UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA

FACULTAD DE ZOOTECNIA Y ECOLOGÍA

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

MODELACIÓN Y ANÁLISIS GEOESPACIAL DE FACTORES DETERMINANTES EN LA PRODUCTIVIDAD DE LOS ECOSISTEMAS, EN EL ESTADO DE CHIHUAHUA, MÉXICO

POR:

M. C. JESÚS ALEJANDRO PRIETO AMPARÁN

TESIS PRESENTADA COMO REQUISITO PARA OBTENER EL GRADO DE DOCTOR IN PHILOSOPHIA CON ÁREA MAYOR EN RECURSOS NATURALES

Chihuahua, Chih. **Marzo 2020**

Modelación y análisis geoespacial de factores determinantes en la productividad de los ecosistemas, en el Estado de Chihuahua, México, disertación presentada por Jesús Alejandro Prieto Amparán como requisito parcial para obtener el grado de Doctor en Philosophia, ha sido aceptada y aprobada por:

Ph. D. Carlos Ortega Ochoa
Director de la Facultad de Zootecnia y Ecología

D. Ph. Agustín Corral Luna
Secretario de Investigación y Posgrado

Ph. D. Iván Adrian García Galicia
Coordinador Académico

D. Ph. Alfredo Pinedo Alvarez
Presidente

12 / MARZO / 2020

Fecha

Comité:

D.Ph. Alfredo Pinedo Alvarez
Dr. Martín Martinez Salvador
Ph.D. Federico Villarreal Guerrero
Dr. Eduardo Santellano Estrada
Dr. Carmelo Pinedo Alvarez

AGRADECIMIENTOS

A Dios, por su sabiduría y la oportunidad de realizar un Doctorado.

Al Consejo Nacional de Ciencia y Técnología (CONACyT), por el financiamiento que permitio la realización del Doctorado.

A la Facultad de Zootecnia y Ecología (FZyE).

A la Universidad Autónoma de Chihuahua (UACH).

Al Dr. Alfredo Pinedo Alvarez, por sus enseñanzas y consejos.

Al Dr. Federico Villarreal Guerrero, por ayuda en la formación de escritura y paciencia.

Al Dr. Martín Martínez Salvador, por su paciencia y enseñanzas.

Al Dr. Eduardo Santellano Estrada, por su enseñanza en el área estadística, paciencia y dedicación.

Al Dr. Carmelo Pinedo Alvarez, por sus reflexiones, paciencia y enseñanzas.

Al Proyecto Tarahumara Sustentable

Al Centro de Investigaciones en Geográfía Ambiental (CIGA), donde me permitieron llevar a cabo mi estancia doctoral.

DEDICATORIA

Compañera de vida:

Rosalinda Ibarra Lopez

“Por su apoyo y paciencia incondicional en esta etapa”

A mis padres y hermano:

Elvira Amparán Hernandez

Socorro Alejandro Prieto Aguayo

Elías Prieto Amparán

“Quienes han estado en los cimientos de mi desarrollo personal y profesional”

A mi familia:

“Quienes me han apoyado desde joven”

A los amigos:

“Quienes han estado presentes, apoyándome indoncionalmente”

CURRICULUM VITAE

El autor nació el 6 de enero de 1991 en la ciudad de Chihuahua, Chihuahua, México.

2009 - 2013	Estudios de Licenciatura en el Ingeniero en Desarrollo Territorial de la Facultad de Ciencias Agrotecnológicas.
2014 – 2015	Estudios de Maestría en Ciencias en Producción Animal y Recursos Naturales de la Facultad de Zootecnia y Ecología.
2015 – 2016	Participación en el Proyecto “Sistema de Monitoreo de Datos e Información de la Sierra Tarahumada (SMDIT-ST)”, a cargo del WWF.
2016 – 2019	Estudios de Doctor in Philosophia en Producción Animal y Recursos Naturales de la Facultad de Zootecnia y Ecología.

Miembro de la Sociedad Latinoamericana de Percepción Remota y Sistemas de Información Espacial (SELPER).

Miembro de la Sociedad Mexicana de Manejo de Pastizales (SOMMAP).

GENERAL ABSTRACT

MODELING AND GEOSPATIAL ANALYSIS OF DETERMINANTS OF ECOSYSTEM PRODUCTIVITY IN THE STATE OF CHIHUAHUA, MEXICO.

BY:

M. C. JESUS ALEJANDRO PRIETO AMPARAN

The spatial modelling and analysis of natural resources is an increasingly urgent need. The results of geospatial modeling can contribute to a better understanding of the current, past and future state of natural resources. The objective of this work is to model and analyze five regions of the state of Chihuahua using geographic information systems, remote sensing data, field information, and geospatial data. The work consisted of five case studies. In the first study, grassland biomass was correlated with spectral data from the Landsat sensor. In the second study, scenarios of land use/land cover change in the forest region of San Juanito were generated. In the third study, a set of Rio Conchos basins were analyzed using multivariate techniques and the compound parameter to find those basins with greater susceptibility to erosion. In the fourth study, the productivity of the forests of the Chinatu Ejido was analyzed, as well as their spatial distribution. The fifth study combined interpolation techniques and multivariate analysis for a set of wells in the Delicias 005 irrigation district, where seven physicochemical parameters were analyzed, to determine their spatial distribution. The results of the five studies can contribute to the analysis of large areas of land and low cost, to know their current status to take priority actions to avoid their deterioration.

RESUMEN GENERAL

MODELACIÓN Y ANÁLISIS GEOESPACIAL DE FACTORES DETERMINANTES EN LA PRODUCTIVIDAD DE LOS ECOSISTEMAS, EN EL ESTADO DE CHIHUAHUA, MÉXICO

POR:

M. C. JESÚS ALEJANDRO PRIETO AMPARÁN

Doctor en Philosophia de Recursos Naturales

Secretaría de Investigación y Posgrado

Facultad de Zootecnia y Ecología

Universidad Autónoma de Chihuahua

Presidente: Dr. Alfredo Pinedo Alvarez

La modelación y análisis espacial de los recursos naturales es una necesidad cada vez más urgente. Los resultados de la modelación geoespacial pueden contribuir a una mejor comprensión del estado actual, pasado y futuro de los recursos naturales. El objetivo de este trabajo es modelar y analizar cinco regiones del estado de Chihuahua mediante sistemas de información geográfica, datos de sensores remotos, información de campo y datos geoespaciales. El trabajo consistió en cinco estudios de caso. En el primer estudio, se correlacionó biomasa de pastizales con datos espectrales del sensor Landsat. En el segundo estudio, se generaron escenarios de cambios en la cobertura de uso de suelo en la región forestal de San Juanito. En el tercer estudio, un conjunto de cuencas del Río Conchos fue analizadas mediante técnicas multivariadas y el parámetro compuesto para encontrar aquellas cuencas con mayor susceptibilidad a la erosión. En el cuarto estudio, se analizo la productividad de los bosques del Ejido

Chinatu, así como su distribución espacial. El quinto estudio combinó las técnicas de interpolación y el análisis multivariado para un conjunto de pozos en el Distrito de Riego de Delicias 005, donde siete parámetros fisicoquímicos se analizaron para conocer su distribución espacial. Los resultados de los cinco estudios pueden contribuir en el análisis de grandes superficies de terreno a bajo costo, para conocer su estado actual y emprender acciones prioritarias tendientes a evitar su deterioro.

CONTENT

LIST OF TABLES

LIST OF FIGURES

GENERAL INTRODUCTION

Natural resources are being altered at an increasingly rapid rate by human activities, such as the combustion of fossil fuels, deforestation, the increase of the agricultural frontier over large pastures, the alteration of the water cycle, among others (Camacho-Olmedo *et al*., 2018).

In this context, modelling and spatial analysis of the current state of the environmental system over time is today a great challenge. This research challenge has several fundamental aspects, including understanding the causes of environmental degradation and analyzing the drivers of change, such as degradation. This supports decision-making in natural resource management and spatial planning (Peagelow y Camacho-Olmedo, 2008).

With the development of computational technologies, multiple computer models have been generated, which support large amounts of data. The processing power of these technologies has made it possible to analyze these data in less time. Currently, the most convenient way to manage and analyze this set of data is through geographic information systems (GIS).

GIS offers the possibility of relating the location data of natural resources with their quantitative and qualitative descriptive characteristics. This offers an integral and geospatial vision of the information, which allows to improve the analytical techniques, including the statistics and the geostatistics (Bocco, 2004). Among these data, there are those derived from remote sensors, field information, geospatial information, among others, which have a territorial context. The combined use of specialized software and geospatial information has a farreaching impact on the mapping, monitoring and management of natural resources on a sustainable basis (Reddy and Singh, 2018).

The spatialization of the current state of natural resources in northern Mexico, especially the state of Chihuahua, makes it possible to detect the current problem of its resources. However, few studies have analyzed the distribution of natural resources in the State of Chihuahua (Manjarrez *et al*., 2015), some others have determined the distribution of the natural potential of some of their ecosystems (Martínez-Salvador *et al*., 2019) and a few others have modeled their possible behavior (Prieto-Ampáran *et al*., 2019).

Therefore, the modeling and analysis of natural resources with a spatially explicit context is of great importance. Information on the nature, extent and spatial distribution of natural resources is a prerequisite for their mapping, monitoring and management on a sustainable basis.

The objective of this research is to model and analyze five ecosystems of the state of Chihuahua, Mexico, through geospatial technologies to know their productivity and generate scenarios of possible behaviors. The specific objectives are: 1) the comparison of radiometric correction methods in the estimation of grassland biomass, 2) evaluate and simulate changes in forest cover in the forest region of the state of Chihuahua, 3) describe the behavior of 31 sub-basins, distributed in the Rio Conchos Basin, based on their geomorphometric parameters to find the sub-basins with greater susceptibility to erosion, 4) analyze the productivity of the Ejido Chinatu forest through a set of variables with spatial context and field data, using multivariate techniques, and 5) analyze eight physicochemical parameters in well water samples from the 005 Irrigation District in Chihuahua, using multivariate techniques and interpolation methods to know its spatial distribution.

LITERATURE REVIEW

Grassland, forest, among others ecosystems play an important role as they provide a wide variety of environmental products and services for human well-being (Yang *et al*., 2009; Weiskittel, Crookston and Radtke, 2011) such as soil retention (Borrelli *et al*., 2017), water yield (Sun *et al*., 2006), and are the habitat of many species of flora and fauna.

However, these ecosystems have experienced notorious degradation in recent years, caused mainly by anthropogenic activities. In turn, these activities have modified the climate regime, prolonging droughts, fragmenting the habitat and reducing its border. The application of improper management practices has been one of the main causes of such degradation (Craine, 2013; Manjarrez *et al*., 2015).

In one hand, grasslands in northern Mexico are extensively used for cattle grazing and have also experienced intensive land-use change due to agricultural activities, which has resulted in changes of the species composition on such ecosystems (Estrada-Castillon *et al*., 2010; Manjarrez *et al*., 2016). In other hand, temperate forests of Mexico occupy 17 % of the national territory, represented by 32 million hectares. In this region, the greatest association of pine and oak forests in the world occurs (González, 2012). Around 23 different species of pines and close to 200 species of oaks live in the ecoregion of Sierra Madre Occidental (Navar, 2009). However, 40 thousand hectares of forests get on average lost annually. This region has the highest deforestation rate in the world (Velázquez *et al*., 2002; Mas *et al*., 2004).

Accurate monitoring and evaluation of these ecosystems is critical for their conservation, and their restoration (He *et al*., 2005; Vazquez-Quintero *et al*., 2016). Even though traditional methods such as field sampling, are costly, as well as time-and labor-consuming, when large pieces of land have to be covered (Xie *et al*., 2009). With the aim of developing more effective monitoring methods, there have been numerous studies on indirect methods that are very useful for analyzing the state of ecosystems, using remote sensing information (Goetz *et al*., 1999; Li *et al*., 2003; Zhao *et al*., 2016), models of land use/land cover changes (LULCC), commonly employed to quantify deforested surfaces, measuring the degree of change in the ecosystem (Lapola *et al*., 2011). Regression methods suchs as the logistic regression have been employed to generate models of LULCC. These models suppose that the relationship between the LULCC and the variables that produce it is a logistic function (Mas *et al*., 2010; Mas *et al*., 2014). Geostatistical-multivariate techniques allow us to study the spatial variability as well as the relationships that may exist among characteristics of the territory, with a special scale, wide distribution and low cost (Prieto-Amparán *et al*., 2018).

Remote Sensing in the Analysis of Natural Resources

The Landsat satellite has provided data since 1972 (Cohen and Goward, 2004) with an extensive global coverage. This is an important resource for monitoring global environmental change (Woodcock *et al*., 2001; Hansen and Loveland 2012; Wulder *et al*., 2008).

However, the satellites usually present a problem of saturation in the images, considered an important factor (Chander *et al*., 2009; Tan *et al*., 2012; Roy *et al*., 2016). This phenomenon of saturation is caused by atmospheric dispersion and absorption, which has substantial variations with time, space, and wavelength. In addition, the surface reflectance is strongly affected by the elevation of the terrain, a more evident effect in mountainous environments (Körner *et al*., 2008).

Different atmospheric and radiometric correction methods have been developed to transform the original values of the images into values of reflectance (Kaufman *et al*., 1997; Chavez, 1988; Janzen *et al*., 2006). Several of these techniques have been developed to estimate the amount of atmospheric bias from thin-surface terrestrial images by assessing the molecular and aerosol dispersion (Lyapustin *et al*., 2004; Richer *et al*., 2006).

Modelling Changes in Forest Land Use Cover

The study of the land use/land cover changes (LULCC) has become a fundamental research topic, since the change in land use/land cover (LULC) affects forest ecosystems and their biodiversity (Gharun *et al*., 2017).

The spatial modeling is a technique contemplating alternative scenarios of LULCC, which could contribute to better explain the key processes influencing LULCC (Pijanowski *et al.,* 2002; Eastman *et al*., 2005; Torrens, 2006; Perez-Vega *et al*., 2012). Thus, one of the main functions of the LULCC models is the establishment of scenarios, with the aim of changing policies and inadequate practices for the sustainable management of natural resources (DeFries *et al*., 2007; Berberoğlu *et al*., 2016).

Several approaches to establish LULCC scenarios have been developed and tested to generate scenarios of LULCC. Ferrerira *et al*. (2013) generated deforestation scenarios to 2050 in the central Brazilian savanna biome finding the possible increase of 13.5 % in deforested areas. Galford *et al*. (2015) used Bayesian Weights of Evidence for policiy scenarios from 2010 a 2050 evaluating plans for agriculture and forest in Democratic Republic of Congo.

Multivariate Techniques in the Spatial Analysis of Natural Resources

Management of soil and water, implies the characterization of the ecosystems inside the watershed and the understanding of the relationships between uplands, lowlands, land use/land cover, geomorphic processes, slope, and soil (Chen *et al*., 2011; Rahaman *et al*., 2015). In watershed management, erosion control is one of the main components (Gajbhiye *et al*., 2015). Thus, the hydrological planning and monitoring of a watershed is important for the development of environmental policies (Sharma *et al*., 2014).

The analysis of morphometry is often carried out based on geographic information systems (GIS, Shrimali *et al*., 2001; Thakker and Dhiman, 2007; Sharma *et al*., 2010; Viramontes-Olivas *et al*., 2008; Tilahun *et al*., 2014). On a spatial scale, morphometric parameters, i.e., the Gravelius compactness coefficient (Zavoianu, 1985) elongation ratio (Shumm, 1956), among others, are important, to know the hydrological configuration of watersheds.

The relationships among these parameters are useful for developing hydrological models, which allow prioritizing watersheds based on their condition, such as erosion susceptibility. To determine the aforementioned relationships, statistical methods, such as multivariate techniques, have been widely used worldwide (Saha *et al*., 2012; Sharma *et al*., 2013).

Within these multivariate techniques we find the princpal component analysis (PCA) and group analysis (GA) (Miranda *et al*., 1996; Castillo-Rodriguez *et al*., 2010; Bateyneh and Zumlot, 2012; Oketola *et al*., 2013; Tritsch *et al*., 2016; Prieto-Amparán *et al*., 2019), as well as multivariate analysis of variance (MANOVA) and the ranking methodology known as compound parameter (*Cp*) (Altaf *et al*., 2014) which have been widely used in recent years for the analysis of environmental data from watersheds. These techniques assist with analyzing the spatial variability of watersheds, their structure, as well as the relationships existing among them.

Interpolation and Multivariate Analysis on Water Quality

Water quality is an important factor that affects human health and ecological systems (Qadir *et al*., 2008). In the rural context, groundwater is the support of agricultural irrigation and it is essential for providing additional food security resources (Morris *et al*., 2003). However, food security can be affected by pollutants present in the irrigation water, causing serious clinical and physiological problems to humans when such pollutants get accumulated in large amounts (Sharma *et al*., 2007; Khan *et al*., 2008).

The evaluation of water quality in most countries has become a critical issue in recent years (Varol and Davraz, 2015). Water quality is subject to constant changes due to seasonal and climatic factors (AlSuhaimi *et al*., 2017). Likewise, spatial variations emphasize the need for water monitoring that provides a representative and reliable estimate (Muangthong and Shrestha, 2015).

Multivariate techniques and exploratory data analyses are appropriate for the synthesis of data and its interpretation (Singh *et al*., 2005). Classification, modeling and interpretation of the monitored data are the most important steps in the evaluation of water quality (Boyacioglu, 2006; Zhao *et al*., 2007; Brogna *et al*., 2017).

The spatial variations emphasize the need for water monitoring that provides a representative and reliable estimate (Muangthong and Shrestha, 2015). Recently, several approaches have been used in for water quality analysis. Multivariate techniques such as PCA and GA could be used for analyzing big water quality databases without losing important information (Helena *et al*., 2000; Singh *et al*., 2005; Wang *et al*., 2013).

