On the delineation of riparian forests using elevation and Ikonos data

Philippe Maillard[1]
Thiago de Alencar-Silva[1]

[1] Universidade Federal de Minas Gerais – UFMG/IGC
Av. Antônio Carlos, 6627, Belo Horizonte, MG 31270-901, Brasil
philippe@ufmg.br, thiago-alencar@uol.com.br

Abstract. Riparian formations are often narrow strips of vegetation on margins of streams governed by good water availability. With the advent of high-resolution satellite, the study riparian forests using space-borne remote sensing has gain more attention. However, the sole use of remotely sensed data is seldom sufficient to delineate riparian forests and their special hydrological and geomorphological context should also be considered. In this study, we propose a two-fold method to delineate riparian forests in a brazilian savannah context. Four sites on the margins of the Pandeiros river were used to test our approach, each in a different hydrological context. In our approach, the hydrographic network and the relief are processed using the “Depth To Water” (DTW) analysis to create a mask of the riparian zone, then the riparian forests are extracted using region-based image classification approach from the SPRING software package. Both steps are fully explained. An average scores of about 90% success was obtained. The discussion focuses on methodological choices issues and the advantages of the approach.

Keywords: riparian forest, object-based classification, Ikonos image floresta ribeirinha, classificação baseada em objetos, imagem Ikonos

1. Introduction

Riparian forests have an exceptionally high ecological value compared to other forests. In Brazil they are the subject of an animated debate between those that preach their protection or restoration and others that claim they should be used for agricultural purposes. In all cases, their mapping and state of conservation is a very actual political issue. Until the advent of high spatial resolution orbital sensors, satellite remote sensing had limited application for mapping and monitoring riparian forests. With the availability of high resolution satellite images, an increasing number of remote sensing scientists have had riparian forests as their object of study (JOHANSEN et al., 2007; OLIVEIRa de SOUZA; JOHANSEN, 2008). Many of them have stumble across the problem of delineating these formations and determining criteria for their spatial definition. It is generally agreed that such criteria has to involve some hydrological or geomorphological parameter as well as a biological or ecological one. Riparian forests are connected to streams which they follow with varying adherence and width. Having more continuous access to water, they are usually denser and lusher than the surrounding forest on dryer ground.

Although high resolution satellite data such as *Ikonos*, *Quickbird*, *Worldview*, *Geoeye* or *Pléiade* have a very good richness of details, the accurate delineation of riparian forest requires the third dimension to be considered. A digital elevation model or surface (DEM/DSM) is perhaps the most natural way to involve the third dimension in remote sensing. A DEM can be used to assess the geomorphological-hydrological counterpart of delimiting a riparian forest. However, this involves processing the data and extract some hydrological-geomorphological properties like the direction of flow and the vertical distance to the nearest stream. With such information, it is possible to determine interactively the width of the riparian zone and thence the riparian forests within that zone. Digital image processing can then be used to extract the forested portion within the riparian zone. Given the narrow width of most riparian forests, high resolution images ($< 5m$) are an appropriate source of data. However, pixel-based methods of classification are often ineffective with high resolution data and it is generally accepted that region-based or object-based classifiers yield better results.

In this article, we propose a hybrid method involving processing a DEM and region-based classification of a high-resolution Ikonos image to delineate the riparian forest of a Brazilian stream in a savannah environment. To do so we used a region-based forest classification in conjunction with the "depth to water" algorithm to extract a riparian zone believed to contain the riparian forest. A short review on the delineation of riparian zones and remote sensing of riparian forests follows.

1.1. Transversal Delineation of a Riparian Zone

Riparian zones have distinct ecological characteristics and their limits are normally accompanied by changes in soil condition, vegetation and other factors such as hydrological regime and geomorphological context (NAIMAN; BILBY; BISSON, 2000). Their precise delineation can be determined through topography, soil and flora (POLVI; WOHL; MERITT, 2011). Verry, Dolloff e Manning (2004) have used a geomorphological criterion to define the riparian ecotone as the floodplain of a particular stream plus 30 meters on each side. Johansen e Phinn (2006) used ten field transects to evaluate the performance of their mapping based on a classification of Ikonos and Landsat TM images. Other authors have simplified the problem by establishing a fixed width within which it was believed the riparian zone was contained (CONGALTON et al., 2002). In a previous study, we have used such an approach to limit the search zone where palm swamps are found (MAILLARD; ALENCAR-SILVA; CLAUSI, 2008).

