driven models and AHP knowledge-based technique. **Environmental Earth Sciences**, volume 77**,** 2018.

BAATZ, M.; SCHÄPE, A. Object-Oriented and Multi-Scale Image Analysis in Semantic Networks. **2nd International Symposium: Operationalization of Remote Sensing**, ITC Netherlands, 1999.

BREIMAN, L.; FRIEDMAN, J. H.; OLSEN, R. A.; STONE, C. J. **Classification and Regression Trees**. Belmont, CA: Wadsworth, 1984.

BUI, D. T.; SHIRZADI, A.; SHAHABI, H.; CHAPI, K.; OMIDAVR, E.; PHAM, B. T.; ASL, D. T.; KHALEDIAN, H.; PRADHAN, B.; PANAHI, M.; AHMAD, B. B.; RAHMANI, H.; GRÓF, G.; LEE, S. A Novel Ensemble Artificial Intelligence Approach for Gully Erosion Mapping in a Semi-Arid Watershed (Iran). **Sensors** (Basel), v.19; 2019.

EUSTACE, A.; PRINGLE, M.; WITTE, C. Give me the dirt: Detection of Gully Extent and Volume Using High-Resolution LIDAR. **Innovations in Remote Sensing and Photogrammetry: Lecture Notes in Geoinformation and Cartography.** Berlin: Springer publisher, pp. 255-269, 2009.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEIN, I. Textural Features for Image Classification. **IEEE Transactions on systems, man, and cybernetics**, v. SMC-3, nº. 6, p. 610-621, 1973.

HESSEL, R.; VAN ASCH, T. Modelling gully erosion for a small catchment on the Chinese Loess Plateau. **Catena**, 54, pp. 131-146, 2003.

JAMES, L. A.; WATSON, D. G.; HANSE, W. F. Using LiDAR data to map gullies and headwater streams under forest canopy: South Carolina, USA. **Catena**, 71, pp. 132-144, 2007.

JOHANSEN, K.; TAIHEI, S.; TINDALL, D.; PHINN, S. Object-based Monitoring of Gully Extent and Volume in North Australia using LiDAR Data. **Proceedings of the 4th GEOBIA**, Rio de Janeiro, pp. 168-173, 2012.

KING, C.; BAGHDADI, N.; LECOMTE, V.; CERDAN, O. The application of remote-sensing data to monitoring and modelling of soil erosion. **Catena**, 62, pp. 79-93, 2005.

MARTÍNEZ-CASANOVAS, J. A.; RAMOS, M. C.; POESEN, J. Assessment of sidewall erosion in large gullies using multi-temporal DEMs and logistic regression analysis. **Geomorphology**, 58, pp. 305-321, 2004.

MARZOLFF, I.; POESEN, J. The potential of 3D gully monitoring with GIS using high-resolution aerial photography and a digital photogrammetry system. **Geomorphology**, volume 111, pp. 48-60, 2009.

MORGAN, R. P. C. **Soil Erosion and Conservation.** Oxford: Blackwell Publishing, 2005.

PERROY, R. L.; BOOKHAGEN, B.; ASNER, G. P.; CHADWICK, O. A. Comparison of gully erosion estimates using airborne and ground-based LiDAR on Santa Cruz Island, California. **Geomorphology**, 118, pp. 288-300, 2010.

REAL, L. S. C.; CRESTANA, S.; FERREIRA, R. R. M.; RODRIGUES, V. G. S. Evaluation of gully development over several years using GIS and fractal analysis: a case study of the Palmital watershed, Minas Gerais (Brazil). **Environmental Monitoring and Assessment,** volume 192, Article number: 434, 2020.

RA'EGA O ESPAÇO GEOGRÁFICO EM ANÁLISE
https://revistas.ufpr.br/raega
ISSN eletrônico 2177-2738
RA'EGA, Curitiba, PR, V.8, n.2, p. XX – XX, 11/2020
DOI: http://dx.doi.org/10.5380/raega.v48i0.74842

SHRUTHI, R. B. V.; KERLE, N.; JETTEN, V. Object based gully feature extraction using high spatial resolution imagery. **Geomorphology,** 134, pp. 260-268, 2011.

SHRUTHI, R. B. V.; KERLE, N.; JETTEN, V. ABDELLAH, L.; MACHMACH, I. Quantifying temporal changes in gully erosion areas with object oriented analysis. **Catena**, v. 128, p. 262-277, 2015.

TSO, B.; MATHER, P. M. **Classification methods for remotely sensed data**. Florida – USA: CRC Press, 2009.

VRIELING, A.; RODRIGUES, S. C.; STERK, G. Evaluating erosion from space: a case study near Uberlândia. **Revista Sociedade & Natureza**, p. 683-696, 2005.

ZABIHI, M.; MOTEVALLI, ALIREZA; DARVISHAN, A. K.; POURGHASEMI, H. R.; ZAKERI, M. A.; SADIGHI, F. Spatial modelling of gully erosion in Mazandaran Province, northern Iran. **Catena**, 161, p.1-13, 2018.

ZADEH, L. A. Fuzzy Sets. **Information and Control**, v. 8, p.338-353, 1965.

ZADEH, L. A. Is there a need for fuzzy logic? **Information Sciences**, v. 178, p. 2751–2779, 2008.

WANG, T.; HE, F.; ZHANG, A.; GU, L.; WEN, Y.; JIANG, W.; SHAO, H. Quantitative Study of Gully Erosion Based on Object-Oriented Analysis Techniques: A Case Study in Beiyanzikou Catchment of Qixia, Shandong, China. **The Scientific World Journal, pp. 1-11, 2014.**

WANG, R.; ZHANG, S.; PU L.; YANG, J.; YANG, C.; CHEN, J.; GUAN, C.; WANG, Q.; CHEN, D.; FU, B.; SANG, X. Gully Erosion Mapping and Monitoring at Multiple Scales Based on Multi-Source Remote Sensing Data of the Sancha River Catchment, Northeast China. **ISPRS - International Journal of Geoinformation**, 5, 200-217, 2016.