Land cover change dynamics and multi-factor analysis in high mountains basins of Colombian Andes

Leonairo Pencue-Fierro
Grupo de Óptica y Láser, Universidad del Cauca
Red Interinstitucional del Cambio Climático y Seguridad Alimentaria - RICCLISA
Popayán, Colombia
leonairo@unicauca.edu.co

Apolinar Figueroa-Casas
Grupo de Estudios Ambientales, Universidad del Cauca
Red Interinstitucional del Cambio Climático y Seguridad Alimentaria - RICCLISA
Popayán, Colombia
apolinar@unicauca.edu.co

***Abstract*—This paper presents an integrated analysis of factors affecting the dynamics of an Andean basin. The study was conducted by using Landsat images, SRTM DEM data from NASA, infrastructure information (cities and roads proximity) and the resulting hydrological data analysis, over a characteristic basin of Colombian valleys. Land cover recognition was validated with *in-situ* data, this allowing to build a temporal variation profile of the land cover over a 26 years period. This information was integrated by clustering to obtain uniform regions according to variables used, resulting in the generation of different trend maps. The proposed approach allows to establish the importance of using information as a measure of anthropogenic affectation levels, thus it is possible to make more specific action plans and efficiently compensate the environmental degradation effects.**

Keywords—Land cover dynamics, multi-factor analysis, clustering analysis, high mountain basins

I. Introduction

The transformation humans exerts on the territory is one factor that offers greater challenges in determining the sustainability of basins ecosystem services. However that territory is determined by factors such as topography, land use dynamics, meteorology, position on the globe, etc., this is leading to understanding and modeling are more difficult [1]. Remote sensing analysis techniques in recent decades have made possible to integrate the information above described. Thus, it has helped to establish new research and effective methods to identify trends and prepare action plans to stop or revert unwanted effects or threaten stability basins [2].

Specifically, the study of the basins dynamics by hydrological models provides useful tools for understanding and analyzing multiple processes occurring on them, allowing the development of prospective analyzes and, with these, to provide integrated systems for decision making. In this same way, integration with land cover change models can determine a analysis basis for determining the long-term trends that establish future scenarios of mountainous areas basins [2],[3].

This paper presents the results of a research about dynamics of land cover change considering integrated analysis by including the multiple factors influencing the current state of the lands. The study has focused on the high mountain basins such as the Colombian Andes ones, characterized by steep hillsides and marked precipitation regimes in environments affected by deforestation and areas sensitive to floods and landslides.

Multi-factor analysis methods are often used as tools for environmental planning and assessments [4]. In this study, the effect of six main factors: i) altitude, ii) slope, iii) land cover, iv) significant change regions, v) cities and roads normalized distance and vi) rivers normalized distance, were selected to analyze the even regions according to contextual information. The multi-factor analysis allows integrating information from different sources using artificial intelligence techniques to establish indicative cluster in even areas, whose behavior is established in terms of the factors involved. In this paper we include aspects of socio-economic infrastructure, such as cities and roads proximity as decisive factors that influence the environmental balance.

II. Study Area Description

In Colombian geomorphology mountain ranges and valleys predominate, and on these lives 70% of the national population, becoming the main region of socio-economic activity in the country. The valleys are basins resulting from geodynamics that formed the mountain ranges and are not only the result of fluvial erosion.

The Colombian Andes basins are characterized by having steep slopes and heavy rainfall regimes at certain times of year, so they are likely to produce suddenly floods, mainly due to loss of retention capacity, given by indiscriminate deforestation to convert lands in space to crops and rangelands. Also, no river buffer areas (riparian forests) are respected, worsening the overall picture of the basins.

Colombian Andean region consists of two main valleys: the Cauca River valley and Magdalena River valley that cross the country from South to North between Central, Western and Eastern mountain ranges. The different tributary watersheds of these rivers are catchment areas that feed them and have become typical Colombian mountainous landscape. For this

This work is funded by the Project "Programa para el Fortalecimiento de Red Interinstitucional de Cambio Climático y Seguridad Alimentaria" – RICCLISA and Universidad del Cauca.

978-1-4673-7119-3/15/$31.00 ©2015 IEEE

study we have chosen an intermediate pattern catchment basin of the Cauca River, specifically in the upper basin, located on the western slopes of the Central mountain range. This is the Palace River basin, in of Cauca Department.