More appropriate approaches used indices derived from a DEM to estimate some local characteristic associated with the riparian zone. Two such indices have received more attention: *Soil Wetness Index* or SWI and *Depth to Water* index or DTW (GÜNTER; SEIBERT; UHLENBROOK, 2004; CASE; MENG; ARP, 2005; MURPHY; OGILVIE; ARP, 2009). SWI uses flow direction and slope to infer the accumulation of humidity in the soil (GÜNTER; SEIBERT; UHLENBROOK, 2004). Conversely, DTW uses the DEM and the drainage network to create a new surface where the sloping of stream has been removed as if they ran on flat terrain but the relief away from the streams maintains its elevation difference relative to the nearest stream. Murphy, Ogilvie e Arp (2009) have shown that the DTW approach outperforms SWI by having a much larger agreement between the area mapped as "wet" and the actual soil wetness index of the study area and attributed this difference to the over-dependency of SWI on flow accumulation computed from the DEM.

1.2. Remote Sensing of Riparian Forests

The number of remote sensing publications exclusively dedicated to the study of riparian vegetation is relatively scarce. Remote sensing studies of riparian zones can be roughly divided in two types: 1) low and medium resolution studies (mainly Landsat) and 2) high resolution.

Most low- and medium-resolution studies of riparian vegetation used Landsat data to extract the limits of the vegetation under fluvial influence. Many, if not most, have reached relatively poor classification results (JOHANSEN; PHINN, 2006). For instance, Congalton et al. (2002) reached an accuracy of 30% while Baker et al. (2006) attained 40% while trying to classify riparian zones and/or wetland. Riparian forests are frequently narrow and rarely surpass a width of 100 m and this is often regarded as the main limiting factor for using medium-resolution data like Landsat (CONGALTON et al., 2002; JOHANSEN; PHINN, 2006), Radarsat (band C radar at 25 m) or even ASTER data (15 m) (MAILLARD; ALENCAR-SILVA; CLAUSI, 2008).

With the availability of high spatial resolution satellite data ($< 5m$), not only could riparian forests could be extracted with increased precision but some authors have demonstrated that the data could also be used to assess the biophysical structure of the riparian vegetation. For instance, authors have used texture analysis in conjunction with high-resolution data such as Ikonos and Quickbird to model structural parameters like canopy openness, leaf area index (JOHANSEN et al., 2007) and biomass (DILLABAUGH; KING, 2008). We believe that high resolution satellite data can contribute significantly to improve the delineation and characterization of riparian forests and other formations.

2. Material and Methods

The general idea of the riparian forest delineation process is illustrated in Figure 1 where the right side shows the DEM-based hydrological-geomorphological processing while the image processing is on the left side. The basic idea consists in limiting the search for riparian forest within a previously defined riparian zone.

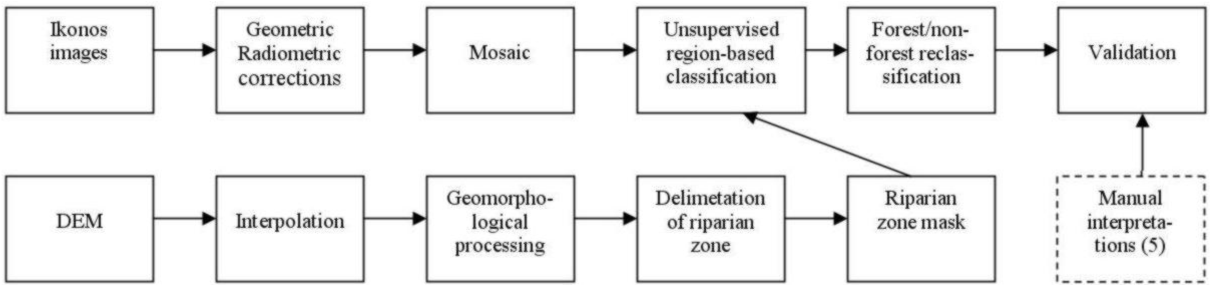


Figure 1: Generic flowchart of the delineation process of the riparian forest.