Furthermore, interpolation methods have been employed to map the spatial distribution of soil properties (Villatoro et al., 2008; Bhunia et al., 2011), heavy metals (Xie *et al*., 2011; Yan *et al*., 2015), population characteristics (Navarrete, 2012), precipitation (Wang *et al*., 2014; Núñez *et al*., 2014), among others. Data interpolation offers the advantage of projecting maps or continuous surfaces from discrete data (Johnston *et al*., 2001). Therefore, spatial interpolation techniques are essential to create a continuous (or predictable) surface from values of sampled points (Wang *et al*., 2014).

Interpolation is an efficient method to study the spatial allocation of elements, their inconsistency, reduce the error variance and execution costs (Behera and Shukla, 2015). The interpolation methods are useful for identifying contamination sources, assessing pollution trends and risks (Markus and McBratney, 2001; Rawlins *et al*., 2006).

REFERENCES

AlSuhaimi, A.O., K.M. AlMohaimidi, and K.A. Momani. 2019. Preliminary assessment for physicochemical quality parameters of groundwater in Oqdus Area, Saudi Arabia. Journal of the Saudi Society of Agricultural Sciences.18:22-31.

Batayneh, A., and T. Zumlot. 2012. Multivariate statistical approach to geochemical methods in water quality factor identification; application to the shallow aquifer system of the Yarmouk basin of north Jordan. Res. J. Environ. Earth Sci. 4:756–768.

Berberoğlu, S., A. Akın, and Clarke, K.C. 2016. Cellular automata modeling approaches to forecast urban growth for adana, Turkey: A comparative approach. Landscape and Urban Planning 153:11-27.

Bhunia, G.S., P.K. Shit, and R. Maiti. 2016. Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC). Journal of the Saudi Society of Agricultural Sciences, 17:114-116.

Borrelli P., D.A. Robinson, L.R. Fleischer, E. Lugato, C. Ballabio, C. Alewell, K. Meusburger, S. Modungo, B. Schütt, V. Ferro, V. Bagarello, K. Oost, L. Montanarella, and P. Panagos. 2017. An assessment of the global impact of 21st century land use change on soil erosion. Nature communications 8:2013.

Boyacioglu, H. 2006. Surface water quality assessment using factor analysis. Water. S.A. 32:389–393.

Brogna, D., A. Michez, S. Jacobs, M. Dufrêne, C. Vincke, and N. Dendoncker. 2017. Linking forest cover to water quality: A multivariate analysis of large monitoring datasets. Water, 9:176.

Castillo-Rodríguez, M., J. López-Blanco, and E. Muñoz-Salinas. 2010. A geomorphologic GIS-multivariate analysis approach to delineate environmental units, a case study of La Malinche volcano (central México). Appl. Geogr. 30:629–638.

Chander, G.B., L. Markham and D.L. Helder. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sens. Environ. 113:893–903.

Chavez, P.S., Jr. 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sens. Environ. 24: 459–479.

Chen, T., R.Q. Niu, P.X. Li, L.P. Zhang, and B. Du. 2011. Regional soil erosion risk mapping using RUSLE, GIS, and remote sensing: A case study in Miyun Watershed, North China. Environ. Earth Sci. 63:533–541.

Cohen, W.B. and S.N. Goward. 2004. Landsat's role in ecological applications of remote sensing. BioScience, 54:535-545.

Craine, J.M. 2013. Long-term climate sensitivity of grazer performance: A cross-site study. PLoS ONE, 8:e67065.

DeFries R., A. Hansen, B.L. Turner, R. Reid, and J. Liu 2007. Land use change around protected areas: management to balance human needs and ecological function. Ecol Appl. 17: 1031-1038.

Eastman J.R., L.A. Solorzano, and M.E. Van Fossen. 2005. Transition potential modeling for land-cover change In: GIS, spatial analysis, and modeling, eds. Maguire D Batty JM, Goodchild MF. ESRI Press, California.

Estrada-Castillón, E., L. Scott-Morales, J.A. Villarreal-Quintanilla, E. Jurado-Ybarra, M. Cotera-Correa, C. Cantú-Ayala, and J. García-Pérez. 2010. Clasificación de los pastizales halófilos del noreste de México asociados con perrito de las praderas (*Cynomys mexicanus*): Diversidad y endemismo de especies. Revista mexicana de biodiversidad, 81:401–416.

Ferreira M.E., Jr. L.G. Ferreira, F. Miziara, and B.S. Soares-Filho. 2013. Modeling landscape dynamics in the central Brazilian savanna biome: future scenarios and perspectives for conservation. Journal of Land Use Science, 8: 403-421.

Gajbhiye, S., S.K. Sharma, and M.K. Awasthi. 2015. Application of Principal Components Analysis for Interpretation and Grouping of Water Quality Parameters. Int. J. Hybrid Inf. Technol. 8:89–96.

Galford G.L., B.S. Soares-Filho, L.J. Sonter, and N. Laporte. 2015. Will passive protection save Congo forests?. PloS one, 10: e0128473.

Gharun M., M. Possell, T.L. Bell, and M.A. Adams. 2017. Optimisation of fuel reduction burning regimes for carbon, water and vegetation outcomes. Journal of environmental management 203:157-170.

Goetz, S. J., S. D. Prince, Goward, S. N.; M. M. Thawley and J. 1999. Small. Satellite remote sensing of primary production: An improved production efficiency modeling approach. Ecol. Model. 122: 239–255.

Hansen, M.C. and Loveland, T.R. 2012. A review of large are monitoring of land cover change using Landsat data. Remote Sens. Environ. 122: 66–74.

He, C., Q. Zhang, Y. Li, X. Li and P. Shi. 2005. Zoning grassland protection area using remote sensing and cellular automata modeling-A case study in Xilingol steppe grassland in northern China. J. Arid Environ. 63:814–826.

Helena, B., R. Pardo, M. Vega, E. Barrado, J.M. Fernández, and L. Fernández. 2000. Temporal evolution of groundwater composition in an alluvial aquifer

(Pisuerga river, Spain) by principal component analysis. Water Res. 34:807–816.

Janzen, T.D., L.A. Fredeen, and D.R. Wheate. 2006. Radiometric correction techniques and accuracy assessment for Landsat TM data in remote forested regions. Can. J. Remote Sens., 32:330-340.

Johnston, K., J.M. Ver Hoef, K. Krivoruchko, and N. Lucas. 2001. Using ArcGIS geostatistical analyst. Redlands ESRI: CA. USA.

Kaufman, Y.J., A.E. Wald, L.A. Remer, B.C. Gao, R.R. Li, and L. Flynn. 1997. The MODIS 2.1-um channel-correlation with visible reflectance for use in remote sensing of aerosol. IEEE Trans. Geosci. Remote. 35:1286-1298.

Khan, S., Q. Cao, Y.M. Zheng, Y.Z. Huang, and Y.G. Zhu. 2008. Health risks of heavy metals in contaminated soils and food crops irrigated with wastewater in Beijing, China. Environ. Pollut. 152:686-692.

Körner, I., L. Oreopoulos, G. Feingold, L.A. Remer, and O. Altaratz. 2008. How small is a small cloud? Atmos. Chem. Phys. 8, 3855-3864.

Li, J.Y., X.C. Yang, Y.X. Jin, Z. Yang, W.G. Huang, L.N. Zhao, T. Gao, H.D. Yu, H.L. Ma and Z.H. Qin. 2003. Monitoring and analysis of grassland desertification dynamics using Landsat images in Ningxia, China. Remote Sens. Environ. 138, 19-26.

Lyapustin, A. D.L. Williams, B. Markham, J. Irons, B. Holben, and Y. Wang. 2004. A method for unbiased high-resolution aerosol retrieval from Landsat. J. Atmos. Sci. 61:1233-1244.

Manjarrez–Domínguez, C., A. Pinedo–Alvarez, C. Pinedo–Alvarez, F. Villarreal–Guerrero and L. Cortes-Palacios. 2015. Vegetation landscape analysis due to land use changes on arid lands. Pol. J. Ecol. 63:272–279.

Markus, J., and A.B. McBratney. 2001. A review of the contamination of soil with lead II. Spatial distribution and risk assessment of soil lead. Environ. Int. 27:399–411.

Miranda, J., E. Andrade, A. López-Suárez, R. Ledesma, T.Cahill, and P. Wakabayashi. 1996. A receptor model for atmospheric aerosols from a southwestern site in Mexico City. Atmos. Environ. 30:3471–3479.

Morris, B.L., A.R. Lawrence, P.J.C. Chilton, B. Adams, R.C. Calow, and B.A. Klinck. 2003. Groundwater and its susceptibility to degradation: a global assessment of the problem and options for management Vol. 3. United Nations Environment Programme: Nairobi, Kenya, 2003. 118. ISBN: 92-807-2297-2.

Muangthong, S., and S. Shrestha. 2015. Assessment of surface water quality using multivariate statistical techniques: case study of the Nampong River and Songkhram River, Thailand. Environ. Model. Assess. 187:548.

Muangthong, S., and S. Shrestha. 2015. Assessment of surface water quality using multivariate statistical techniques: case study of the Nampong River and Songkhram River, Thailand. Environ. Model. Assess. 187:548.

Navarrete Álvarez, M. 2012. Modelos geoestadísticos del precio de la vivienda: aproximación al conocimiento intraurbano de la ciudad de Madrid. Doctoral dissertation. Madrid, España.

Núñez López, D., E.J. Treviño Garza, V.M. Reyes Gómez, C.A. Muñoz Robles, O.A. Aguirre Calderón, and J. Jiménez Pérez. 2014. Uso de modelos de regresión para interpolar espacialmente la precipitación media mensual en la cuenca del río Conchos. Rev. Mex. Cien. Agrí. 5:201-213.

Oketola, A.A., S.M. Adekolurejo, and O. Osibanjo. 2013. Water Quality Assessment of River Ogun Using Multivariate Statistical Techniques. J. Environ. Prot. 4:466–479.

Pérez-Vega, A., J.F. Mas, and A. Ligmann-Zielinska. 2012. Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. Environ. Modell. Sofw. 29: 11-23.

Pijanowski B.C., D.G. Brown, B.A. Shellito, and G.A. Manik. 2002. Using neural networks and GIS to forecast land use changes: a land Transformation Model. Comput. Environ. Urban, 26:553–575.

Prieto-Amparán, J.A., E. Santellano-Estrada, F. Villarreal-Guerrero, M. Martinez-Salvador, A. Pinedo-Alvarez, G. Vázquez-Quintero, M.C. Valles-Aragón, and C. Manjarrez-Domínguez. 2019. Spatial Analysis of Temperate Forest Structure: A Geostatistical Approach to Natural Forest Potential. Forests 10:168.

Prieto-Amparán, J.A., B.A. Rocha-Gutiérrez, M.D.L. Ballinas-Casarrubias, M.C. Valles-Aragón, M.D.R. Peralta-Pérez, A. Pinedo-Alvarez, 2018. Multivariate and Spatial Analysis of Physicochemical Parameters in an Irrigation District, Chihuahua, Mexico. Water, 10:1037.

Qadir, A., R.N. Malik, and S.Z. Husain. 2008. Spatio-temporal variations in water quality of Nullah Aik-tributary of the river Chenab, Pakistan. Environ. Monit. Assess. 140:43-59.

Rahaman, S.A., S.A. Ajeez, S. Aruchamy, and R. Jegankumar. 2015. Prioritization of Sub Watershed Based on Morphometric Characteristics Using Fuzzy Analytical Hierarchy Process and Geographical Information

System—A Study of Kallar Watershed, Tamil Nadu. Aquat. Procedia 4:1322–1330.

Rawlins, B.G., R.M. Lark, R. Webster, and K.E. O'Donnell. 2006. The use of soil survey data to determine the magnitude and extent of historic metal deposition related to atmospheric smelter emissions across Humberside, UK. Environ. Pollut. 143:416–426.

Richter, R., D. Schlapfer, and A. Muller. 2006. An automatic atmospheric correction algorithm for visible/NIR imagery. Int. J. Remote Sens. 27:2077-2085.

Roy, D.P., H.K. Zhang, J. Ju, J.L. Gomez-Dans, P.E. Lewis, C.B. Schaaf, Sun, Li Q., J.; H. Hung and V. Kovalskyy. 2016. A general method to normalized Landsat reflectance data to nadir BRDF adjusted reflectance. Remote Sens. Environ. 176:255-271.

Saha, S., and S. Mukhopadhyay. 2012. Determination of Significant Geomorphic Parameters of Kunur River Basin through Principal Component Analysis. J. Eng. Comput. Appl. Sci. 4:232–243.

Schumm, S.A. 1956. Evolution of drainage systems and slopes in badlands at perth amboy, new jersey. GSA Bull. 67:597.

Sharma, R.K., M. Agrawal, and F. Marshall. 2007, Heavy metal contamination of soil and vegetables in suburban areas of Varanasi, India. Ecotox. Environ. Safe. 66:258-266.

Sharma, S.K., G.S. Rajput, S. Tignath, and R.P. Pandey. 2010. Morphometric analysis of a watershed using GIS. J. Indian Water Res. Soc. 30:33–39.

Sharma, S.K., S. Gajbhiye, R.K. Nema, and S. Tignath. 2014. Assessing vulnerability to soil erosion of a watershed of tons River basin In Madhya Pradesh using Remote sensing and GIS. Int. J. Environ. Res. Dev. 4:153–164.

Sharma, S.K., S. Tignath, S. Gajbhiye, and R. Patil. 2013. Application of principal component analysis in grouping geomorphic parameters of Uttela watershed for hydrological modeling. Int. J. Remote Sens. Geosci. 2:63–70.

Shrimali, S., S. Aggarwal, and J. Samra. 2001. Prioritizing erosion-prone areas in hills using remote sensing and GIS—A case study of the Sukhna Lake catchment, Northern India. Int. J. Appl. Earth Obs. Geoinf. 3:54–60.

Singh, K.P., A. Malik, and S. Sinha. 2005. Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques: a case study. Anal. Chim. Acta. 538:355e37.

Sun G., G. Zhou, Z. Zhang, X. Wei, S.G. McNulty, and J.M. Vose. 2006. Potential water yield reduction due to forestation across China. Journal of Hydrology 328:548-558.

Tan, K.C., H. San, M. Lim, Z. MatJafri, K. Abdullah. 2012. A comparison of radiometric correction techniques in the evaluation of the relationship between LST and NDVI in Landsat imagery. Environ. Monit. Assess. 184:3813-3829.

Thakker, A.K., and S.P. Dhiman. 2007. Morphometric analysis and prioritization of miniwatersheds in Mohr watershed, Gujarat using Remote sensing and GIS techniques. J. Indian Soc. Remote Sens. 35:313–321.

Tilahun, S.A., C.D. Guzman, A.D. Zegeye, E.K. Ayana, A.S. Collick, B. Yitaferu, and T.S. Steenhuis. 2014. Spatial and temporal patterns of soil erosion in the semi-humid Ethiopian highlands: A case study of Debre Mawi watershed. In Nile River Basin, 1st ed.; Assefa, M., Melesse, W., Abtew Shimelis, G., Eds.; Springer: Cham, Switzerland. 149–163.

Torrens PM. 2006. Geosimulation and its application to urban growth modeling. Springer-Verlag: London, U. K.

Tritsch, I., P. Sist, I.D.S. Narvaes, L. Mazzei, L. Blanc, C. Bourgoin, G. Cornu, and V. Gond. 2016. Multiple Patterns of Forest Disturbance and Logging Shape Forest Landscapes in Paragominas, Brazil. Forests 7:315.

Varol, S., and A. Davraz. 2015. Evaluation of the groundwater quality with WQI (Water Quality Index) and multivariate analysis: a case study of the Tefenni plain (Burdur/Turkey). Environ. Earth. Sci. 73:1725-1744.

Vázquez-Quintero, G., R. Solís-Moreno, M. Pompa-García, F. Villarreal-Guerrero, C. Pinedo-Alvarez, and A. Pinedo-Alvarez. 2016. Detection and projection of forest changes by using the Markov Chain Model and Cellular Automata. Sustainability. 8:236.

Villatoro, M., C. Henríquez, and F. Sancho. 2008. Comparación de los interpoladores IDW y Kriging en la variación espacial de pH, Ca, CICE y P del suelo. Agron. Costarric. 32.

Viramontes-Olivas, O.A., V.M. Reyes-Gómez, L.F. Escoboza-García, J.A. Román-Calleros, A. Pérez-Márquez, C. Pinedo-Álvarez, P.B. Sánchez-Fernández, and N. Miranda Ojeda. 2008. Hidrología de los suelos de las cuencas media y baja del Río Conchos, Chihuahua, México. Rev. Latinoam. Recur. Nat. 4:31–45.

Wang, S., G.H. Huang, Q.G. Lin, Z. Li, H. Zhang, and Y.R. Fan. 2014. Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada. Int. J. Climatol. 34:3745-3751.

Wang, Y., P. Wang, Y. Bai, Z. Tian, J. Li, X. Shao, L.F. Mustavish, and B.L. Li. 2013. Assessment of surface water quality via multivariate statistical techniques: a case study of the Songhua River Harbin region, China. J Hydro-Environ. Res 7:30-40.

Weiskittel A.R., N.L. Crookston, P.J. and Radtke. 2011. Linking climate, gross primary productivity, and site index across forests of the western United States. Can J. Forest Res. 41:1710-1721.

Woodcock, C.E., S.A. Macomber, Pax-Lenney, M. W.C. Cohen. 2001. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. Remote Sens. Environ. 78:194-203.

Wulder, M.A., J.C. White, S.N. Goward, J.G. Masek, J.R. Irons, M. Herold, W.B. Cohen, T.R. Loveland and C.E. Woodcock. 2008. Landsat continuity: Issues and opportunities for land cover monitoring. Remote Sens. Environ. 112:955-969.

Xie, Y., T.B. Chen, M. Lei, J. Yang, Q.J. Guo, B. Song, and X.Y. Zhou. 2011. Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: accuracy and uncertainty analysis. Chemosphere, 82:468-476.