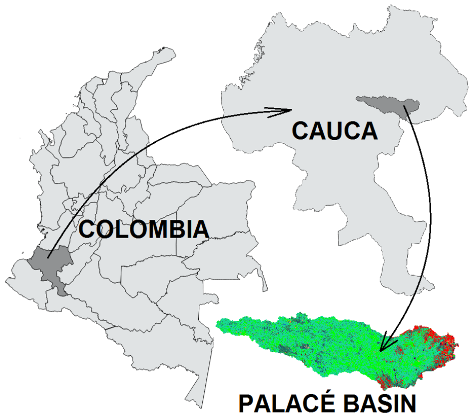

Fig. 1. Study area for Andean basins

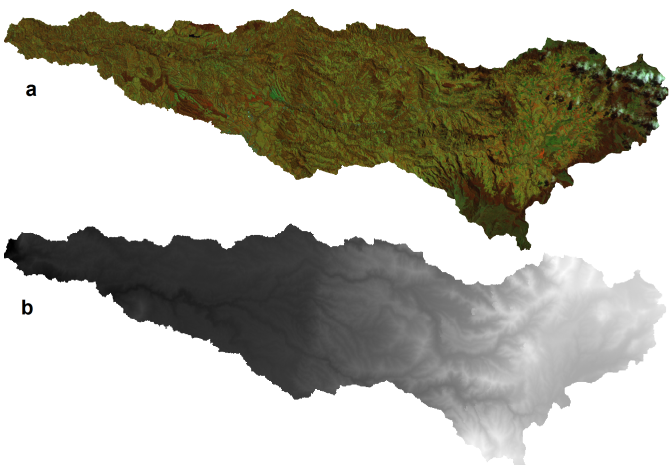

Fig. 2. (a) Landsat 453 composition, (b) DEM map

The Palace River basin has a size of 645 km^2 and is located between 1300 and 3700 meters above sea level, has an annual rainfall of 1770 mm, relative humidity of 79% and an average temperature of 15°C. Fig. 1 shows the location of the basin in Colombia and in the Cauca Department. Land cover distributed in the basin are: páramo (a very important intertropical high mountain ecosystem which it is situated about 3000 meters above sea level) vegetation, dense forest, open forest, lagoons and marshes, rural construction, annual crops, land in transition, natural pastures, natural grassland associated with trees, planted forests and bare soils. Natural land cover occupies 58.4% of the area of the basin and is responsible for the ecological, water, climate and edaphic balance of the region. The remaining area is being exerted by human pressure activities [5].

Fig. 2 shows the enhanced Landsat 453 composition (a) and elevation map (b) of the same basin. What we can see is the diversity of land cover regarding mainly forests, crops, páramos and rangelands. From the elevation map there can be observed 3 main areas: the highlands (right, lighter), the intermediate land in the center, marked by the river canyons and its tributaries and, in the bottom side (left, darker), when the river runs in the valley before flows into the Cauca river.

III. Proposed Methodology

A. Satellite data

The multispectral Landsat data was used as source for derivation of land cover information. We selected images acquired between 1989 and 2015, all these multispectral data are nearly cloud-free and at similar phenological conditions. To classify the multispectral data, we used a supervised classification algorithm to identify the landcover's classes, a series of ground truth data were used as reference for the classification validation. The Normalized Difference Vegetation Index (NDVI) it is also generated for estimate the quantity, quality and vegetation development.

For hydrology characterization and analysis, we used the Shuttle Radar Topography Mission (STRM) data from NASA. The SRTM elevation data (or Digital Elevation Model, DEM) had to be previously filtered because it has a series of impulsive and small noisy area, especially in mountain regions. This involved the implementation of an algorithm for detecting and compensating errors in DEM maps. The algorithm for this task consists in detecting very abrupt changes, the correction of the artifacts consists of making successive dilations on the detected area from the neighborhood data to cover the entire detection. The DEM filtered data is then used to determine a set of indices of interest for analyzing: slope, transverse and vertical curvatures (horizontal and vertical concavities), local minima for streams of water, local maxima for watersheds, roughness, fractality, and all hydrological interest primitives as flow accumulation, catchment areas, topographic wetness index (TWI), and flow velocity. Thus, proximity to water maps are obtained, these will be part of the factors used later.