2.1. The "Depth to Water" Geomorphological Analysis Algorithms

The DTW model was developed by Murphy et al. (2006) and was chosen as the geomorphological processing algorithm. The DTW requires processing the whole watershed to be consistent. It takes two grids as input: an hydrographic grid identifying the location of all streams of the watershed and a slope grid. The model determines for each cell in the watershed the path of least slope to the nearest stream cell, it then computes a "cumulative slope value" for each cell in the watershed. Doing so, it takes both distance and slope into account. In Geomatics term, the DTW can be resumed to a "cost" surface between the DEM and the hydrographic network (Equation 1).

$$DTW = \left[\sum \frac{\Delta z_i}{\Delta x_i} \times \alpha\right] x_c \tag{1}$$

Where $\frac{\Delta z_i}{\Delta x_i}$ is the slope cell i, x_c is the cell resolution and α is equal 1.0 when the cell is contiguous by its side and $\sqrt{2}$ when the cell is diagonally connected. The process produces a surface where stream elevation difference (from headwaters to river mouth) have been removed so that a single elevation threshold can be set to serve as a riparian zone delimiter.

2.2. Image Segmentation and Classification with SPRING

A region-based unsupervised classification approach was adopted using tools available from the SPRING free software package (© 1991-2011, DPI/INPE). SPRING offers two segmentation algorithms: watershed and traditional region-growing with some modification (BINS et al., 1996). Since watershed tends to over-segment (MEINEL; NEUBERT, 2004) region growing was preferred. Region growing is initiated using single pixel seeds and proceeds by choosing and merging the most likely (similarity parameter) neighbour at each iteration. Similarity is based on the euclidian (D) distance between the mean vectors of the two regions (a single pixel at the first iteration):

$$D_{R_j,R_k} = ||M_i - M_k|| \tag{2}$$

Then, if:

$$D_{R_j,R_k} \leq D_{R_j,R_l} \text{ for every } l \in N(R_j), \tag{3}$$

R_k is merged with R_j. $N(R_j)$ represents the set of all contiguous neighbours of region R_j. The iterations stop when no "joinable" regions remain or when a certain size is attained (set by the user). In the last iteration, small regions are merged with larger adjacent ones, in accordance with the area threshold defined. Meinel e Neubert (2004) have evaluated and compared SPRING with e-Cognition and other commercial software packages and have reported good results but slower performance and difficult handling.

2.3. Study Area

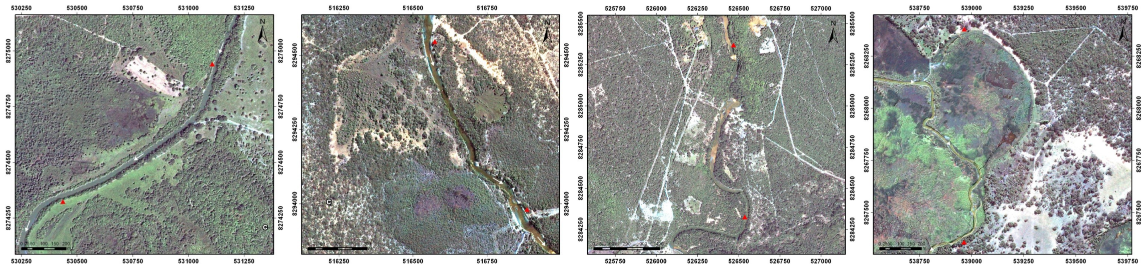
The study areas are situated along the margins of the Pandeiros River in Northern Minas Gerais, an environmental protection area. The Pandeiros River is an affluent of the *São Francisco* River. The total area of the Pandeiros watershed is 3921,00 km^2 and its elevation varies from 450 m to 850 m. The study sites are all about 1 km^2 along slightly meandering stretches of the river (Figure 2e). The climate is marginally semiarid with about 900 mm of precipitation and an average temperature of over 25°C. Precipitation varies from 124 mm per month between October and April to less than 2 mm between May and September.

Figure 2a-d shows Ikonos sub-scenes of the four study sites in true-colour composition: *Catolé*, *Balneário*, *Agropop* and *Pântano*. In the three first sites, both sides of the stream are considered. In the Pântano site, only the left margin was used since the right margin is in fact a wide wetland for which no field data was available. The green areas located along the river and on the bottom right of Figure 2a correspond to riparian forest and savannah formations, respectively. The light green to brown zones represent herbaceous areas. *Veredas*, characterized by a specific texture, can be seen on the left hand side of Figure 2a. The others light tone areas are bare soil. On the Ikonos scene riparian forest often appears similar to cerrado formations.