Yan, W., Q. Mahmood, D. Peng, W. Fu, T. Chen, Y. Wang, S. Li, J. Chen, and D. Liu. 2015. The spatial distribution pattern of heavy metals and risk assessment of moso bamboo forest soil around lead–zinc mine in Southeastern China. Soil. Till. Res. 153:120-130.

Yang, Y., J. Fang, Y. Pan and C. Ji. 2009. Aboveground biomass in Tibetan grasslands. J. Arid Environ. 73:91-95.

Zavoianu, I. 1985. Morphometry of Drainage Bassins, 2nd ed.; Editura Academiei: Bucharest, Romania; Elsevier Science Publishers: Amsterdam, The Netherlands. 9–235.

Zhao, P., L. Dengsheng, W. Guangxing, W. Chuping, H. Yujie, and Y. Shuquan 2016. Examining spectral reflectance saturation in Landsat imagery and corresponding solutions to improve forest aboveground biomass estimation. Remote Sens. 8:469.

STUDY I. ATMOSPHERIC AND RADIOMETRIC CORRECTION ALGORITHMS FOR THE MULTITEMPORAL ASSESSMENT OF GRASSLANDS PRODUCTIVITY

BY:

Prieto-Amparán J.A.[1], Federico Villarreal-Guerrero[1], Martin Martinez-Salvador[1], Carlos Manjarrez-Domínguez[2], Eduardo Santellano-Estrada[1], y Alfredo Pinedo-Alvarez[1],*

[1]Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada km 1., C.P. 31453. Chihuahua, Chih. México.
[2]Facultad de Ciencias Agrotecnológicas, Universidad Autónoma de Chihuahua, Campus I, s/n., C.P. 31350. Chihuahua, Chih., México.

*Corresponding author: Alfredo Pinedo Alvarez; Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada km 1., C.P. 31453. Chihuahua, Chih. México. apinedo@uach.mx, tel. +52-(614)-4-34-03-63, DOI: https://doi.org/10.3390/rs10020219.

ABSTRACT

ATMOSPHERIC AND RADIOMETRIC CORRECTION ALGORITHMS FOR THE MULTITEMPORAL ASSESSMENT OF GRASSLANDS PRODUCTIVITY

BY:

M.C. JESUS ALEJANDRO PRIETO AMPARAN

A key step in the processing of satellite imagery is the radiometric correction of images to account for reflectance that water vapor, atmospheric dust, and other atmospheric elements add to the images, causing imprecisions in variables of interest estimated at the earth's surface level. That issue is important when performing spatiotemporal analyses to determine ecosystems' productivity. In this study, three correction methods were applied to satellite images for the period 2010–2014. These methods were Atmospheric Correction for Flat Terrain 2 (ATCOR2), Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH), and Dark Object Substract 1 (DOS1). The images included 12 sub-scenes from the Landsat Thematic Mapper (TM), Landsat Enhanced Thematic Mapper Plus (ETM+) y Landsat Operational Land Imager (OLI) sensors. The images corresponded to three Permanent Monitoring Sites (PMS) of grasslands, 'Teseachi', 'Eden', and 'El Sitio', located in the state of Chihuahua, Mexico. After applying the corrections to the images, they were correlated with the biomass information and subsequently evaluated in terms of their accuracy for biomass estimation. For that, biomass production was measured during the study period at the three PMS to calibrate production models developed with simple and multiple linear regression (SLR and MLR) techniques. When the estimations were made with MLR, DOS1 obtained an R^2 of 0.97 ($P<0.05$) for 2012 and values greater than 0.70 ($P<0.05$) during 2013–2014. The rest of the algorithms did not show significant results and DOS1, which is the simplest algorithm, resulted in the best biomass estimator. Thus, in the multitemporal analysis of grassland based on spectral information, it is not necessary to apply complex correction procedures. The maps of biomass production, elaborated from images corrected with DOS1, can be used as a reference point for the assessment of the grassland condition, as well as to determine the grazing capacity and thus the potential animal production in such ecosystems.

Key words: Landsat; ATCOR2; DOS1; FLAASH: spatio temporal

RESUMEN

ATMOSPHERIC AND RADIOMETRIC CORRECTION ALGORITHMS FOR THE MULTITEMPORAL ASSESSMENT OF GRASSLANDS PRODUCTIVITY

POR:

M. C. JESÚS ALEJANDRO PRIETO AMPARÁN

Doctor en Philosophia de Recursos Naturales

Secretaría de Investigación y Posgrado

Facultad de Zootecnia y Ecología

Universidad Autónoma de Chihuahua

Presidente: Dr. Alfredo Pinedo Alvarez

Un paso clave en el procesamiento de imágenes satelitales es la corrección radiométrica, la cual ayuda a cuantificar la reflectancia del vapor de agua, el polvo atmosférico y otros elementos atmosféricos encontrados en las imágenes, causando imprecisiones en las variables de interés estimadas a nivel de la superficie terrestre. Esta cuestión es importante cuando se desea realizar análisis espacio-temporales para determinar la productividad de los ecosistemas. En este estudio, se aplicaron tres métodos de corrección a las imágenes de satélite para el período 2010-2014. Estos métodos fueron Atmospheric Correction for Flat Terrain 2 (ATCOR2), Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH), y Dark Object Substract 1 (DOS1). Las imágenes incluían 12 subescenas de los sensores Landsat Thematic Mapper (TM), Landsat Enhanced Thematic Mapper Plus (ETM+) y Landsat Operational Land Imager (OLI). Las imágenes corresponden a tres Sitios Permanentes de Monitoreo de pastizales (SPMP), ‘Teseachi’, ‘Eden’, y ‘El Sitio’, ubicados en el estado de

Chihuahua, México. Después de aplicar las correcciones a las imágenes, éstas fueron correlacionadas con la información de la biomasa y posteriormente evaluadas en términos de su precisión para la estimación de biomasa. Para ello, se midió la producción de biomasa durante el periodo de estudio en los tres PMS para desarrollar los modelos de producción desarrollados con técnicas de regresión lineal simple y múltiple (RLS y RLM). Cuando se hicieron las estimaciones con RLM, el modelo DOS1 obtuvo un R^2 de 0.97 ($P<0,05$) para 2012 y valores superiores a 0.70 ($P<0,05$) durante 2013-2014. El resto de los algoritmos no mostraron resultados significativos y DOS1, que es el algoritmo más simple, resultó en el mejor estimador de biomasa. Por lo tanto, en el análisis multitemporal de los pastizales basado en información espectral, no es necesario aplicar procedimientos de corrección complejos. Los mapas de producción de biomasa, elaborados a partir de imágenes corregidas con DOS1, pueden utilizarse como punto de referencia para la evaluación del estado de los pastizales, así como para determinar la capacidad de pastoreo y, por tanto, la producción animal potencial en dichos ecosistemas.

Palabras clave: Landsat; ATCOR2; DOS1; FLAASH; espacio-temporal.

INTRODUCTION

Grassland ecosystems play an important role in biodiversity conservation, ecosystem services provision, and the global carbon cycle (Yang *et al*., 2009). They also play a key role in biogeochemical cycles and the exchange of energy (Jin *et al*., 2009). However, grasslands have experienced a notorious degradation in recent years due to long droughts, climate variability, fragmentation and anthropic intervention. The application of improper management practices has been one of the main causes of such degradation (Craine, 2013; Manjarrez *et al*., 2015). Specifically, grasslands in northern Mexico are extensively used for cattle grazing and have also experienced intensive land-use change due to agricultural activities, which has resulted in changes of the species composition on such ecosystems (Estrada-Castillon *et al*., 2010; Manjarrez *et al*., 2016).

Accurate monitoring and evaluation of grasslands' conditions is critical for their conservation and, in some cases, their restoration (He *et al*., 2005). In the past decades, biomass inventory of grasslands has been driven by traditional methods of evaluation, which include extensive field sampling (Fan *et al*., 2007; Ruppert *et al*., 2014). Even though these methods are accurate, they are costly, as well as time-and labor-consuming, when large pieces of land have to be covered (Xie *et al*., 2009). With the aim of developing more effective monitoring methods, there have been numerous studies on indirect methods to estimate the biomass of grasslands using remote sensing information (Goetz *et al*., 1999; Li *et al*., 2003; Zhao *et al*., 2016). In this endeavor, optical sensors, radar, and Lidar systems have been used (Naesset *et al*., 2013). In general, all these studies have sought to find relationships between grassland structural variables and satellite image spectral data (Schino *et al*., 2003).

For instance, Marsett *et al*. (2006) used image-processing algorithms to quantify the total cover, height, and biomass of grasslands. Comparisons of remote sensor estimates with independent field measurements yielded values of R^2 = 0.80, 0.85, and 0.77 and Nash Sutcliffe coefficient values of 0.78, 0.70, and 0.77 for coverage, plant height, and biomass, respectively. Dusseux *et al*. (2015) correlated spectral data from the Satellite Pour l'Observation de la Terre (SPOT) sensor converted to indices, such as the leaf area index, the Normalized Difference Vegetation Index, and the vegetation cover fraction, with data of biomass measured in the field. The coefficients of determination found were R^2 = 0.68, 0.30, and 0.50. These studies demonstrated the good relationships between the spectral information from the remote sensors and the biomass inventories in the field. Rodríguez-Maturino *et al*. (2017) correlated 3-year data from Landsat TM5 as well as field measurements of coverage of grass canopy and grass height, obtaining values of R^2 greater than 0.70.

The Landsat satellite has provided data since 1972 (Cohen and Goward, 2004) with an extensive global coverage. This is an important resource for monitoring global environmental change (Woodcock *et al*., 2001; Hansen and Loveland 2012; Wulder *et al*., 2008). However, the problem of saturation of data in the images is considered an important factor affecting the results when estimating the biomass of vegetation (Chander *et al*., 2009; Tan *et al*., 2012; Roy *et al*., 2016). This phenomenon of saturation is caused by atmospheric dispersion and absorption, which has substantial variations with time, space, and wavelength. In addition, the surface reflectance is strongly affected by the elevation of the terrain, a more evident effect in mountainous environments (Körner *et al*., 2008). All of these factors determine the importance of atmospheric correction to attenuate the effects of noise occurring in the capture of satellite images (Zelazowski and Sayer, 2011; Pons *et al*., 2014), especially when multitemporal studies are carried out (Pinedo-Alvarez *et al*., 2007; Prieto-Amparán *et al*., 2016; Vazquez-Quintero *et al*., 2016).

Different atmospheric and radiometric correction methods have been developed to transform the original values of the images into values of reflectance (Chavez, 1988; Kaufman *et al*., 1997; Janzen *et al*., 2006). Several of these techniques have been developed to estimate the amount of atmospheric bias from thin-surface terrestrial images by assessing the molecular and aerosol dispersion (Lyapustin *et al*., 2004; Richer *et al*., 2006). For instance, the algorithm called Fast Line of sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) serves to derive surface and atmospheric reflectance properties using a MODTRAN accuracy model developed by Spectral Science Inc., Burlington, MA. U.S. (Kruse *et al*., 2004), which explains the effects of adjacency associated with the dispersion of the atmosphere. The method of Atmospheric Correction for Flat Terrain 2 (ATCOR2) removes the brightness of the image, as well as the possible effects of fog or clouds, to obtain the values of the terrestrial surface (Neubert y Meinel, 2005). The method of Dark Object Substract 1 (DOS1) is based on the properties of the image and is the algorithm most widely used for the detection of land-use changes (Paolini *et al*., 2006; Cui *et al*., 2014).

There have been some studies on the comparison of different approaches to correct the atmospheric effects. In this regard, El Hajj *et al*. (2008) compared relative radiometric normalization and a 6S algorithm employing SPOT5 data. Chang *et al*. (2008) evaluated the correction methods TOA, GDOS, and AC performed on Landsat images. Calliceco and Dell'Acqua (2011) compared the algorithms 6S and MODTRAN. Agrawal *et al*. (2011) compared the FLAASH and QUAC algorithms. Nazer *et al*. (2014) compared five atmospheric correction algorithms, 6S, FLAASH, ATCOR, DOS, and ELM, over sand, turf, grass, and water surfaces. López-Serrano *et al*. (2016) evaluated the performance of the COST, ATCOR2, FLAASH, 6S, and TOA algorithms for the estimation of forest above-ground biomass. Martins *et al*. (2017) compared the 6S, ACOLITE, and Sen2Cor methods applied to the new platform Sentinel 2-MSI. However, these studies lack the multitemporal component and only a few studies (Vicente-Serrano *et al*., 2008; Nguyen *et al*., 2015; Raab *et al*., 2015) have been developed to compare the methods of radiometric correction on different dates.

Based on the aforementioned, most of the studies are not multitemporal; they neither compare correction methods at different dates nor determine the grasslands biomass productivity in semi-arid regions. The objective of this study was to compare three correction methods based on their precision for the estimation of grassland biomass on the semi-arid ecosystems of Chihuahua, Mexico. Two atmospheric (ATCOR2, FLAASH) correction methods and one radiometric (DOS1) correction method were evaluated. The results may serve grassland owners for decision-making on animal load adjustments. Likewise, government institutions and non-governmental organizations working in areas such as forestry, agriculture, livestock research, and rural development could use this information for planning, decision–making, and the development of public policies.

MATERIAL AND METHODS

Location of the Study Area

The study included three Permanent Monitoring Sites (PMS), which belong to the National Livestock Oriented Land Monitoring System. The first PMS is called ‘Teseachi’, located at the central coordinates 28°53′35″ (N), 107°26′49″ (W); the second PMS, called ‘El Sitio’, is located at 27°35′17″ (N), 106°16′30″ (W); the third PMS, called ‘Eden’, is located at 27°06′50″ (N) and 105°26′46″ (W) (Figure 1). Each PMS is composed of nine monitoring stations, where biomass sampling was carried out once a year. The three PMS are located in the semi-arid region of Chihuahua, where the vegetation is dominated by grasses. Besides the grasslands, this region also houses chaparral vegetation and dunes (CONABIO, 2014). Grasslands provide habitat for wildlife, serve as reservoirs of carbon, and help mitigate global environmental change impacts (Jurado *et al*., 2013). In these ecosystems, it is possible to find mammals, such as Bison bison and Cynomys ludovicanus, which have multiple effects on grasslands, and both of them are considered key in maintaining grassland habitats (Samson *et al*., 2004). In addition, mammals such as Puma concolor can also be found, which is in a decreasing status based on the International Union for Conservation of Nature (IUCN) red list (IUCN, 2016).

Biomass Sampling

The data on biomass employed for this study comes from values registered on the field during the period 2010–2014. During the first year of sampling, the sites were plenty identified. The center site and the corner boundaries were marked with flags. These points got recorded with a global positioning system

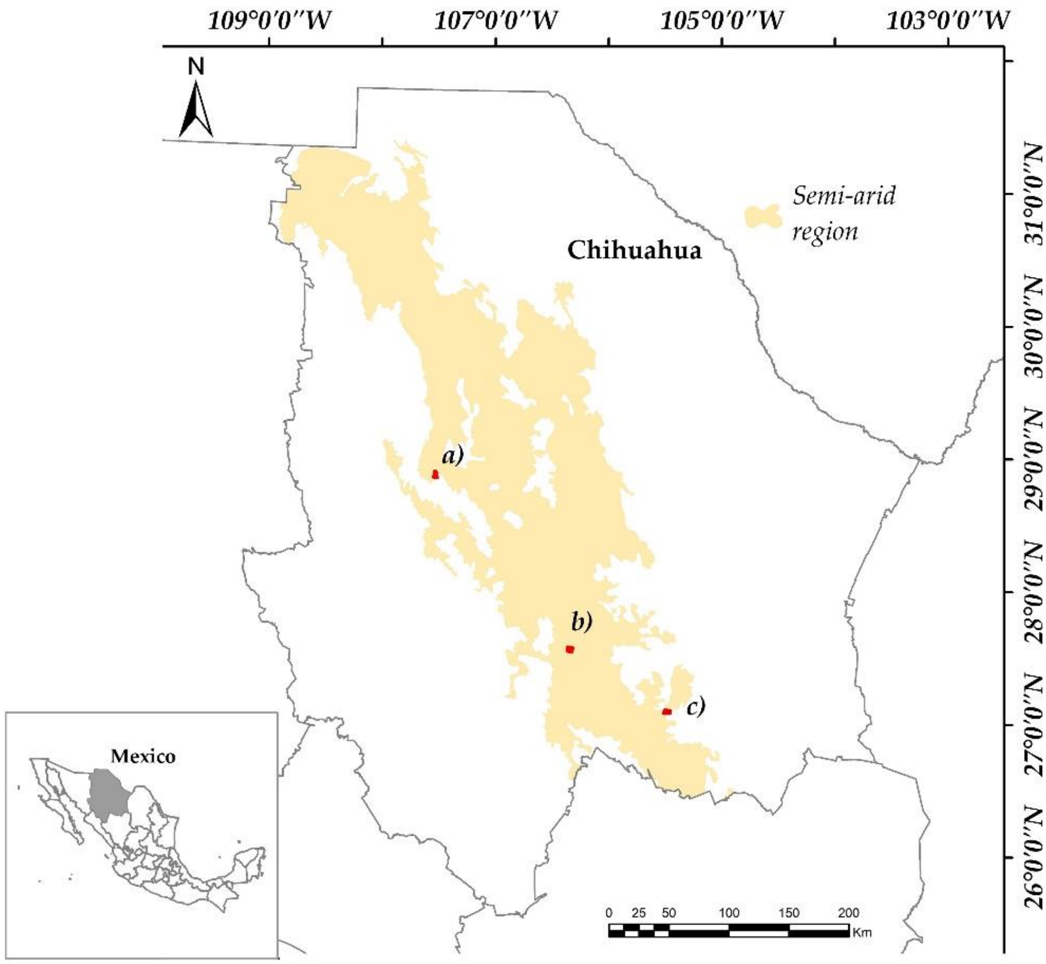

Figure 1. Location of the three permanent monitoring sites. The pictures depict representative biomass features of the sites. (a) Teseachi; (b) El Sitio; (c) Eden.