B. Human disturbances

The integration of information must be supplemented by anthropogenic available data such as the road infrastructure. Previous studies have established how a road can set conditions for human settlements and affect the surrounding area [6]. This way, auxiliary maps of settlements and road infrastructure, such as roads, were used to make proximity maps. These maps were modulated with slopes and land cover classes in different scenarios to establish the capacity to influence in terms of local conditions. This approach allows to obtain influence maps according to the parameters to be taken into account for the study of the basin. Thus, it was possible to establish the importance of the driving forces for reshaping the land and how the landscape has been established through the time.

C. K-means clustering

The clustering is one of the most versatile methods for data analysis, as in pattern recognition, data mining, image analysis, etc. It is valuable in remote sensing analysis too, where regions with similar features (regular areas regarding the factors) are clustered together without any previous knowledge. It has become the important algorithm of unsupervised classification, providing maps of regions with similar behavior, with which we can make further analysis.

For clustering calculation, an error function must be perform. The error function used is the sum of the distances that each point is from its cluster's centroid. Assuming that there are k clusters, the error function (E_k) is defined as:

$$E_k = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \sum_{m=1}^{M} \left(F_{i,jm}^{*} - C_{im} \right)^2 \quad (1)$$

where E_k is the error function, $F^{*}_{i,jm}$ is the m^{th} feature value, C_{im} is the centroid of the m^{th} feature of the i^{th} cluster, thus $(F_{*i,jm}-C_{im})^2$ is the distance measure between a data point and the cluster centroid to which it is assigned and n_i is the number of data points in cluster i. The feature values for this case are the maps of altitude, slope, changes, land cover and proximity factors for each point (feature vector from each pixel).

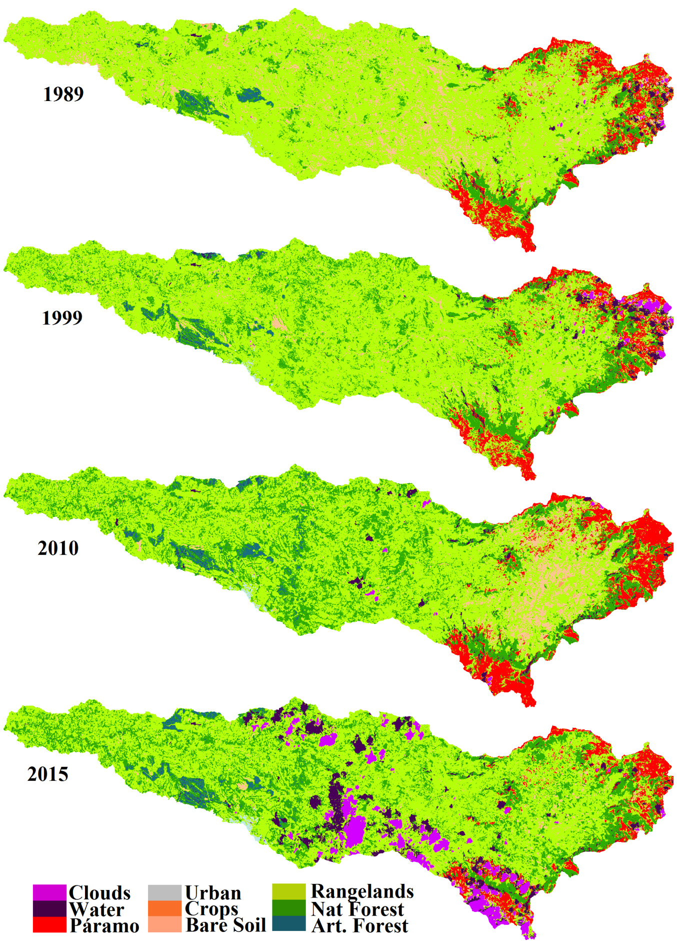

Fig. 3. Land cover changes for years 1989, 1999, 2010 and 2015

IV. Experimental Results and Discussion

The selected multispectral images were processed and analyzed by obtaining land cover maps for four years as shown in Fig. 3. In this figure we can observe changes in the different land covers in the study area. There are many elements to see in detail, but first significant changes are observed in the páramo and rangeland. A more detailed analysis shows influences regarding the geomorphological and hydrological characteristics.