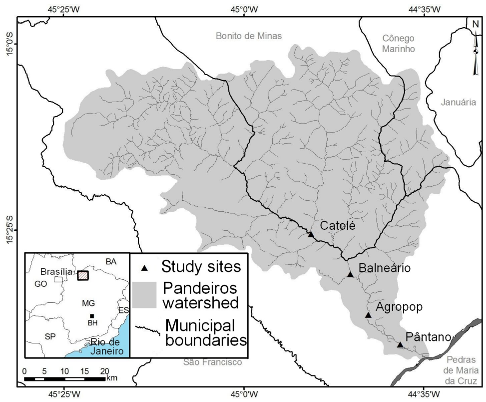
2.4. Validation

Since ground truth was rather limited for these four sites, we decided to compare the results with that of human interpreters. Five geography students (three graduated and two undergraduated in their last year) were asked to interpret the four sub-scenes with knowledge of the particular context of a *cerrado* landscape. They were to digitally trace the main vegetation classes on a computer screen using a GIS software. Amongst the different classes, they had to distinguish riparian forest. These students were picked based on their experience in image interpretation and were informed not to have contact between themselves. Upon completion, their individual interpretation were reclassified to contain but two classes: "riparian forest" and "other classes". The five thematic classifications were converted to raster and were compiled together using a majority processing. The class with the most "hit" won the process. Since the images had only two classes and an odd number of subject was used, no possible conflict remained and one of the two classes necessarily won the majority.

Then, difference images were computed between each of the results of the four sites and their respective "ground truth" image. For each pixel, when the class of the "automated"



(a) *Agropop* (b) *Catolé* (c) *Balneário* (d) *Pântano*



(e) Pandeiros watershed

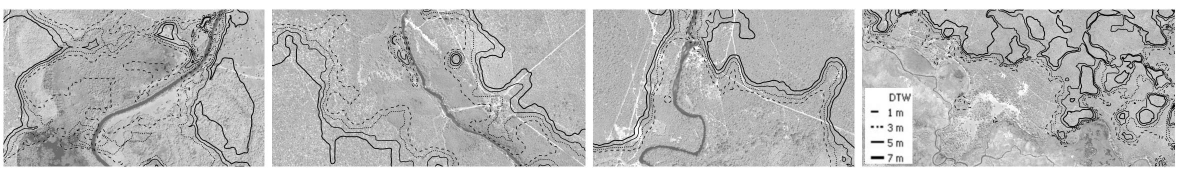
Figure 2: Ikonos subscenes of the four study sites and their location in the watershed.

classification was the same as the ground truth, a value of "2" would be set, otherwise, the pixel would be given value of "0" for omission errors and "1" for commission errors. The final score would be simply computed by summing all pixel with a value of "2" in the difference image.

3. Results and Discussion

3.1. Setting the "Depth to Water" threshold

The most significant limitation for our use of the DTW algorithm was attributed to the relatively poor resolution of the ASTER GDEM data with a cell size of 30 m. Because the topography of the region is rather smooth, it has a relatively predictable behaviour and we used a minimum curvature interpolation (a variation of bi-cubic spline) with a tension parameter to artificially resample the DEM grid to 4 m (the same as the multispectral Ikonos images). We realised that this is not ideal but since our goal was to demonstrate a concept of how to limit the search zone for riparian forest, we felt it was still a valid option. The results tended to prove us right and we found that a threshold of 5 m was a good approximation of the riparian zone in all four sites (Figure 3). To illustrate this, Figure 3 shows all four site with DTW thresholds of 1.0, 3.0, 5.0 and 7.0 m.

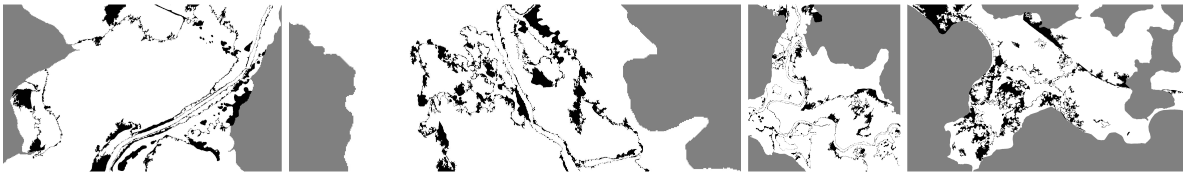


(a) *Agropop* (b) *Catolé* (c) *Balneário* (d) *Pântano*

Figure 3: "Depth to Water" thresholds of 1.0, 3.0, 5.0 and 7.0 m.