(GPS) to ensure that the biomass sampling was performed at the same locations during all of the years studied. The biomass was sampled during the month of October. During this period, the maximum peak of biomass is achieved and the vegetation offers a strong reflectance (Ni, 2004).

The sampling design for the biomass collection in each PMS was based on the shape and size of a pixel of a satellite image. The PMS is a quadrangular area of 225 ha (1.5 km × 1.5 km). In its interior, the PMS includes nine sampling stations of 1.0 ha each (100 m × 100 m). The centers of the sampling stations were located 200 m apart. The sampling stations consisted of a cross, which was marked from North to South and East to West. The cross served to locate its four corresponding quadrants. Within each quadrant, a circled area of 1.6 m of diameter was randomly marked. The circles were built with heat-resistant, orange color, plastic tubes of 0.02 m diameter, which are commonly employed for electrical wire protection. The biomass inside the quadrants was cut with sizers and placed in paper bags. The bags were tagged to clearly identify the sites where the biomass came from. The bags were then oven dried for 48 h at 70 °C and the database was built with the values of these dry weights ($Kg \times ha^{-1}$).

A detailed description of the field sampling design and data collection protocol can be found in (CGG-SAGARPA-COLPOS, 2009).

Satellite Data

A total of 12 images, including scenes taken by the Landsat Thematic Mapper (TM), Landsat Enhanced Thematic Mapper Plus (ETM+), and the Operational Land Imager (OLI), available at the United States Geological Survey (USGS, 2016), were used. The scenes had a spatial resolution of 30 m × 30 m and correspond to the period of 2010–2014 (Table 1). The sensors operate with several bands filtered from the electromagnetic spectrum. The Landsat TM and ETM+ are equipped with band 1 (0.45–0.52 μm), band 2 (0.52–0.60 μm), band 3 (0.63–0.69 μm), band 4 (0.76–0.90 μm), band 5 (1.55–1.75 μm), and band 7 (2.08–2.35 μm). Likewise, the OLI is equipped with band 2 (0.45–0.51 μm), band 3 (0.53–0.59 μm), band 4 (0.64–0.67 μm), band 5 (0.85–0.88 μm), band 6 (1.57–1.65 μm), and band 7 (2.11–2.29 μm) (Barsi *et al*., 1998).

Correction Methods

The correction methods (CMs) used to eliminate the noise in the satellite images were: ATCOR2, FLAASH, and DOS1. The ATCOR2 method removes the brightness of the image, as well as the possible effects of fog or clouds, to obtain the values of the terrestrial surface (Neubert y Meinel, 2005). It also uses predetermined sensor calibration values as well as solar angles to obtain reflectance values (Janzen *et al*., 2006). This method is based on the MODerate resolution atmospheric TRANsmission (MODTRAN) radioactive transfer model (Berk *et al*., 1998). The main characteristics of ATCOR2 are: a pre-classification of the scene (soil, water, fog, and clouds), recovery of atmospheric parameters (aerosol optical thickness, water vapor) and surface reflectance recovery (Marcello *et al*., 2016). The surface reflectance (ρ_{SUP}) is obtained by Equation (1).

$$\rho_{SUP} = \frac{1}{a_1}\left(\frac{d^2\,\pi\,L_{TOA}}{E_{TOA}\,cos\theta_i} - a_0\right) \tag{1}$$

where: *d* is the direct distance to the sun, *LTOA* is the spectral radiance of the satellite, *ETOA* is the solar spectral radiance on a surface perpendicular to the rays of the sun outside the atmosphere, and θ_i is the solar zenith angle. To obtain

Table 1. Characteristics of Landsat Thematic Mapper (TM), Landsat Enhanced Thematic Mapper Plus (ETM+), and Landsat Operational Land Imager (OLI) scenes used in the study.

PMS	Scene ID	Date of Image	Sensor	Path/Row
Eden	LT50310412010279EDC00	6 October 2010	TM	31/41
	LT50310412011298EDC00	25 October 2011		
	LC80310412013319LGN00	15 November 2013	OLI	
	LC80310412014274LGN00	1 October 2014		
El Sitio	LT50320412010302EDC00	29 October 2010	TM	32/41
	LT50320412011305EDC00	1 November 20111	ETM+	
	LE70320412012028EDC00	28 October 2012		
	LC80320412013278LGN00	5 October 2013	OLI	
	LC80320412014281LGN00	8 October 2014		
Teseachi	LT50330402010309EDC00	5 November 2010	TM	33/40
	LT50330402011296EDC00	23 October 2011	ETM+	
	LE70330402012339EDC00	4 December 2012		
	LC80330402013285LGN00	12 October 2013	OLI	
	LC80330402014288LGN00	15 October 2014		

PMS=Permanent monitoring sites.

the coefficients a_0 and a_1, the standard atmospheric parameters (aerosol type, visibility or optical thickness, and water vapor column) are required. Such parameters are available in the Software ERDAS© (v.2014). For this study, the solar model used was rural with a scene visibility of 40 km. In addition, the parameter of *tropical_rural* was chosen based on the dates the scenes were taken.

The mean reflectance of the studied area $\bar{\rho}_{SUP_i}$ is calculated to correct for the adjacency effect. Therefore, Equation (2) describes the relation to obtain the reflectance of the free surface of the adjacency effect.

$$\rho'_{SUP} = \rho_{SUP} + \left(\int_{\lambda}^{\lambda_2} \frac{\tau_{0dif}}{\tau_{0dir}} R d\lambda\right)\left(\rho_{SUP} - \sum_{i=1}^{n_R} \bar{\rho}_{SUP_i} w_i\right) \quad (2)$$

where: τ_{0dif} and τ_{0dir} are the diffuse and direct transmittance, respectively, *R* is the sensor-specific spectral response curve, and w_i defines the weighting coefficients as a function of the distance-dependence. The atmospheric correction was carried out with the ATCOR2 module of the software ERDAS© (v.2014).

The FLAASH algorithm is also based on the MODTRAN radiative transfer model (Alder-Golden *et al*., 1998; Anderson *et al*., 2002). It is designed to eliminate the atmospheric effects caused by the molecular dispersion of particles in the atmosphere. It is determined by Equation (3).

$$L_{TOA} = \left(\frac{A\rho_{SUP}}{1-\rho_e S}\right)\left(\frac{B\rho_e}{1-\rho_e S}\right) + L_o \quad (3)$$

where: L_{TOA} is the spectral radiance reached by the satellite, ρ_{SUP} is the reflectance of the pixel surface, ρ_e is the reflectance of the average surface of the pixel of the surrounding region, *S* is the spherical albedo of the atmosphere, L_o is the radiance backscattered by the atmosphere, and *A* as well as *B* are coefficients that depend on the atmosphere and geometric conditions.

The first term of the equation corresponds to the reflectance of the surface that travels directly into the sensor while the second term corresponds to the luminosity of the surface that is dispersed by the environment. The distinction between ρ_{SUP} and ρ_e explains the "adjacency effect" (spatial blending of radiation between nearby pixels) caused by the atmospheric dispersion. The values of *A*, *B*, *S*, and L_o can be determined empirically from the MODTRAN4 standards. The vision and the solar angles of the measurement and the nominal values for the surface elevation, aerosol shape, and visible range of the scene must be specified (Marcello *et al*., 2016). For this study, the standard model used was tropical, which is recommended for locations with Latitudes around 30° when the scenes are taken in September–October. The terrain elevation values used were 1818.13, 1450.5, and 2114.5 meters above the sea level for El Sitio, Eden, and Teseachi, respectively. In addition, the aerosol type was chosen as rural with a scene visibility of 40 km, corresponding to zones with clean weather conditions, as this is the case for the three sites analyzed in this study. This correction method was carried out with the FLAASH module of the software ENVI© (v.5.1).

The DOS1 method is based on the properties of the image. This correction method is the most widely used for the detection of land-use changes. Elements such as water, forests, and shadows are considered as dark objects when their values of reflectance are close to zero. Dark objects are detected automatically when the pixel reflectance value is less than or equal to 1.0 %. The assumption is that some pixels within the image receive 0 % of the solar radiation (100 % of shade), mainly due to the effect of topography, and the value of radiances corresponding to these pixels registered by the satellite correspond to atmospheric dispersion (Chavez, 1988). If a dark object is found in the image, the minimum reflectance value in the histogram is assigned to such an object. From this minimum, it is possible to correct the entire scene by the effects of the atmospheric dispersion (Paolini *et al*., 2006; Cui *et al*., 2014). To obtain the surface reflectance, Equation (4) is used.

$$\rho_{SUP} = \frac{d^2 \, \pi \, (L_{TOA} - L_0)}{E_{TOA} \, cos\theta_i} \tag{4}$$

where: *d* is the direct distance to the sun, L_{TOA} is the spectral radiance to the satellite, L_0 is the backscatter glow through the atmosphere, E_{TOA} is the solar spectral radiance on a surface perpendicular to the sun's rays outside the atmosphere, and θ_i is the solar zenith angle. The radiometric correction was carried out with the Semi-Automatic Classification plugin developed by Congedo (Congedo, 2013) and included in the software QGis (v.2.18).

Accuracy of the Correction Methods

A visual analysis of false color compositions was performed to compare the correction methods (CMs). Statistical analyses included an ANOVA and a means comparison of the spectral signatures per band, carried out in the software SAS© (v.9.1.3). In addition, a simple linear regression (SLR) analysis was carried out between the values of each CMs, separately per band, and the biomass. Moreover, a multivariate principal component analysis (PCA) was performed. The components that explained at least 80 % of the total variability of the reflectance values of each CMs were selected. The accuracy of the CMs was determined by

SLR when only one main component was selected and by multiple linear regression (MLR) in the cases when two or more components were selected. In any case, the coefficient of determination (R^2) and the root mean square error (RMSE) were quantified to find the CMs that best fitted the data of each year and each sampling site analyzed. Likewise, the P-value served to determine the reliability of the CMs to estimate the biomass. A flow diagram explaining the methodology followed in this study is represented in Figure 2.

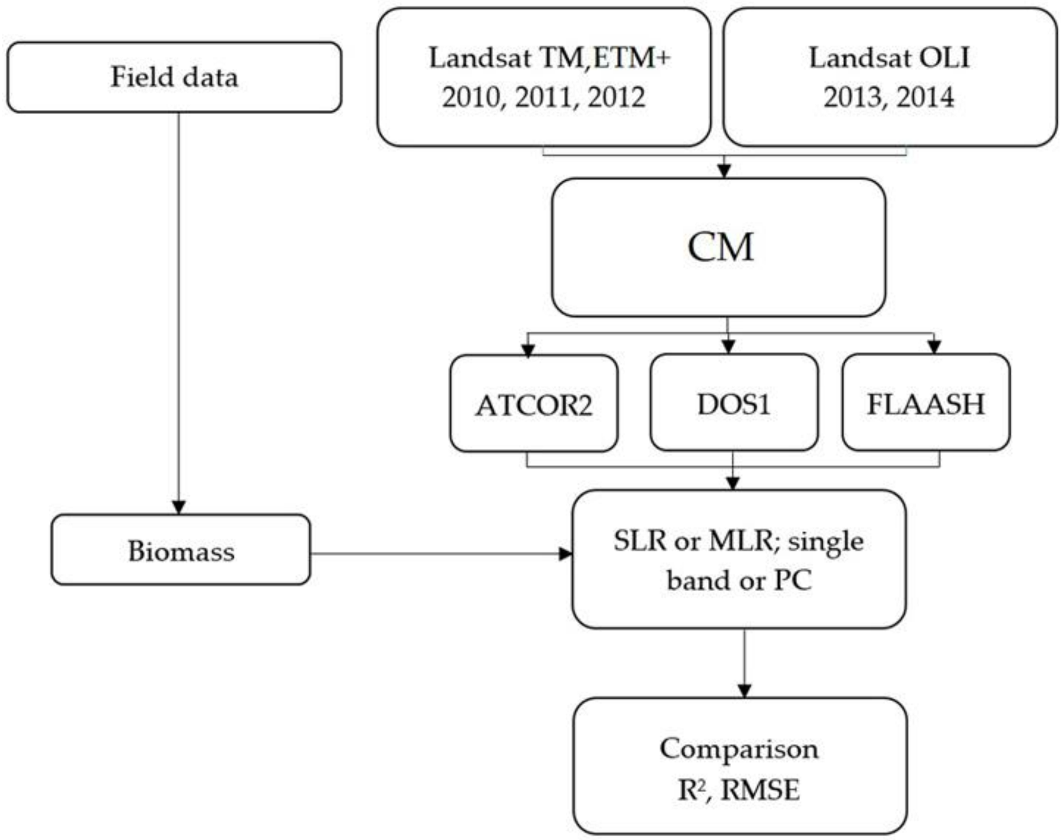

Figure 2. Flow diagram representing the procedure followed for the analysis. CM: correction method; SLR: simple linear regression; MLR: multiple linear regression; PC: principal component; RMSE: root mean square error.

RESULTS AND DISCUSION

Composition of Radiometrically Corrected Images

In a first analysis, the radiometrically corrected false-color images for Eden (years 2010, 2011, 2013, and 2014), as well as El Sitio and Teseachi (years 2010, 2011, 2012, 2013, and 2014), as is shown in Figures 3–5, were visually compared. The correction methods were configured using the parameter of rural zone, which is a pre-calibrated value for zones not affected by urban zones or industrial activities (Marcello *et al*., 2016).

Comparative Analysis of the Correction Methods

The reflectance means from the three PMS, obtained after applying the CMs for the period 2010–2014, were compared (Figures 6–8). The differences among the CMs in general varied. The spectral signature of the grassland showed low reflectance values for the bands blue, green, and red during the studied period. Conversely, high values of reflectance were obtained by the bands corresponding to the Near Infra-Red (NIR) and Shortwave Infra-Red (SWIR) regions for the three PMS. That was possibly due to a strong chlorophyll absorption.

The DOS1 and ATCOR2 methods presented higher values of surface reflectance than the FLAASH method in the visible region for most of the years. This situation may have been induced due to the configuration of DOS1 and ATCOR2, which ignore the effects of atmospheric dispersion on the spectral signatures. In particular, DOS1 does not have the capability to simulate the atmospheric absorption and produces a decrement of surface reflectance (Lu *et al*., 2002).

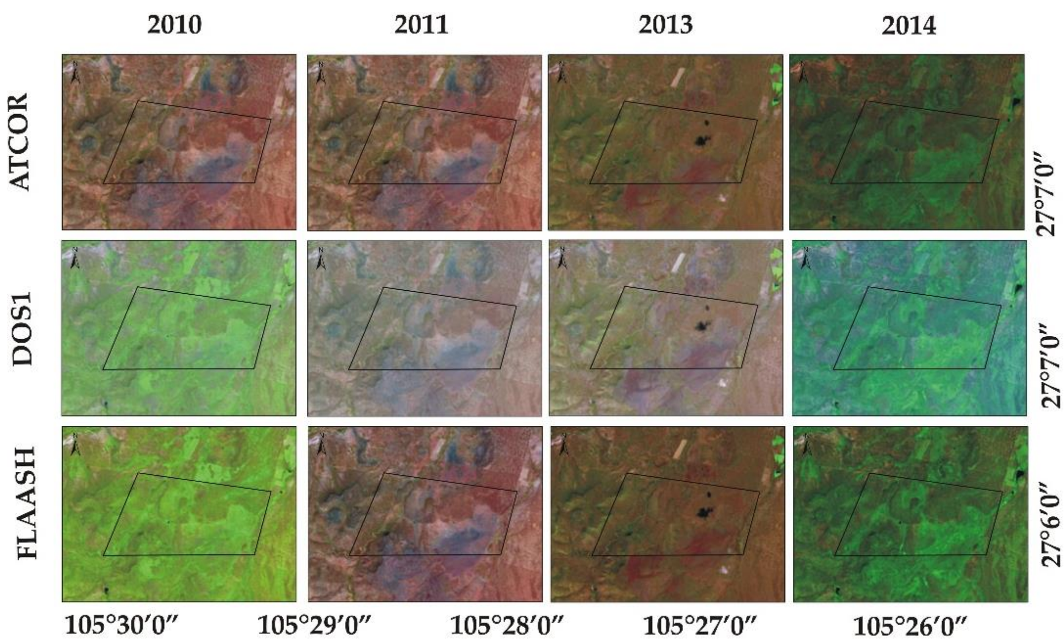

Figure 3. False-color images from the different radiometric correction algorithms. Site: Eden.

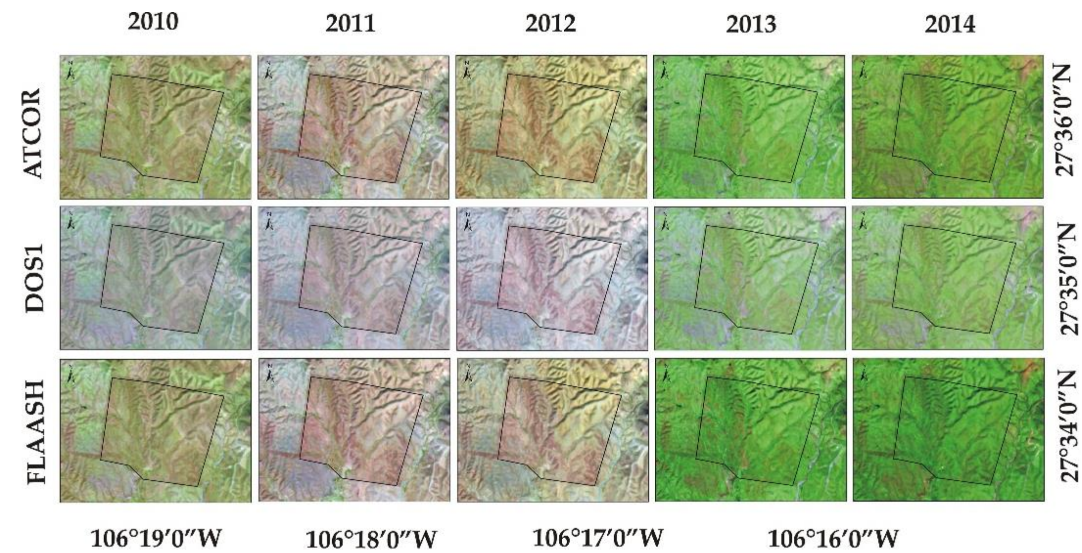

Figure 4. False-color images from the different radiometric correction algorithms. Site: El Sitio.