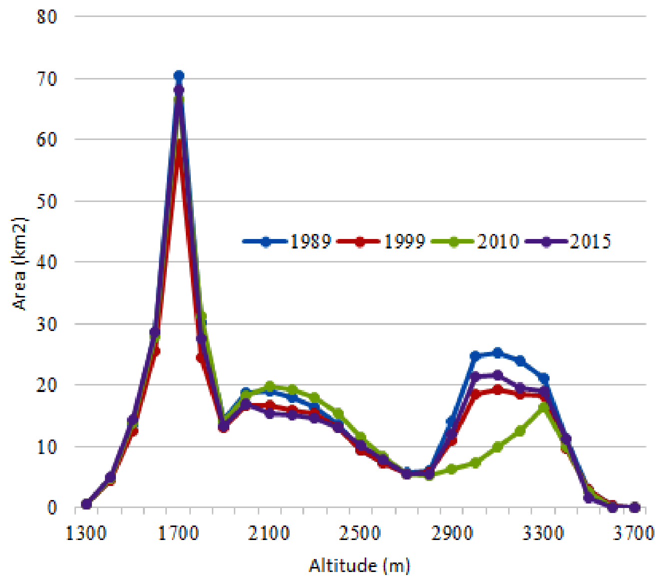

Fig. 4. Evolution of land cover (Natural forests)

Fig. 4 shows the evolution of forest cover in the same basin of Fig. 3. In this case, we can verify changes according to altitude over the years. Generally, stable scenarios are observed at lower altitudes, less than 2800 meters. There are significant variations in specific ranges and between 2800 m and 3500 m, an aspect which primarily explains the variation forest and páramo cover seen in Fig. 3. It is important to note that it was not possible to obtain reliable data in other years due to the almost constant presence of clouds over the area. In the 2015 image we observe some cloud clusters. These clusters are detected to generate a mask that allows comparative analysis regardless the areas covered in the other years analyzed. Major changes were observed in the dynamics of natural forests, this is due to two main factors: human intervention through deforestation and affectation of ENSO phenomenon of 2008-2009 that was particularly strong.

This allows us to identify covers that change more noticeably. The most sensitive are forests and rangeland, for this study we took into account changes in forests because its decrease is due to the rangeland growth for livestock pastures, so they are complementary. With this we were able to build a map showing the most important changes for the time period analyzed. For this, change points of the forest cover were accumulated in the years available and consolidated into a general map of change susceptibility. In this way was possible to construct a map of the points of major changes as a factor in the clustering analysis.

After obtaining the primary data it is possible to generate the clusters that describe the landscape in terms of the factors involved. Two types of tests were performed: (1) including the cities and roads proximity factor (anthropogenic factors) and (2) neglecting this factor. For land cover, a factor was included in (1) generic terms (NDVI) and in (2) terms of the most affected natural cover (natural forests).

This proposal shows the generation of clustering regions taking into account the different factors and their importance for obtaining areas of interest. The number of clusters (*k*) was 20 for the whole scene. This number was chosen because it allowed the scene classifying in representative sets previously known in field measures. Fig. 5 and 6 show the results for clustering analysis. The colors only represent clusters established by the algorithm but should not be interpreted as a specific land cover. What matters here is the regions segmentation and the contextual meaning according to the dynamic factors included in each case.

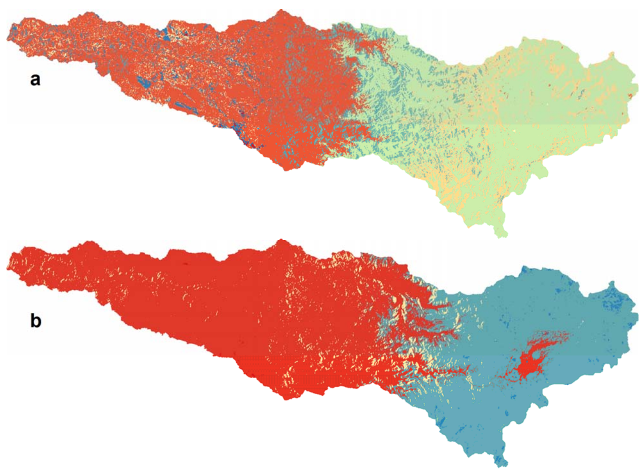

Fig. 5. Clustering analysis with NDVI (a) without proximity factors (b) with all factors