3.2. Results of the Unsupervised Region-Based Classification

One of the drawbacks of using region-based segmentation and classification resides in setting one or more parameters that usually control how spatial and spectral criteria interact in the iterative process. There is no systematic way to determine the parameter values that would yield the best results. So, in all cases, numerous tests were done with each parameter to determine the best combination. In addition to the segmentation parameters the number of classes has to be defined. Although only two classes are sought, it was necessary to set the number of classes to a larger value so that the within-class variance is kept to lower values and class confusion (spectral) is kept low. The strategy of using a greater number of classes than are really required comes from the simple fact that classes can easily be ultimately merged whereas "excessively-inclusive" classes cannot be split. Figure 4 shows the difference images generated by Boolean operators between the classification result of each site and the validation map produces by the interpreters. In the images, white represents the agreement, black, the difference and gray are areas outside the riparian zone. A large proportion of the mismatches are located on edges where chances of confusion are greater both for the classification program and the interpreters.



(a) *Agropop* (b) *Catolé* (c) *Balneário* (d) *Pântano*

Figure 4: Difference images between the classification approach and the validation map for the four sites. Legend: white=agreement; black=difference; grey=outside riparian zone (DTW threshold).

SPRING does segmentation and classification in two entirely different phases beginning by a region-growing segmentation. Apart from the number of class desired, two parameter need to be set for the segmentation to take place: "similarity" and "area". Values of "similarity" ranging from 5 to 100 and "area" varying from 1000 to 2500 were tested to determine the best results. It was found that in the case of SPRING, the initial number of classes of 7 to 12 always yielded the best results with 12 so we increased the maximum number of classes until results began to worsen. It was found that the best results were obtained with 10 to 30 classes depending on the site. Table 1 shows the parameters for all best results.

Table 1: Parameters and classification performance of the region-based classification using SPRING for the best results for each site.

Site	**Parameters**			**Class. Performance**		
	Similarity	Area	n classes	R.F.	O.C.	Overall
Catolé	10	2000	30	85.5	95.0	91.2
Balneário	30	1000	30	88.3	93.4	90.8
Agropop	20	2500	25	83.3	94.4	90.2
Pântano	5	1000	10	80.4	93.2	85.5

Results between 85.5 and 91.2% were considered excellent. The main drawback of SPRING is its inherent complexity of manipulation and a relatively slow processing. The segmentation in SPRING apparently tends to be dominated by the spatial component that make it merge spectrally different classes too easily. This is why we had to overcome this tendency by having ten classes so that riparian forest would not be merged with other land cover. This also implied some difficulties to chose and merge the right classes after the initial classification.

3.3. Discussion

The riparian zone delineation is clearly more complex than can be estimated using a fixed width from both margins of the stream. In most cases, delineating a riparian zone can prove difficult even in the field. This is especially true for the flatter parts of the river such as is the case for the *Pântano* and *Agropop* sites. The riparian zone was modelled using a fixed vertical elevation difference from the stream which is also a simplification but more appropriate than using an horizontal fixed distance. This was verified by the shape and width of the riparian forest in the four sites which behave very differently according to the geomorphological context. In the flat wetland part of the river, the riparian zone has a very erratic shape that does not appear to be much related to the path of the river. In more embedded section upstream (*Balneário*) the riparian zone and forest have a width and shape that can be more easily related to the stream course. Although still an approximation, the DTW approach is a much more realistic approach than a fixed width.

The region-based classification in SPRING yielded very good results with AN average score of 89.43%. The fact that the segmentation parameters need to be set by the user with little cue as to how to choose a value is perhaps the most important drawback in using segmentation. The same can be said for the number of classes. A systematic approach that worked well here and can be applied in any situation implies setting the number of class to a value larger than actually needed. How much larger will depend on the complexity of the scene and application.

The approach adopted in this article consisted in the first of two steps to 1) delineate and 2) characterize riparian forests of the Pandeiros river. It appeared clear to us that the delineation stage should take advantage of the peculiar position of riparian forests, not only horizontally from the river, but also vertically. This could possibly receive the name "hydrological distance" as it relates directly to the hypsometry as an hydrological variable. The second stage consists in the structural characterization of the riparian forest which has been simplified by not involving a separation criteria and which is being the object of another research article.