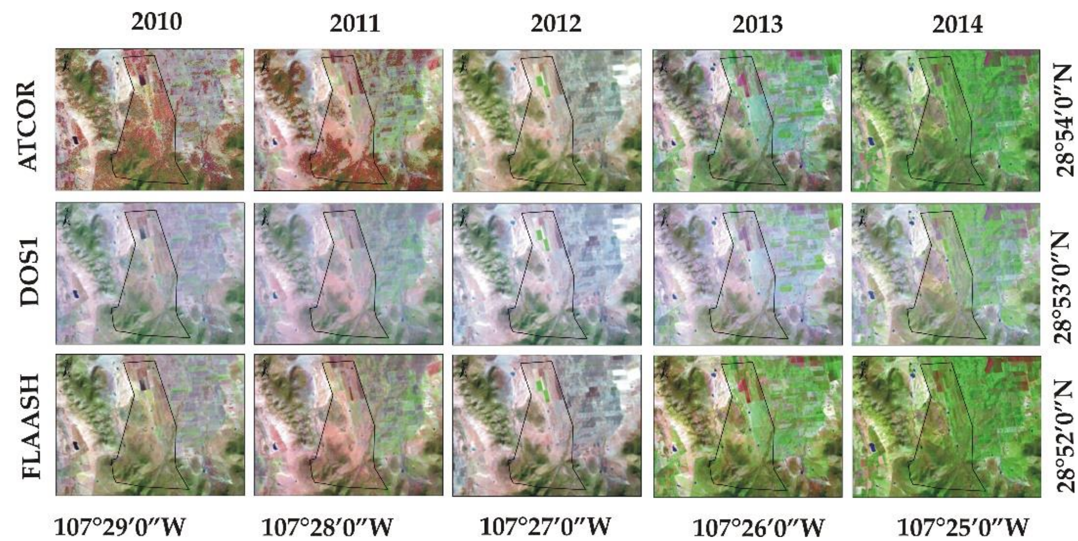

Figure 5. False-color images from the different radiometric correction algorithms. Site: Teseachi.

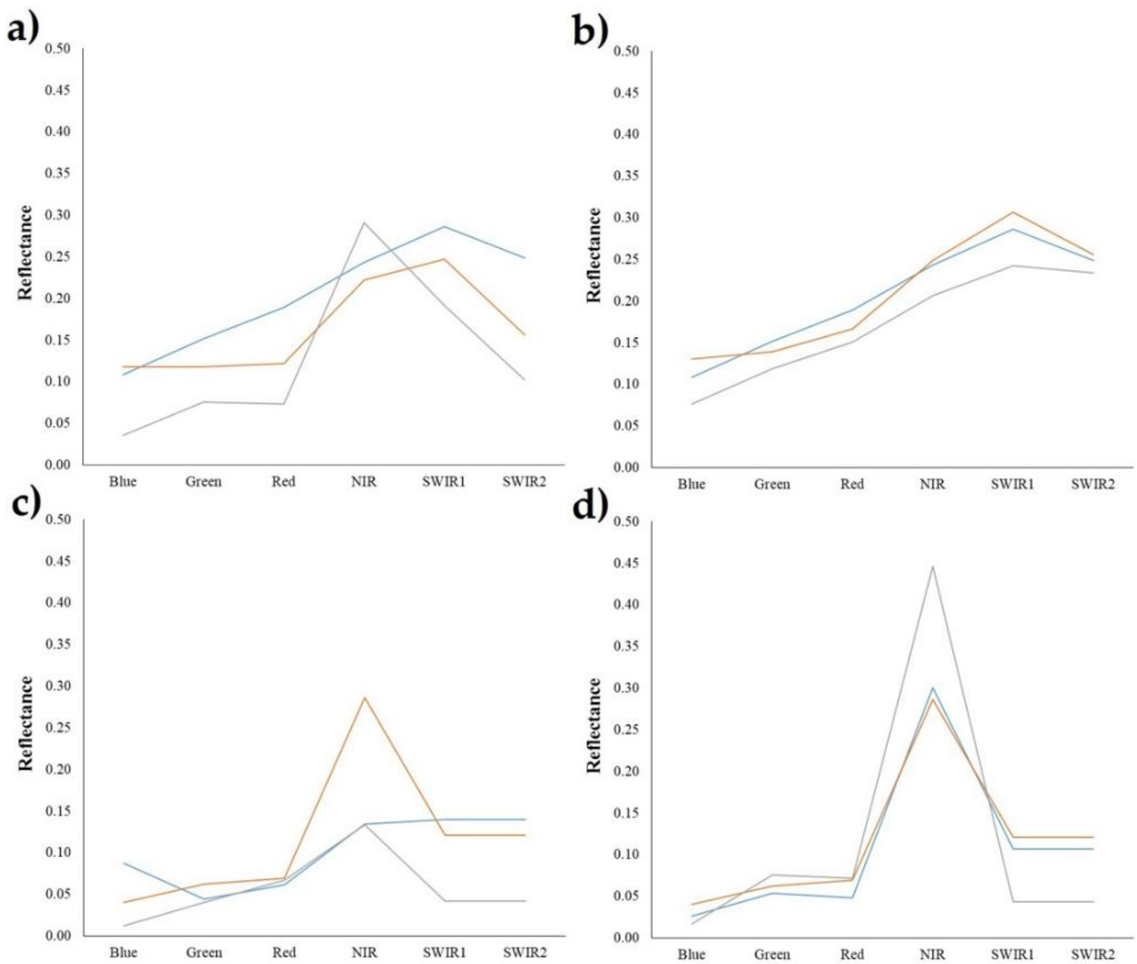

Figure 6. Spectral response of each correction method in Eden for 2010 a); 2011 (b); 2012 (c); 2013 (d); 2014 (e). ATCOR2 (—), DOS1 (—) and FLAASH (—).

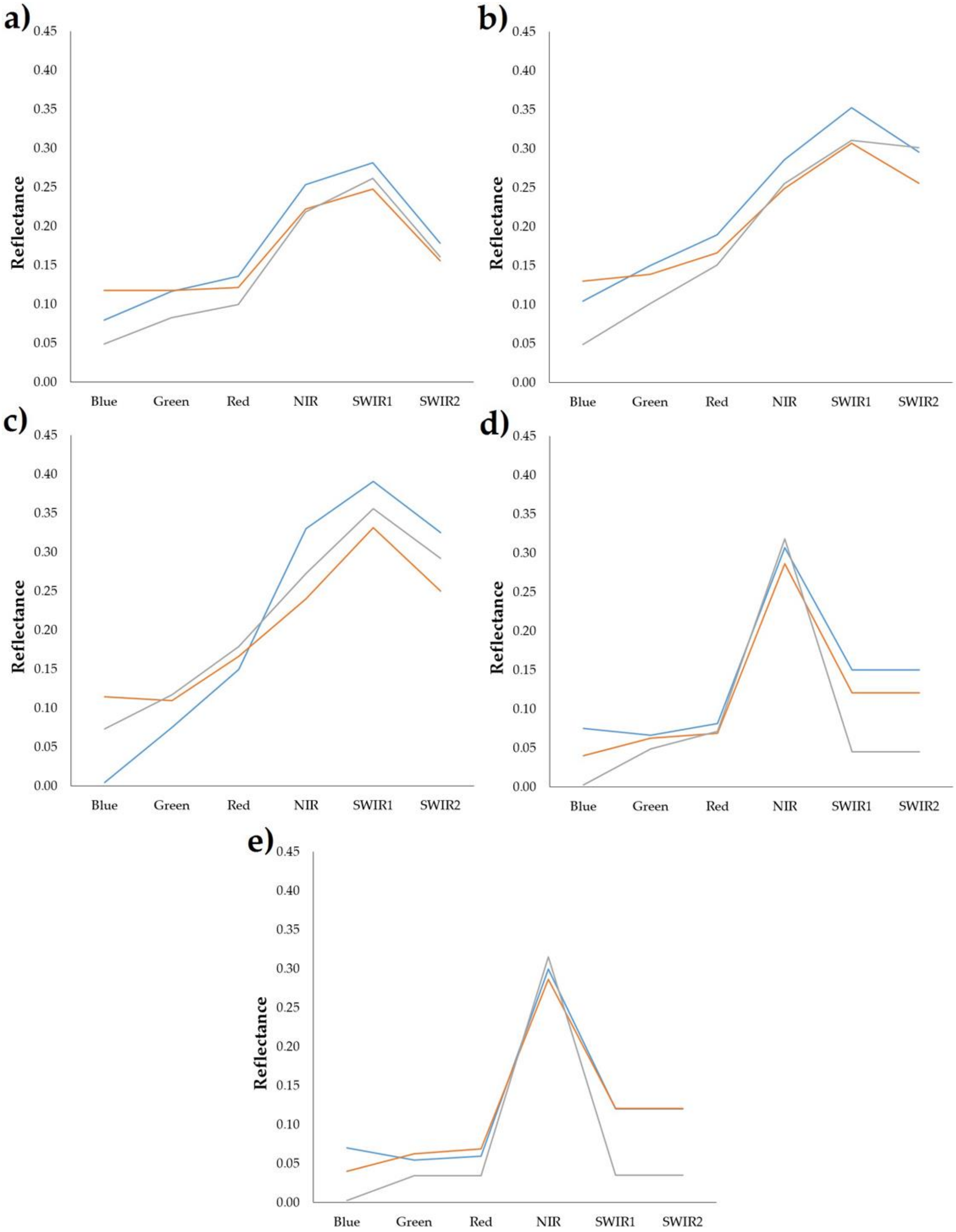

Figure 7. Spectral response of each correction method in El Sitio for 2010 a); 2011 (b); 2012 (c); 2013 (d); 2014 (e). ATCOR2 (—), DOS1 (—) and FLAASH (—).

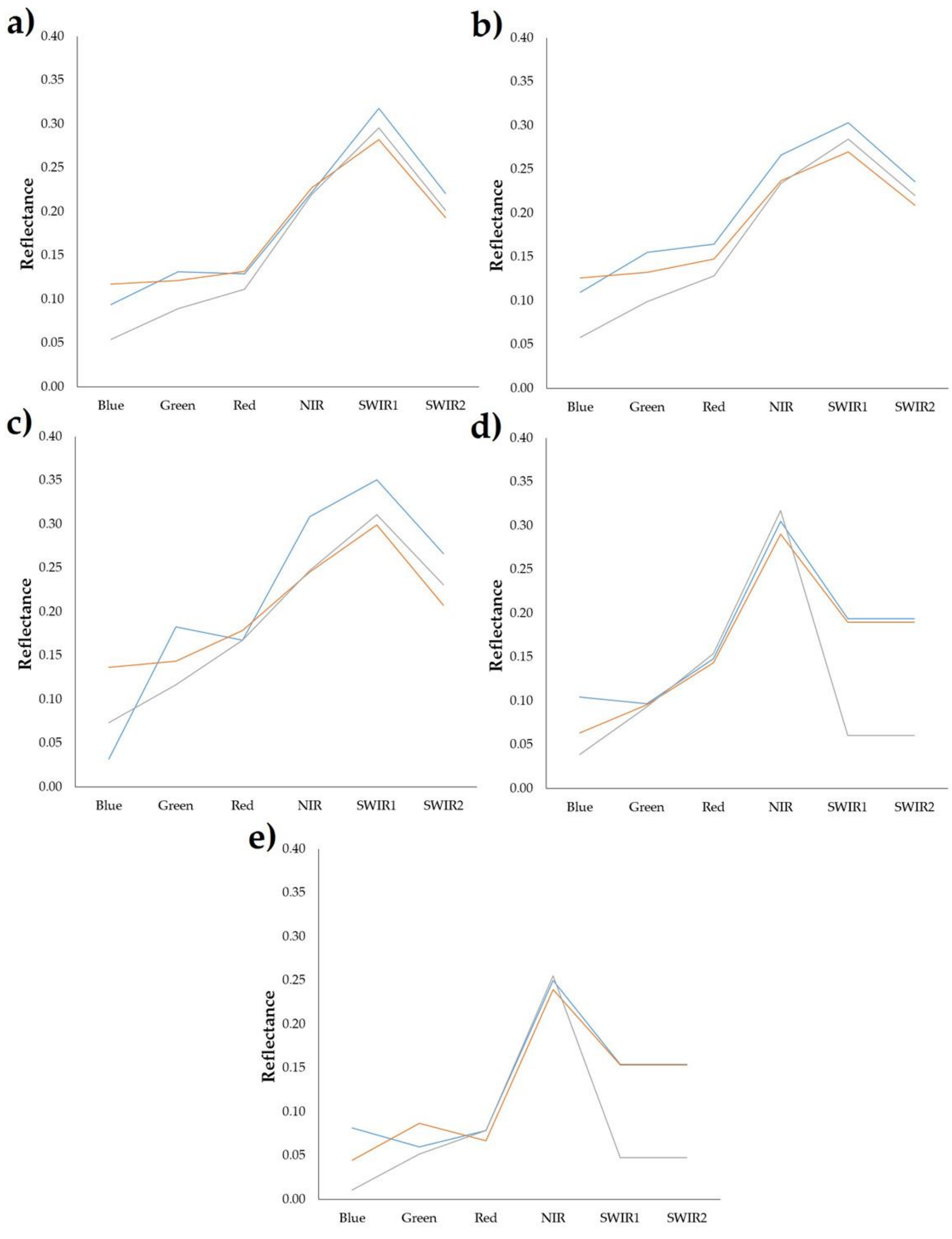

Figure 8. Spectral response of each correction method in Teseachi for 2010 a); 2011 (b); 2012 (c); 2013 (d); 2014 (e). ATCOR2 (—), DOS1 (—) and FLAASH (—).

These dissimilarities are also due to some combinations of adjustments in the radiometric calibration (Manakos *et al*., 2011). Furthermore, it could also be due to the effects caused by the heterogeneity of the sites and the grassland itself.

The ANOVA applied to the data from the three PMS determined that there were significant differences ($P<0.05$) for all of the CMs in the visible region of the spectrum. The NIR showed a greater number of non-significant results for the three sites. In the case of Eden, all of the CMs showed significant differences for all of the bands, except for the SWIR 1 in 2011 (ATCOR2 and DOS1) and the red region in 2013 (DOS1 and FLAASH). Similarly, significant differences were also detected for El Sitio, with the exception of the NIR with the ATCOR2 method in 2010. Another two exceptions were the SWIR 1 in 2010 (ATCOR2) and the red region in 2013 (ATCOR2). Finally, the ANOVA detected significant differences for Teseachi; however, this site showed the greatest number of non-significant results.

Estimated Annual Biomass

Table 2 shows the accuracy of the biomass estimation during 2010–2014 when applying the CMs to the scenes of each PMS. Variability on the values of R^2 are observed among all the years and CMs. The contribution of the spectral bands to each SLR model was calculated. The results showed that, in Eden, the bands were not good biomass estimators ($P>0.05$) when used separately. For El Sitio, the results showed significant values of R^2 ($P<0.05$) for the NIR in 2014 for the three CMs. In Teseachi, a larger number of significant values of R^2 ($P<0.05$)

Table 2. Accuracy of the correction methods in the three permanent monitoring sites (PMS). The column of Band (spectrum region) corresponds to the spectral bands and the corresponding spectrum region that contributed the most for the biomass estimation.

No	Site	CA	Year	Band (Spectrum Region)	R^2	RMSE ($Kg \cdot ha^{-1}$)
1	Eden	ATCOR2	2010	1 (Blue)	0.35	72.01
2			2011	1 (Blue)	0.22	23.92
3			2012	n/d	n/d	n/d
4			2013	2 (Blue)	0.38	153.31
5			2014	6 (SWIR1)	0.05	321.27
6		DOS1	2010	2 (Green)	0.02	88.36
7			2011	1 (Blue)	0.12	25.42
8			2012	n/d	n/d	n/d
9			2013	5 (NIR)	0.19	174.70
10			2014	5 (NIR)	0.48 *	237.99
11		FLAASH	2010	1 (Blue)	0.19	80.11
12			2011	1 (Blue)	0.20	24.22
13			2012	n/d	n/d	n/d
14			2013	2 (Blue)	0.31	161.73
15			2014	7 (SWIR2)	0.04	323.04
1	El Sitio	ATCOR2	2010	1 (Blue)	0.08	98.32
2			2011	1 (Blue)	0.09	54.14
3			2012	2 (Green)	0.07	54.14
4			2013	2 (Blue)	0.40	34.43
5			2014	5 (NIR)	0.81 *	84.16
6		DOS1	2010	5 (SWIR1)	0.10	97.21
7			2011	7 (SWIR2)	0.26	302.29
8			2012	2 (Green)	0.07	59.06
9			2013	5 (NIR)	0.46 *	32.54

10			2014	5 (NIR)	0.81 *	84.16
11			2010	1 (Blue)	0.02	101.54
12			2011	1 (Blue)	0.27	29.10
13		FLAASH	2012	2 (Green)	0.08	58.94
14			2013	5 (NIR)	0.67 *	26.24
15			2014	5 (NIR)	0.79 *	86.66
1			2010	5 (SWIR1)	0.26	82.58
2			2011	3 (Red)	0.40	52.92
3		ATCOR2	2012	5 (SWIR1)	0.42	70.76
4			2013	7 (SWIR2)	0.76 *	48.79
5			2014	5 (NIR)	0.57 *	40.90
6			2010	7 (SWIR2)	0.44 *	71.42
7			2011	3 (Red)	0.40	52.93
8	Teseachi	DOS1	2012	7 (SWIR2)	0.72 *	48.65
9			2013	7 (SWIR2)	0.76 *	48.79
10			2014	5 (NIR)	0.42 *	40.90
11			2010	7 (SWIR2)	0.44 *	71.44
12			2011	3 (Red)	0.40	53.00
13		FLAASH	2012	7 (SWIR2)	0.72 *	60.76
14			2013	7 (SWIR2)	0.76 *	77.97
15			2014	6 (SWIR1)	0.57 *	40.61

* = Significant level (P<0.05).
n/d = No data.
CA = Correction algorithm.

were observed for the red region, NIR, and SWIR 2. Finally, in Teseachi, the NIR and SWIR are the most relevant spectral regions for the prediction of biomass.