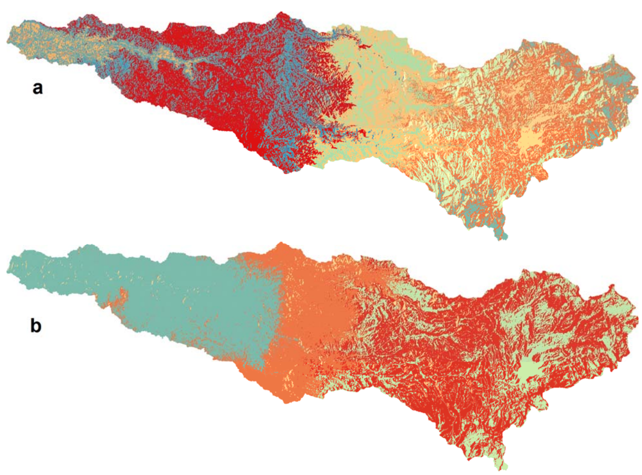

Fig. 6. Clustering analysis with natural forests land cover (a) without proximity factors (b) with all factors

In Fig. 5 we can see clustering analysis with generic land cover factor, first the analysis (a) without proximity factors, two main groups divide the basin into halves, there are some isolated areas in the region corresponding to dynamic conversion of forests to crops or rangelands. For case (b), including the anthropic factor, there is a tendency to uniform segmentation, expanding the lower region, obtaining fewer classes, which means that, from the point of view of the proximity factor, the interventions in this region are more likely notable, remaining differentiated only in the upper basin. The red cluster on the right side could draw the attention and corresponds to a intervened area with highly intensive crops.

In Fig. 6 forests factor (instead of the NDVI), as an important agent for the integrity of the basin is used in the clustering analysis. In (a) we observe structured zones in terms of geomorphology in which we can differentiate final stretch of the river as a distinct cluster; in this case more diversity of regions is obtained in the studied areas, indicating that more detailed analysis must be done. In the case of (b) three main areas are produced and again we get a level of uniformity warning about susceptibility to human disturbance. What we can see from this analysis is that the landscape has many facets that are appearing in terms of integrated variables and it is important to perform other inspections to obtain a tool to make better decisions on land management.

V. Conclusions

A proposal for factors integrated analysis from various sources was presented. Major intervention regions were obtained by analyzing changes oriented towards natural forest cover. Segmentation of multi-factor regions was performed and cluster regions were obtained. The inclusion of anthropogenic factors shown to be important to assess the possibility of more significant degradation due to the standardization of the areas with the variables involved. This will allow more realistic improvement plans and greater effectiveness in the complexity of Andean ecosystems. Further analysis should be performed to compare the results with other basins systems.

Acknowledgment

This work was developed within the project “Programa para el Fortalecimiento de la Red Interinstitucional de Cambio Climático y Seguridad Alimentaria” – RICCLISA, funded by Colciencias and Universidad del Cauca. Authors acknowledge the support of Universidad del Cauca to carry out this research.

References

[1] N. Lu, A. J. Hernandez and R. D. Ramsey. “Land cover dynamics monitoring with Landsat data in Kunming, China: a cost-effective sampling and modelling scheme using Google Earth imagery and random forests” Geocarto International, 2015.

[2] D. Lu, S. Hetrick, E. Moran and G. Li, “Application of Time Series Landsat Images to Examining Land-use/Land-cover Dynamic Change” Photogrammetric Engineering and Remote Sensing. Vol. 78, No. 7, July 2012.

[3] A. B. Brink and H. D. Eva, “Monitoring 25 years of land cover change dynamics in Africa: A sample based remote sensing approach”, Applied Geography, Volume 29, Issue 4, December 2009.

[4] N. Shan, X. Ruan and J. Ao. “RS and GIS Based Temporal-spatial Variation and Multi-factor Spatial Analysis on Nonpoint Source Pollution”, Geoinformatics, 2010 18th International Conference, pp.1,4, 18-20. 2010.

[5] http://www.acueductopopayan.com.co/gestion-ambiental/fundacion-procuenca-rio-las-piedras/fuentes-de-abastecimiento/cuenca-palace/. Acueducto y Alcantatillado de Popayán SA ESP. (March 16, 2015)

[6] N. Rodríguez Eraso. Deforestación y cambio en la cobertura del suelo en colombia: Dinámica espacial, factores de cambio y modelamiento. PhD Thesis, Universidad Autónoma de Barcelona. 2011.