4. Conclusion

Having the objective of delineating riparian forests we propose a hybrid method based on a two-fold approach. First using the DEM and the hydrographic network, a Depth-to-water surface is created from which a riparian zone can be defined by a simple threshold. We propose the name "hydrologic distance" for this surface. Although this threshold is bound to vary from region to region, we were able to set a single value that applied to all our testing sites. The second step involves extracting the forested areas within the riparian zone using a region-based classification scheme in four test sites from the Pandeiros river with different characteristics. In order to minimize the effects of the particular parameters settings, a large number of tests were performed. SPRING yielded an average classification score of 89.43%. A special validation scheme using a voting principle of five independent interpretations was used to calculate accuracy. This proved a practical way to perform validation on whole regions (sites) instead of using small samples.

Acknowledgment

This work is part of the doctoral thesis of Thiago Alencar-Silva under the supervision of Philippe Maillard at the *Universidade Federal de Minas Gerais* and was made possible through the financial support of CAPES. The authors wish to thank Dr. Yule Roberta Ferreira Nunes and her team at the Department of Ecology of the *Unimonte* University for providing field data. ASTER GDEM is produced and provided free of charge by METI and NASA.

References

BAKER, C. et al. Mapping wetlands and riparian areas using Landsat ETM+ imagery and decision-tree-based models. *Wetlands*, v. 26, n. 2, p. 465–474, 2006.

BINS, L. et al. Satellite imagery segmentation: a region growing approach. In: *Proceedings of VIII Brazilian Remote Sensing Symposium*. Salvador, Bahia: [s.n.], 1996. p. 4.

CASE, B.; MENG, F.; ARP, P. Digital elevation modelling of soil type and drainage within small forested catchments. *Canadian Journal of Soil Science*, v. 85, p. 127–137, 2005.

CONGALTON, R. et al. Evaluating remotely sensed techniques for mapping riparian vegetation. *Computers and Electronics in Agriculture*, v. 37, n. 1-3, p. 113–126, 2002.

DILLABAUGH, K. A.; KING, D. J. Riparian marshland composition and biomass mapping using Ikonos imagery. *Canadian Journal of Remote Sensing*, v. 34, n. 2, p. 143–158, 2008.

GÜNTER, A.; SEIBERT, J.; UHLENBROOK, S. Modelling spatial patterns of saturated areas: an evaluation of different terrain indices. *Water Resource Research*, v. 40, p. 1–19, 2004.

JOHANSEN, K. et al. Application of high spatial resolution satellite imagery of riparian and forest ecosystem classification. *Remote Sensing of Environment*, v. 110, p. 29–44, 2007.

JOHANSEN, K.; PHINN, S. Mapping structural parameters and species composition of ripirian using IKONOS and Landsat ETM+ data in australian tropical savannahs. *Photogrammetric Engeneering and Remote Sensing*, v. 72, n. 1, p. 71–80, 2006.

MAILLARD, P.; ALENCAR-SILVA, T.; CLAUSI, D. A. An evaluation of radarsat-1 and aster data for mapping *veredas* (palm swamps). *Sensors (MDPI)*, v. 8, p. 6055–6076, 2008.

MEINEL, G.; NEUBERT, M. A comparison of segmentation programs for high resolution remote sensing data. In: *Proceedings of the XXth ISPRS Congress*. Istanbul: ISPRS, 2004. p. 1097–1102.

MURPHY, P. et al. Forest operations planning based on high resolution wet areas mapping: verifications. In: ACKERMAN, P.; LäNGIN, D.; ANTONIDES, M. (Ed.). *Proceedings of International Precision Forestry Symposium*. Stellenbosch, South Africa: Stellenbosch University, 2006. p. 477–488.

MURPHY, P.; OGILVIE, J.; ARP, P. Topographic modelling of soil moisture conditions: a comparison and verification of two models. *European Journal of Soil Science*, v. 60, p. 94–109, 2009.

NAIMAN, R.; BILBY, R.; BISSON, P. Riparian ecology and managementin the Pacific coastal rain forest. *BioScience*, v. 50, n. 11, 2000.

OLIVEIRa de SOUZA, A. M.; JOHANSEN, K. In: ARIZPE, D.; MENDES, A.; RABAçA, J. (Ed.). *Sustanainable Riparian Zones: a management guide*. Valencia, Spain: Generalitat Valenciana, 2008. p. 148–156.

POLVI, L.; WOHL, E.; MERITT, D. Geomorphic and process domains controls on riparian zones in the Colorado front range. *Geomorphology*, v. 125, p. 504–516, 2011.

VERRY, E.; DOLLOFF, C.; MANNING, M. Riparian ecotone: a functional definition and delineation for resource assessment. *Water, Air and Soil Pollution*, Focus 4, p. 67–94, 2004.