The results obtained could be related to the specific atmospheric parameters included as inputs in each of the models. The DOS1 method does not refer to the atmospheric profile (Chavez *et al*., 1996) and FLAASH uses global values for its atmospheric parameters (Mattew *et al*., 2000). Figures 9 and 10 show the variations of the values of R^2 and RMSE obtained by the SLR analysis. The RMSE fluctuated in the three CMs, being DOS1 the one with the most stable and the lowest values.

The results of the PC analysis for biomass estimation in the three PMS are shown in Table 3. Two main components were considered for the PCA, which represented at least 80 % of the total data set variance. The analysis of the principal components served to group the spectral variance and to establish its relation to biomass production. Thus, each resulting component represents a reduced percentage of variability.

In the site Eden, the highest values of R^2 were obtained for DOS1 when grouping the visible and the SWR1, as well as the SWIR2, regions. The FLAASH correction algorithm showed moderate values of R^2 by grouping the visible region of the spectrum in PC1 and NIR, as well as SWIR in PC2. This result is in agreement with the findings by Hadjimitsis *et al*. (2004) who obtained reliable PCs by grouping similar regions of the spectrum. The site El Sitio showed values of R^2 greater than 0.71 and the components derived from DOS1 were the best biomass estimators. The rest of the methods did not obtain significant results. Teseachi showed values of R^2 between 0.41 and 0.98. Table 3 shows the way the Spectral

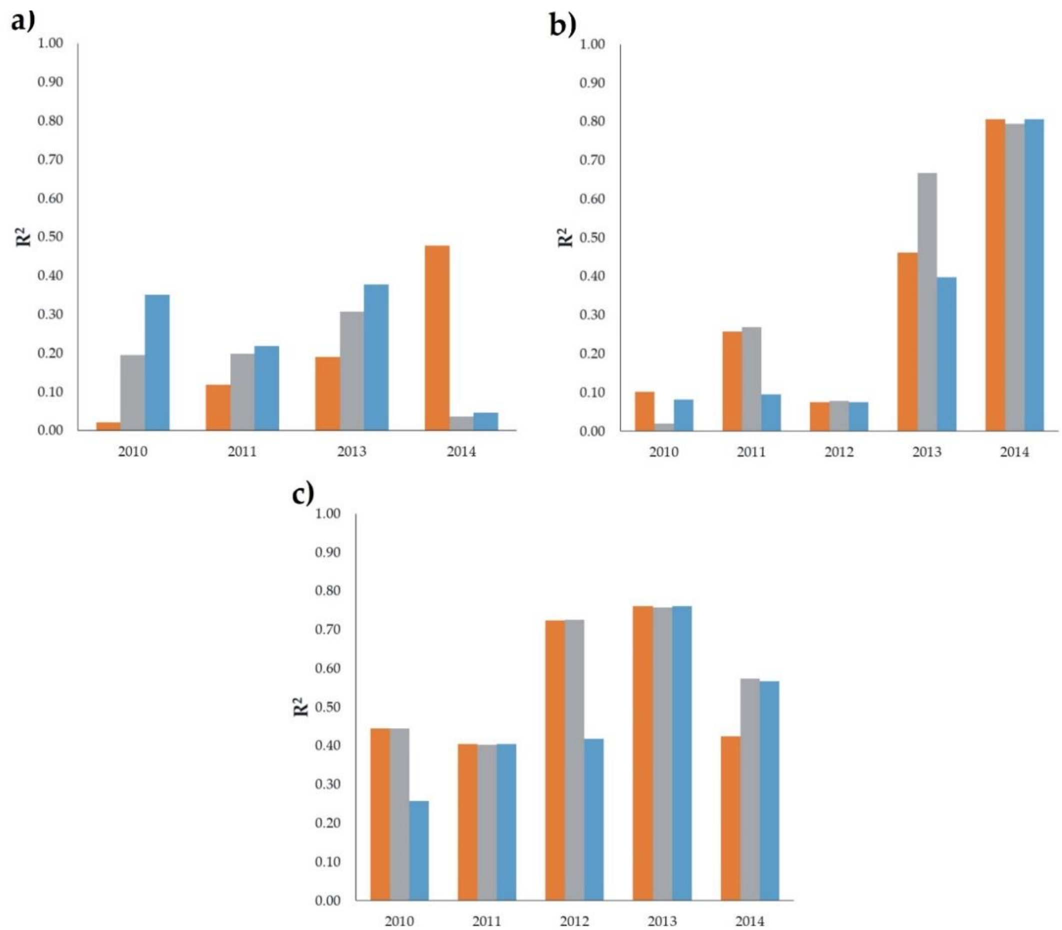

Figure 9. Variation in the precision of the correction methods represented by values of R^2. Eden (a); El Sitio (b); Teseachi (c); ATCOR2 (—), DOS1 (—) and FLAASH (—) through single band.

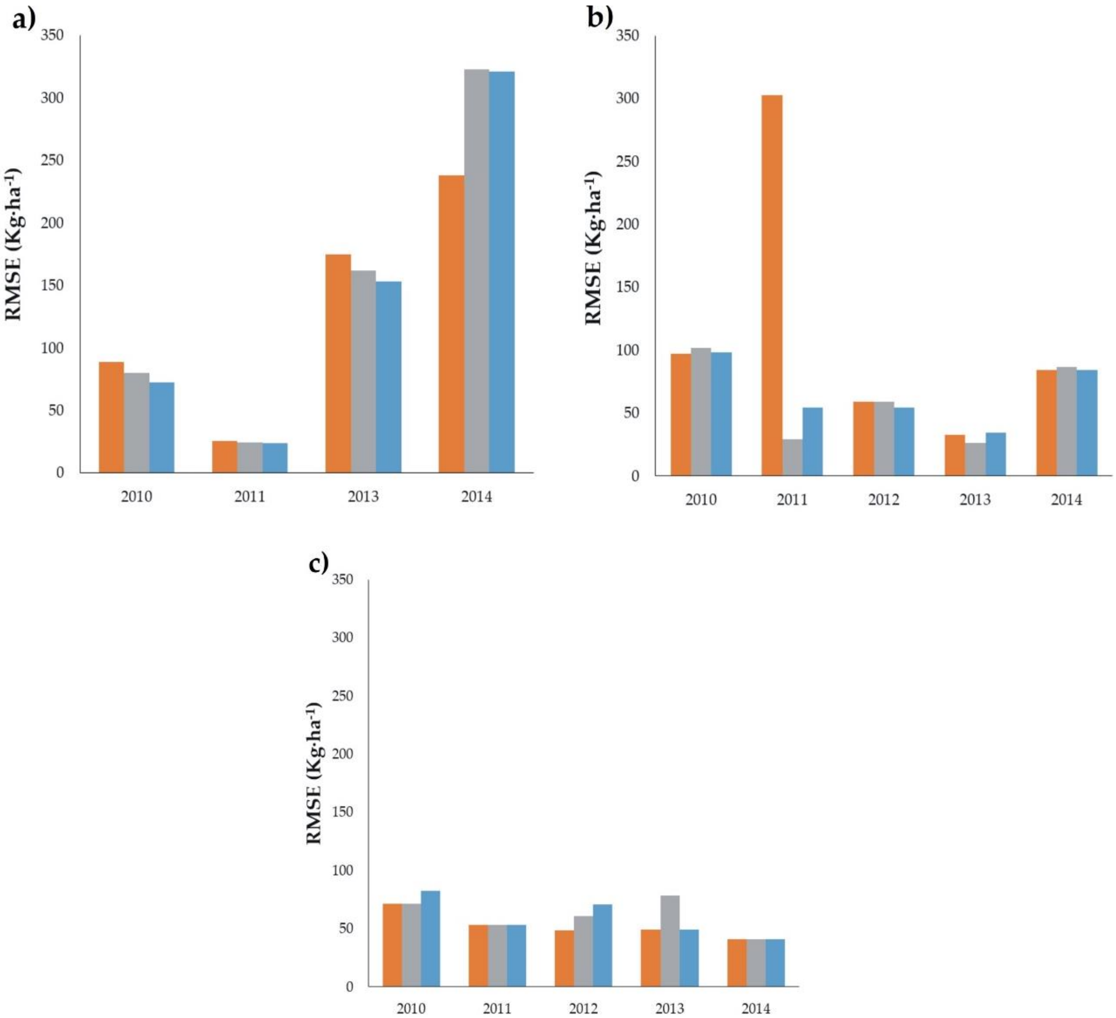

Figure 10. Variation in the precision of the correction methods represented by the RMSE (Kg·ha⁻¹). Eden (a); El Sitio (b); Teseachi (c); ATCOR2 (—), DOS1 (—) and FLAASH (—) through single band.

Table 3. Accuracy of the correction methods represented by the values of R^2 and RMSE for biomass estimation in the three PMS and the spectral bands forming the principal components.

No	Site	Correction Method	Year	PC1 Bands	PC2 Bands	R^2	RMSE ($Kg \times ha^{-1}$)
1	Eden	ATCOR2	2010	1 2 3 4	5 7	0.48	69.08
2			2011	1 2 3	4 5 7	0.45	21.54
3			2012	n/d	n/d	n/d	n/d
4			2013	2 3 4	6 7	0.55	139.53
5			2014	2 3 4	6 7	0.32	292.99
6		DOS1	2010	1 2 3 4 5 7		0.01	88.78
7			2011	1 2	3 4 5 7	0.97 *	4.97
8			2012	n/d	n/d	n/d	n/d
9			2013	2 3 4 6 7	5	0.77 *	98.49
10			2014	2 3 4 6 7	5	0.84 *	141.18
11		FLAASH	2010	1 4 7	3	0.39	75.30
12			2011	4 7	1 2 3	0.45	21.56
13			2012	n/d	n/d	n/d	n/d
14			2013	3 4 5 6 7	2	0.42	158.38
15			2014	2 3 4	5 6 7	0.63 *	214.41
16	El Sitio	ATCOR2	2010	1 2 3	4 7	0.97 *	19.10
17			2011	1 2 3	4 5	0.93 *	15.82
18			2012	2 3 4 5 7	1	0.94 *	15.20
19			2013	3 4 6 7	2 5	0.86 *	17.36
20			2014	2 3 4 6 7	5	0.95 *	44.52
21		DOS1	2010	2 3 4 5 7		0.99 *	95.61
22			2011	1 2 3 4 5 7	7	0.92 *	106.37
23			2012	2 3 4 5 7	1	0.97 **	11.46

24			2013	2 3 4 6 7	5	0.85 *	18.09
25			2014	2 3 4 6 7	5	0.94 *	43.24
26			2010	1 2 3 4 5 7		0.00	102.50
27			2011	2 3 4	1 5 7	0.92 *	9.90
28		FLAASH	2012	1 2	3 4 5 7	0.97 **	11.39
29			2013	3 4 6 7	2 5	0.71	34.29
30			2014	2 3 4 6 7	5	0.95 **	45.61
31			2010	1 2 3 4	5 7	0.87 *	36.92
32			2011	1 2 3		0.41	52.32
33		ATCOR2	2012	1 2 3 4	5 7	0.88 *	77.96
34			2013	3 4 5	2 6 7	0.94 *	25.92
35			2014	2 3 4 5 6 7		0.60 *	25.35
36			2010	1 2 3 4	5 7	0.97 **	14.69
37			2011	1 2 3		0.41	13.09
38	Teseachi	DOS1	2012	1 2 3 4	5 7	0.88 *	33.62
39			2013	2 3 6 7	4 5	0.93 *	28.03
40			2014	2 3 4 5 6 7		0.63 *	23.53
41			2010	1 2 3 4	5 7	0.98 **	14.63
42			2011	1 2 3		0.41	13.24
43		FLAASH	2012	1 2 3 4	5 7	0.88 *	33.49
44			2013	2 3 6 7	4 5	0.84 *	42.16
45			2014	2 6 7	3 4 5	0.71	25.50

* = Significant level at P <0.05.
** = Significant level at P <0.01.
n/d = No data
PC1= Principal component 1.
PC2 = Principal component 2.

bands were grouped to form the principal components. Among the three CMs, DOS1 showed the most consistent outputs for all of the years.

Figures 11 and 12 show the variations of R^2 and RMSE obtained by the SLR analysis between biomass and the spectral information. The site that showed the highest precision was El Sitio. The method of DOS1 produced the most stable and the most precise results among the sites.

The site Eden obtained low values of the coefficient of determination (0.3–0.5) when corrected with ATCOR2. By correcting the data with DOS1, the precision improved significantly ($P<0.05$) for the data of 2011–2014 with R2 values greater than 0.77. This indicates that DOS1 can estimate biomass production with a greater precision in spite of the accelerated changes in the succession of the grassland and the great density, as well as diversity of plants, in the site.

El Sitio obtained precise results for biomass estimations by applying the three CMs, which indicates homogeneity in the reflectance of the grassland. For Teseachi, good yields were obtained in 2010, 2012, and 2013 with the three CMs, with a similar precision to El Sitio.

In this study, we have reviewed three correction methods of satellite images applied them to a temporal series of 12 scenes. The precision of each method was assessed through values of R^2 and RMSE. The DOS1 method, which is the simplest, provided a reasonable correction in the bands of the visible spectrum (Song *et al*., 2001; Cui *et al*., 2014). Given that the input parameters for DOS1 are derived from the image itself, it makes the method relatively easy to

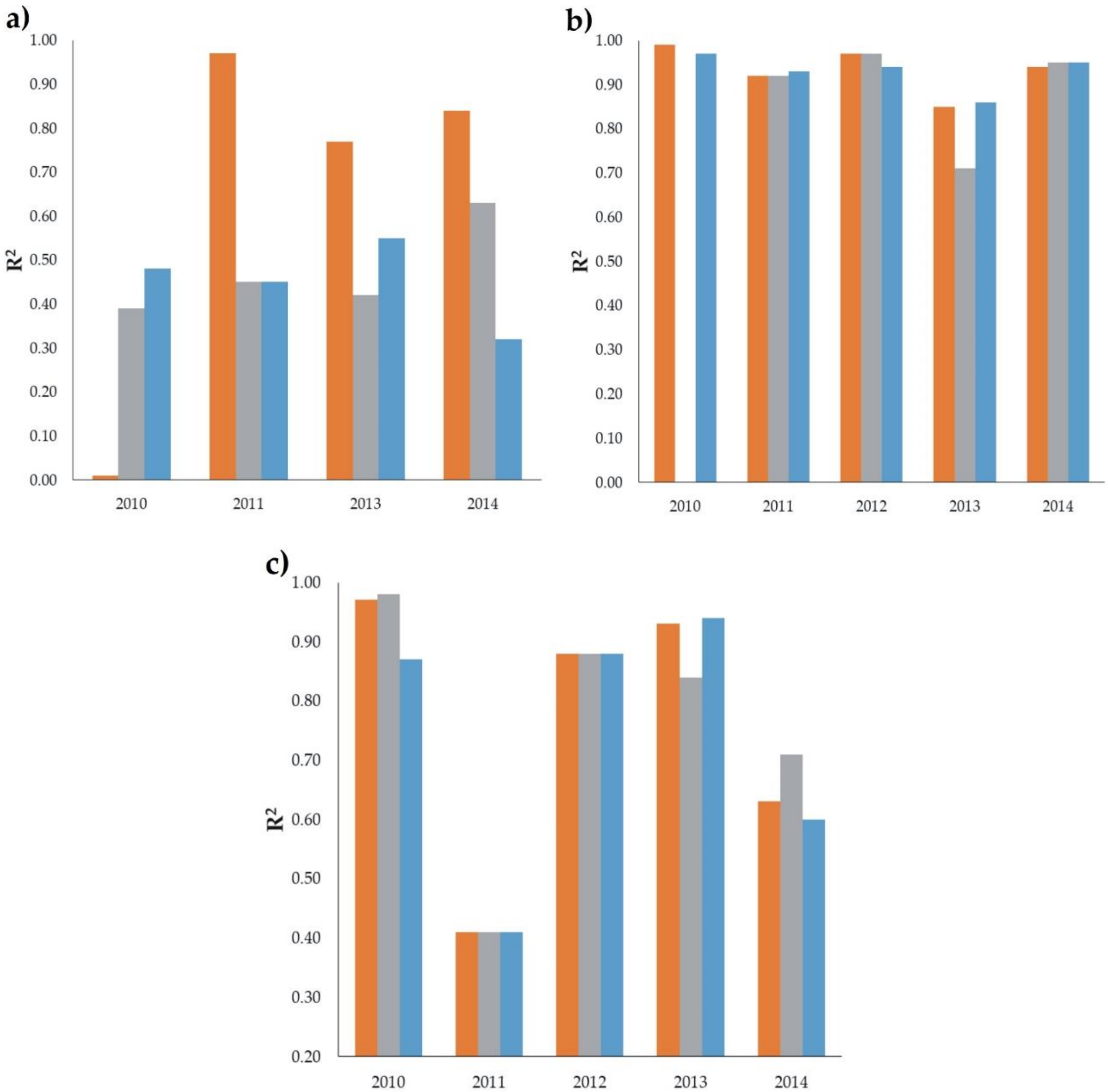

Figure 11. Variation in the precision of the correction methods represented by values of R^2. Eden (a); El Sitio (b); Teseachi (c); ATCOR2 (—), DOS1 (—), and FLAASH (—) through PC.

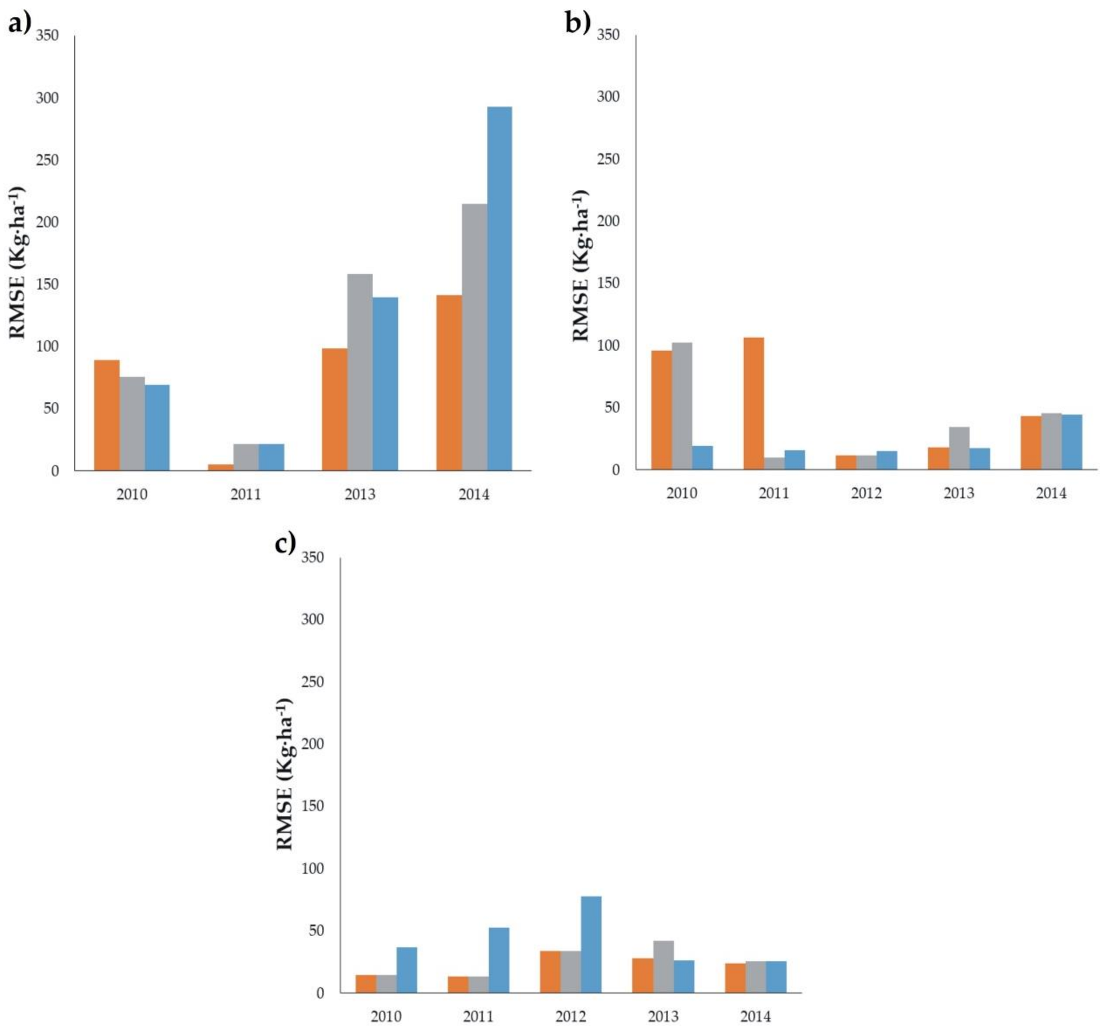

Figure 12. Variation in the precision of the correction methods represented by RMSE (kg·ha^{-1}). Eden (a); El Sitio (b); Teseachi (c); ATCOR2 (—), DOS1 (—), and FLAASH (—) through PC.

implement. Thus, it is preferred over more sophisticated methods that require the acquisition of atmospheric or meteorological data (Chavez, 1996; Kaufman *et al*., 1997; Liang *et al*., 2002). The time required for each method can be a crucial factor when using multiple sets of images (Mahiny and Turner, 2007). The methods of ATCOR2 and FLAASH required much more processing time than DOS1. These requirements limit their application, especially when the historical atmospheric information is limited (Janzen *et al*., 2006; Cui *et al*., 2014; Mason *et al*., 2015). The differences in the results from ATCOR2 and DOS1 were probably due to the availability of reliable atmospheric historical data, which may have conferred to DOS1 a better performance (Janzen *et al*., 2006). Likewise, the differences between ATCOR2 and FLAASH, when performing radiometric correction on spectral data from vegetation, may be due to the water content of such vegetation (Wu *et al*., 2005).

The CMs allowed for the conversion of digital numbers to reflectance values. The spectral reflectance of grasslands was low in the visible region for all of the three PMS during 2010–2014. Chlorophyll absorbs most of the light received on the photosynthetically active radiation range of the spectrum. Consequently, reflectance was higher in the NIR, SWIR 1, and SWIR 2, indicating a contrast between these and the aforementioned visible regions of the electromagnetic spectrum (Chuvieco *et al*., 2010; Sonobe *et al*., 2017). The variation in biomass production estimation for the period 2010–2014 can be largely explained by changes in vegetation, its growth conditions, and its distribution. As mentioned by Yan *et al.* (2015), the growth conditions in semi-arid and arid regions are largely affected by temperature and precipitation. In addition, human activities, such as grazing and farming, importantly affect the vegetation distribution. It was observed that in the period when the scenes were taken, the spectral signature of the grassland varied along the years. This can be explained by the effects of density, weight, coverage, and shade, which are variable in grass communities (Manakos *et al*., 2011; Rodriguez-Maturino *et al*., 2017).

The comparison of the different CMs was based on the bands or principal components that contributed the most to the estimations of biomass. This study proved that the DOS1 method may be enough for radiometric correction in grassland areas, given the data required, which include only a few parameters (Wu *et al*., 2005). The results obtained in this study indicate that DOS1 is sufficient to correct images used in the estimation of structural variables of the grassland. In this sense, it may not be necessary to employ complex algorithms when evaluating areas of grassland (Song *et al*., 2001; Janzen *et al*., 2006)

The ANOVA applied to the three PMS revealed significant differences among DOS1, FLAASH, and ATCOR2 in the visible and the infra-red regions of the spectrum. This is consistent with the results reported by Vicente-Serrano *et al*. (2008) and Nazeer *et al*. (2014). In general, band 4 showed the most non-significant differences, followed by bands 7 and 3.

The correlations between biomass production and the spectral data obtained the highest values when the bands of red and infra-red were employed. In contrast, the relationships between biomass and the spectral values of the visible region were weak, indicating the sensitivity of this spectral range to the atmospheric variation, which agrees with that reported by Roy and Ravan (1996). Thus, in the temporal estimation and quantification of biomass, the application of a method for radiometric correction is necessary to eliminate the temporal variability. For most of the studied years, we found that FLAASH showed the highest reflectance values in the NIR. Such high values affected the relationships between biomass and spectral values. Previous studies have shown that high values of reflectance affect the estimations of biomass and the determination of potential areas for grassland production (Song *et al*., 2001).

The low biomass production values predicted for the three PMS could be due to site conditions or the date of data collection. Thus, the evaluation period may not be the best period to estimate biomass by using multitemporal data from the Landsat sensor. Our results confirmed a lack of association between spectral and biomass data. The relationship between the biomass and the bands (Table 2), and between biomass and the principal components (Table 3), for all of the PMS evaluated confirmed that there was a high variability. Such variability could neither be explained by the spectral response of the bands nor by the principal components. It is possible that the high values of reflectance in the infra-red region that showed in the site Eden after applying the DOS1 method are due to the atmospheric dispersion (Lu *et al*., 2012).

For the atmospheric and radiometric corrections, we have documented the differences between the reflectance values after applying the CMs to the data from different sites. In comparison with other multitemporal studies (Tan *et al*., 2012; Nguyen *et al*., 2015; Raab *et al*., 2015), we have tested different CMs with a change of platform from Landsat 5 to Landsat 8 in grassland areas. In the site Eden, we found that the deviations between the estimated and measured biomass were high for 2013 and 2014. This happened when the biomass was estimated with both the SLR and the MLR. Thus, such deviations can be attributed to the change of platform (TM to ETM+ and ETM+ to OLI). This change can cause instability in the time series (Schroeder *et al*., 2006). Eden was the site with the lowest values of the coefficient of determination. The MLR with PCs was performed to include all of the spectral variability and correlate it to biomass; however, the results showed low correlation values. Therefore, such spectral variability may be influenced by other sources of variation, such as climate, topography, and invasive species (Xie *et al*., 2009). For El Sitio, and analyzing the data with SLR, low values of R^2 were obtained; in contrast, higher correlations were found with the MLR. This shows that biomass can be greatly estimated from the spectral data. More precise estimates, evidenced with higher R^2 values, were produced with the data from Teseachi when analyzed with both the SLR and the MLR. The correlation of biomass and spectral data determined through MLR showed that, in at least two of the three PMS, the variation can be explained with the set of bands of the visible and infra-red ranges of the spectrum. Conversely, in one of the sites, it is necessary to collect more variables to explain the biomass variation. The inclusion of PCs in the biomass estimation allowed for an explanation of the spectral variability more effectively.

CONCLUSIONS AND RECOMMENDATIONS

Atmospheric correction is a crucial step in the pre-processing of satellite images. Landsat images and biomass data from the field were employed to test the precision of three algorithms of atmospheric correction. According to the results, DOS1 presented the highest correlation values, pointing it out as a good method for the atmospheric correction of Landsat images and its application for the estimation of grassland biomass production.

Determination of the grassland production for Teseachi obtained the lowest deviations between estimated and measured biomass when modelling this variable with both single and multiple linear regression. The precisions of the estimates were closely related to the temporal spectral stability of the images. For Eden, the results were the least precise, which indicates that there is a great variation in the terrain that is difficult to explain by the satellite images. Biomass estimation using field and spectral data, coupled with an adequate atmospheric correction method, can accurately reflect grassland characteristics. For future studies, models that consider the effects of climate, minimum and maximum temperatures, precipitation, and topographic data, such as elevation, slope, and aspect, could be tested for higher precisions.

Three correction methods have been proposed and tested in this study. The simplest algorithm, DOS1, provided a reasonable correction and estimated biomass accurately when employing bands in the visible and infra-red regions of the spectrum, at least for cloud-free scenes. Operationally, the DOS1 method, which derives its input parameters from the image itself and is relatively easy to implement, may be more reliable to implement over more sophisticated methods, which require the acquisition of atmospheric or meteorological historical data.

In the multitemporal estimation of grassland biomass production employing spectral information, it is not necessary to apply complex radiometric correction procedures. The use of the DOS1 method provided good results given its nature of providing reliable results on dark surfaces. However, when the spectral signal is affected by various sources of variation, it may be necessary to apply high-precision radiometric corrections, such as ATCOR2 or FLAASH. To make continuous estimates of biomass by remote sensors, it is preferable to employ a set of homogeneous Landsat images provided by a single platform.

Precise estimates of biomass will allow for the quantification of potential carbon stores by grasslands, serve to regulate animal load based on temporal estimation, and ensure control on the use of the grassland biomass. Estimating the spatial distribution of biomass is of great importance to support the study of grassland ecology and its socioeconomic environment. This study proved that it is possible to estimate grassland biomass production by remote sensing through an SLR analysis. Grassland biomass maps can be used as a reference to assess the grassland condition, the grazing capacity, and potential animal production. The use of remote sensing tools in grassland ecosystems is important for their monitoring, conservation, and protection.

REFERENCES

Adler-Golden, S., A. Berk, L.S. Bernstein, S. Richtsmeier, P.K. Acharya, M.W. Matthew, and J.H. Chetwynd. 1998. FLAASH, A MODTRAN4 atmospheric correction package for hyperspectral data retrievals and simulations. In Proceedings of the 7th Annal JPL Airborne Earth Science Workshop, Pasadena, CA, USA, 4–8 March, pp. 9.14.

Agrawal, G., J. Sarup, and M. Bhopal, 2011. Comparision of QUAC and FLAASH atmospheric correction modules on EO-1 Hyperion data of Sanchi. Int. J. Adv. Eng. Sci. Technol., 4:178–186.

Anderson, G.P., G.W. Felde, Hoke, M.L. A.J. Ratkowski, T.W. Cooley, J.H. Chetwynd, Jr., and L.S. 2002. Bernstein, MODTRAN4-based atmospheric correction algorithm: FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes. In Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery VIII; Society of Photo Optics: Orlando, FL, USA, Volume 4725, pp. 65-72.

Barsi, J.A., K. Lee, G. Kvaran, B.L. Markham, and J.A. Pedelty. 2014. The spectral response of the Landsat-8 operational land imager. Remote Sens. 6:10232-10251.

Berk, A., L.S. Bernstein, G.P. Anderson, P.K. Acharya, D.C. Robertson, J.H. Chetwynd, and S.M. Adler-Golden. 1998. MODTRAN cloud and multiple scattering upgrades with application to AVIRIS. Remote Sens. Environ. 65:367-375.

Callieco, F., and F. Dell’Acqua, A comparison between two radiative transfer models for atmospheric correction over a wide range of wavelengths. Int. J. Remote Sens. 2011, 32:1357-1370.

CGG-SAGARPA; COLPOS. 2009. Manual de Campo Para la Red Básica del Sistema Nacional de Monitoreo Terrestre Orientado a la Ganadería; Coordinación General de Ganadería de la Secretaria de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación: México, D.F.

Chander, G. B., L. Markham and D.L. Helder. 2009. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. Remote Sens. Environ. 113:893–903.

Chang, J., D.E. Clay, L. Leigh, D. Aaron, K. Dalsted, and M. Volz. 2008. Evaluating modified atmospheric correction methods for Landsat Imagery: Image-based and model-based calibration methods. Commun. Soil Sci. Plan., 39:1532-1545.

Chavez, P.S. 1996. Image-based atmospheric corrections revisited and improved. Photogramm. Eng. Remote Sens. 62:1025-1036.

Chavez, P.S., Jr. 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote Sens. Environ. 24: 459–479.

Chuvieco, E., I. Aguado, M. Yebra, H. Nieto, J. Salas, M.M. Pilar, L. Vilar, J. Martínez, S. Martin, and P. Ibarra. 2010. Development of a framework for fire risk assessment using remote sensing and geographic information system technologies. Ecol. Model. 221:46-58.

Cohen, W.B. and S.N. Goward. 2004. Landsat's role in ecological applications of remote sensing. BioScience, 54:535-545.

CONABIO (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad). 2014. La Biodiversidad en Chihuahua: Estudio de Estado; Comisión Nacional para el Conocimiento y Uso de la Biodiversidad: Distrito Federral, México.

Congedo, L. 2013. Semi-Automatic Classification Plugin for QGIS; Technical Report; Sapienza University, ACC Dar Project: Rome, Italy.

Craine, J.M. 2013. Long-term climate sensitivity of grazer performance: A cross-site study. PLoS ONE, 8:e67065.

Cui, Q.Y., M.J. Gaillard, G. Lemdahl, L. Stenberg, S. Sugita, and Zernova, G. 2014. Historical land-use and landscape change in southern Sweden and implications for present and future biodiversity. Ecol. Evol., 4:3555-3570.

Dusseux, P., L. Huber-Moy, T. Corpetti and F. Vertès. 2015. Evaluation of SPOT imagery for the estimation of grassland biomass. Int. J. Appl. Earth Obs. 38:72-77.

El Hajj, M., A. Bégué, B. Lafrance, O. Hagolle, G. Dedieu, and Rumeau, M. 2008. Relative radiometric normalization and atmospheric correction of a SPOT 5 time series. Sensors, 8:2774–2791.

Estrada-Castillón, E., L. Scott-Morales, J.A. Villarreal-Quintanilla, E. Jurado-Ybarra, M. Cotera-Correa, C. Cantú-Ayala, and J. García-Pérez. 2010. Clasificación de los pastizales halófilos del noreste de México asociados con perrito de las praderas (*Cynomys mexicanus*): Diversidad y endemismo de especies. Revista mexicana de biodiversidad, 81:401–416.

Fan, J.W., H.P. Zhong, W. Harris, G.R. Yu, S.Q. Wang, Z.M. Hu, Y.Z. Yue. 2007. Carbon storage in the grassland of China based on field measurement of above- and below- ground biomass. Clim. Chang. 86:375–396.

Geosystems, 2013. Haze Reduction, Atmospheric and Topographic Correction. User Manual ATCOR2 and ATCOR3; Geosystems GmbH: Geneva, Switzerland.

Goetz, S. J., S. D. Prince, Goward, S. N.; M. M. Thawley and J. 1999. Small. Satellite remote sensing of primary production: An improved production efficiency modeling approach. Ecol. Model. 122: 239–255.

Hadjimitsis, D.G., C.R.I. Clayton, and V.S. Hope. 2004. An assessment of the effectiveness of atmospheric correction algorithms through the remote sensing of some reservoirs. Int. J. Remote Sens. 25:3651–3674.

Hansen, M.C. and Loveland, T.R. 2012. A review of large are monitoring of land cover change using Landsat data. Remote Sens. Environ. 122: 66–74.

He, C., Q. Zhang, Y. Li, X. Li and P. Shi. 2005. Zoning grassland protection area using remote sensing and cellular automata modeling-A case study in Xilingol steppe grassland in northern China. J. Arid Environ. 63:814–826.

IUCN (International Union for Conservation of Nature). Available online: http://www.iucnredlist.org/details/18868/0 (accessed on 20 November 2016).

Janzen, T.D.; L.A. Fredeen, D.R. Wheate. 2006. Radiometric correction techniques and accuracy assessment for Landsat TM data in remote forested regions. Can. J. Remote Sens., 32:330-340.

Jin, Y.X., X.C Yang, J.J. Qiu, J.Y. Li, T. Gao, Q. Wu, F. Zhao, H.L. Ma, H.D. Yu, B. Xu. 2014. Remote sensing-based biomass estimation and its spatio-temporal variations in temperate grassland, Northern China. Remote Sens. 6:1496-1513.

Jurado, G.P. T.R.A. Saucedo, N.C.R. Morales, and S.M. Martínez. 2013. Almacén y Captura de Carbono en Pastizales y Matorrales de Chihuahua; INIFAP: Chihuahua, Mexico.

Kaufman, Y.J.; A.E. Wald, L.A. Remer, B.C. Gao, R.R. Li, L. Flynn. 1997. The MODIS 2.1-um channel-correlation with visible reflectance for use in remote sensing of aerosol. IEEE Trans. Geosci. Remote. 35:1286-1298.

Körner, I., L. Oreopoulos, G. Feingold, L.A. Remer, O. Altaratz. 2008. How small is a small cloud? Atmos. Chem. Phys., 8, 3855-3864.

Kruse, F.A. 2004. Comparison of ATREM, ACORN, and FLAASH atmospheric corrections using low-altitude AVIRIS data of Boulder, CO. In Proceedings of the 13th JPL Airborne Geoscicence Workshop, Pasadena, CA, USA, 31 March–2 April.

Li, J.Y., X.C. Yang, Y.X. Jin, Z. Yang, W.G. Huang, L.N. Zhao, T. Gao, H.D. Yu, H.L. Ma and Z.H. Qin. 2003. Monitoring and analysis of grassland desertification dynamics using Landsat images in Ningxia, China. Remote Sens. Environ. 138, 19-26.

Liang, S.; H. Fang, J.T. Morisette, M. Chen, C.J. Shuey, C.L. Walthall, C.S. Daughtry. 2002. Atmospheric correction of Landsat ETM+ land surface imagery. II. Validation and applications. IEEE Trans. Geosci. Remote Sens., 40:736-2746.

López-Serrano, P.M., J.J. Corral-Rivas, R.A. Díaz-Varela, J.G. Álvarez-González, López-Sánchez, C.A. 2016. Evaluation of radiometric and atmospheric correction algorithms for aboveground forest biomass estimation using Landsat 5 TM data. Remote Sens. 8:369.

Lu, D., P. Mausel, E. Brondizio, and E. Moran. 2002. Assessment of atmospheric correction methods for Landsat TM data applicable to Amazon basin LBA research. Int. J. Remote Sens. 23:2651-2671.

Lyapustin, A.; D.L. Williams, B. Markham, J. Irons, B. Holben, and Y. Wang. 2004. A method for unbiased high-resolution aerosol retrieval from Landsat. J. Atmos. Sci., 61:1233-1244.

Mahiny, A.S. and B.J. Turner. 2007. A comparison of four common atmospheric correction methods. Photogramm. Eng. Remote Sens. 73:361-368.

Manakos, I., K. Manevski, C. Kalaitzidis, and D. Edler. 2011. Comparison between atmospheric correction modules on the basis of Worldview-2 imagery and in situ spectroradiometric measurements. In Proceedings of the 7th EARSeL Workshop on Imaging Spectroscopy, Edinburgh, UK, 11–13 April; pp. 1–12.

Manjarrez–Domínguez, C., A. Pinedo–Alvarez, C. Pinedo–Alvarez, F. Villarreal–Guerrero and L. Cortes-Palacios. 2015. Vegetation landscape analysis due to land use changes on arid lands. Pol. J. Ecol. 63:272–279.

Marcello, J., F. Eugenio, U. Perdomo, and Medina, A. 2016. Assessment of atmospheric algorithms to retrieve vegetation in natural protected areas using multispectral high resolution imagery. Sensors. 16:1624.

Marsett, C.R., J. Qi, P. Heilman, H.S. Biedenbender, C.M. 2006. Watson, S. Amer, M. Weltz, D. and Goodrich, R. Marsett. Remote sensing for grassland management in the arid Southwest. Rangel. Ecol. Manag. 59:530-540.

Martins, V.S., C.CF. Barbosa, L.A.S. deCarvalho, D.S.F. Jorge, F.D.L. Lobo, E.M.L.D.M Novo. 2017. Assessment of atmospheric correction methods for Sentinel-2 MSI images applied to Amazon floodplain lakes. Remote Sens. 9: 322.

Mason, S., A. Kruczkiewicz, P. Ceccato, and A. Crawford. 2015. Accessing and Using Climate Data and Information in Fragile, Data-Poor States; International Institute for Sustainable Development: Winnipeg, MB, Canada,.

Matthew, M.W., S.M. Adler-Golden, A. Berk, S.C. Richtsmeier, R.Y. Levine, L.S. Bernstein, P.K. Acharya, G.P. Anderson, G.W. Felde, and M.P. Hoke. 2000. Status of atmospheric correction using a MODTRAN4-based algorithm. In Proceedings of the Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI, Orlando, FL, USA, 24−26 April; pp. 199−207.

Naesset, E., T. Gobakken, O.M. Bollandsås, T.G. Gregoire, R. Nelson., G. Ståhl, 2013. Comparison of precision of biomass estimates in regional field sample surveys and airborne LiDAR-assisted surveys in Hedmark County, Norway. Remote Sens. Environ. 130:108-120.

Nazeer, M., J.E. Nichol, Yung, Y.K. 2014. Evaluation of atmospheric correction models and Landsat surface reflectance product in an urban coastal environment. Int. J. Remote Sens. 35: 6271–6291.

Neubert, M. and G. Meinel. 2005. Atmospheric and terrain correction of IKONOS imagery using ATCOR3. In ISPRS Workshop 2005: High-Resolution Earth Imaging for Geospatial Information; Hannover, Germany, 17–20.

Nguyen, H.C., J. Jung, J. Lee, S.U. Choi, S.Y. and Hong, J. Heo. 2015. Optimal atmospheric correction for above-ground forest biomass estimation with the ETM+ remote sensor. Sensors. 15:18865-18886.

Ni, J. 2004. Estimating net primary productivity of grasslands from field biomass measurements in temperate northern China. Plant Ecol. 174:217-234,

Paolini, L., F. Grings, J.A. Sobrino, J.C. Jiménez Muñoz, and H. Karszenbaum, 2006. Radiometric correction effects in Landsat multi-date/multi-sensor change detection studies. Int. J. Remote Sens. 27:685-704.

Pinedo, A.C.; A.A. Pinedo, Q.R.M. Martinez. 2007. Análisis de áreas deforestadas en la región centro-norte de la Sierra Madre Occidental, Chihuahua, México. TECNOCIENCIA Chihuahua, 1:36-43.

Pons, X.; L. Pesquer, J. Cristóbal and O. González-Guerrero. 2014. Automatic and improved radiometric correction of Landsat imagery using reference values from MODIS surface reflectance images. Int. J. Appl. Earth. Obs., 33, 243–254.

Prieto-Amparán, J.A., A. Pinedo-Alvarez, F. Villarreal-Guerrero, C. Pinedo-Alvarez, C. Morales-Nieto, C. 2016. Manjarrez-Domínguez. Past and future spatial growth dynamics of Chihuahua city, Mexico: Pressures for land use. ISPRS Int. Geo-Inf. 5:235.

Raab, C., B. Barrett, F. Cawkwell, and S. Green. 2015. Evaluation of multi-temporal and multi-sensor atmospheric correction strategies for land-cover accounting and monitoring in Ireland. Remote Sens. Lett. 6:784–793.

Richter, R. D. Schlapfer, A., and Muller. 2006. An automatic atmospheric correction algorithm for visible/NIR imagery. Int. J. Remote Sens. 27:2077-2085.

Rodríguez-Maturino, A., J.H. Martínez-Guerrero, I. Chairez-Hernández, M.E. Pereda-Solis, F. Villarreal-Guerrero, M. Renteria-Villalobos, A. Pinedo-Alvarez. 2017. Mapping land cover and estimating the grassland structure in a priority area of the Chihuahuan desert. Land. 6:70.

Roy, D.P., H.K. Zhang, J. Ju, J.L. Gomez-Dans, P.E. Lewis, C.B. Schaaf, Sun, Li Q., J.; H. Hung and V. Kovalskyy. 2016. A general method to normalized Landsat reflectance data to nadir BRDF adjusted reflectance. Remote Sens. Environ. 176:255-271.

Roy, P.S., S.A. Ravan. 1996. Biomass estimation using satellite remote sensing data-An investigation on possible approaches for natural forest. J. Biosci., 21, 535-561.

Ruppert, J.C. and A. Linstädter. 2014. Convergence between ANPP estimation methods in grasslands - A practical solution to the comparability dilemma. Ecol. Indic., 36:524–531.

Samson, F.B., F.L. Knopf, and W.R. Ostlie. 2004. Great Plains ecosystems: Past, present, and future. Wildl. Soc. Bull. 32:6-15.

Schino, G.; Borfecchia, F.; De Cecco, L. Satellite estimate of grass biomass in a mountainous range in central Italy. Agrofor. Syst. 2003, 59:157-162.

Schroeder, T.A., W.B. C. Cohen, Song, M.J. Canty, Yang, Z. 2006. Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. Remote Sens. Environ.103, 16-26.

Song, C., C.E. Woodcock, K.C. Seto, M.P. Lenney, and S.A. Macomber. 2001. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects?. Remote Sens. Environ. 75:230–244.

Sonobe, R. Y., Yamaya, H. Tani, X. Wang, N. Kobayashi, and K. Mochizuki. 2017. Mapping crop cover using multi-temporal Landsat 8 OLI imagery. Int. J. Remote Sens. 38:4348-4361.

Tan, K.C., H. San Lim, M. Z. MatJafri, K. Abdullah. 2012. A comparison of radiometric correction techniques in the evaluation of the relationship between LST and NDVI in Landsat imagery. Environ. Monit. Assess. 184:3813-3829.

United States Geological Survey. Available online: http://glovis.usgs.gov (accessed on 20 June 2016).

Vázquez-Quintero, G.; R. Solís-Moreno, M. Pompa-García, F. Villarreal-Guerrero, C. Pinedo-Alvarez, A. 2016. Pinedo-Alvarez. Detection and projection of forest changes by using the Markov Chain Model and Cellular Automata. Sustainability. 8:236.

Vicente-Serrano, S.M., F. Pérez-Cabello, T. Lasanta. 2008. Assessment of radiometric correction techniques in analyzing vegetation variability and change using time series of Landsat images. Remote Sens. Environ. 112:3916-3934.

Woodcock, C.E., S.A. Macomber, Pax-Lenney, M. W.C. Cohen. 2001. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. Remote Sens. Environ. 78:194-203.

Wu, J., D. Wang, and M.E. Bauer. 2005. Image-based atmospheric correction of QuickBird imagery of Minnesota cropland. Remote Sens. Environ. 99:315-325.

Wulder, M.A., J.C. White, S.N. Goward, J.G. Masek, J.R. Irons, M. Herold, W.B. Cohen, T.R. Loveland and C.E. Woodcock. 2008. Landsat continuity: Issues and opportunities for land cover monitoring. Remote Sens. Environ. 112:955-969.

Xie, Y. C.; Z. Y. M. Sha, Yu, Y. F. Bai. and L.A. Zhang. 2009. Comparison of two models with Landsat data for estimating above ground grassland biomass in Inner Mongolia, China. Ecol. Model. 220:1810-1818.

Yan, J., C. Zhu, W. Liu, F. Luo, J. Mi, Y. Ren, J. Li, T. Sang. 2015. High photosynthetic rate and water use efficiency of Miscanthus lutarioriparius characterize an energy crop in the semiarid temperate region. GCB Bioenergy. 7:207-218.

Yang, Y., J. Fang, Y. Pan and C. Ji. 2009. Aboveground biomass in Tibetan grasslands. J. Arid Environ. 73:91-95.

Zelazowski, P.; A.M. Sayer, G.E. Thomas and R.G. Grainger. 2011. Reconciling satellite-derived atmospheric properties with fine-resolution land imagery: Insights for atmospheric correction. J. Geophys Res. Atmos., 116.

Zhao, P.; L. Dengsheng, W. Guangxing, W. Chuping, H. and Yujie, Y. 2016. Shuquan, Examining spectral reflectance saturation in Landsat imagery and corresponding solutions to improve forest aboveground biomass estimation. Remote Sens., 8:469.

STUDY II. SPATIAL NEAR FUTURE MODELING OF LAND USE AND LAND COVER CHANGES IN THE TEMPERATE FORESTS OF MEXICO

BY:

Jesús A. Prieto-Amparan[1], Federico Villarreal-Guerrero[1], Martin Martínez-Salvador[1], Carlos Manjarrez-Domínguez[2], Griselda Vázquez-Quintero[2], Alfredo Pinedo-Alvarez[1].

[1]Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada km 1., C.P. 31453. Chihuahua, Chih. México.
[2]Facultad de Ciencias Agrotecnológicas, Universidad Autónoma de Chihuahua, Campus I, s/n., C.P. 31350. Chihuahua, Chih., México.

*Corresponding author: Alfredo Pinedo Alvarez; Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada km 1., C.P. 31453. Chihuahua, Chih. México. apinedo@uach.mx, tel. +52-(614)-4-34-03-63., https://doi.org/10.7717/peerj.6617.

ABSTRACT

SPATIAL NEAR FUTURE MODELING OF LAND USE AND LAND COVER CHANGES IN THE TEMPERATE FORESTS OF MEXICO

BY:

M. C. JESUS ALEJANDRO PRIETO AMPARAN

The loss of temperate forests of Mexico has continued in recent decades despite wide recognition of their importance to maintaining biodiversity. This study analyzes land use/land cover change scenarios, using satellite images from the Landsat sensor. Images corresponded to the years 1990, 2005 and 2017. The scenarios were applied for the temperate forests with the aim of getting a better understanding of the patterns in land use/land cover changes. The Support Vector Machine (SVM) multispectral classification technique served to determine the land use/land cover types, which were validated through the Kappa Index. For the simulation of land use/land cover dynamics, a model developed in Dinamica-EGO was used, which uses stochastic models of Markov Chains, Cellular Automata and Weights of Evidence. For the study, a stationary, an optimistic and a pessimistic scenario were proposed. The projections based on the three scenarios were simulated for the year 2050. Five types of land use/land cover were identified and evaluated. They were primary forest, secondary forest, human settlements, areas without vegetation and water bodies. Results from the land use/land cover change analysis show a substantial gain for the secondary forest. The surface area of the primary forest was reduced from 55.8 % in 1990 to 37.7 % in 2017. Moreover, the three projected scenarios estimate further losses of the surface are for the primary forest, especially under the stationary and pessimistic scenarios. This highlights the importance and probably urgent implementation of conservation and protection measures to preserve these ecosystems and their services. Based on the accuracy obtained and, on the models generated, results from these methodologies can serve as a decision tool to contribute to the sustainable management of the natural resources of a region.

Keywords: Markov chains, Scenarios, Cellular automata, Pessimistic scenario, Temperate forests, Dinamica-EGO, Weights of evidence, Land use/land cover change, Remote sensing.

RESUMEN

SPATIAL NEAR FUTURE MODELING OF LAND USE AND LAND COVER CHANGES IN THE TEMPERATE FORESTS OF MEXICO

POR:

M. C. JESÚS ALEJANDRO PRIETO AMPARÁN

Doctor en Philosophia de Recursos Naturales

Secretaría de Investigación y Posgrado

Facultad de Zootecnia y Ecología

Universidad Autónoma de Chihuahua

Presidente: Dr. Alfredo Pinedo Alvarez

La pérdida de los bosques templados de México ha continuado en las últimas décadas a pesar del amplio reconocimiento sobre su importancia para el mantenimiento de la biodiversidad. Este estudio analiza los escenarios de cambio de cobertura de suelo/uso de suelo, utilizando imágenes satelitales del sensor Landsat. Las imágenes corresponden a los años 1990, 2005 y 2017. Los escenarios se aplicaron a los bosques templados con el objetivo de obtener una mejor comprensión de los patrones de cambio de cobertura de suelo/uso de suelo. La técnica de clasificación Support Vector Machine (SVM) sirvió para determinar los tipos de cobertura de suelo/uso de suelo, los cuales fueron posteriormente validados a través del Índice Kappa. Para la simulación de la dinámica de cambio de cobertura de suelo/uso de suelo, se utilizó un modelo desarrollado en Dinamica-EGO, que utiliza modelos estocásticos de las cadenas de Markov, autómatas celulares y pesos de evidencia. Para el estudio se propuso un escenario estacionario, optimista y pesimista. Las proyecciones basadas en los tres escenarios fueron simuladas para el año 2050. Se identificaron y evaluaron cinco tipos de coberturas de suelo/uso de suelo. Estos son: bosques primarios, bosques secundarios, asentamientos humanos, áreas sin vegetación y cuerpos de agua. Los resultados del análisis del cambio de cobertura de suelo/uso de suelo muestran una ganancia sustancial para el bosque secundario. La superficie del bosque primario se redujo del 55.8 % en 1990 al 37.7 % en 2017. Además, los tres escenarios proyectados estiman que las pérdidas adicionales de la superficie son para el bosque primario, especialmente bajo los escenarios estacionario y pesimista. Esto pone de híncapie la importancia y probablemente la urgencia de la aplicación de medidas de conservación y protección para preservar estos ecosistemas y sus servicios. A partir de la precisión obtenida y de los modelos generados, los resultados de estas metodologías pueden servir como una herramienta de decisión para contribuir al manejo sostenible de los recursos naturales de una región.

Palabras clave: Cadenas de markov, escenarios, automatas celulares, escenario pesimista, bosques templados, Dinamica-EGO, pesos de evidencia, cambio de cobertura de suelo/uso de suelo, percepción remota.

INTRODUCTION

Forest ecosystems are important because they provide a wide variety of products and services for the human well being (Hall *et al*., 2006; Fischer and Lindenmayer, 2007; Weiskittel *et al*., 2011) harvested products (Houghton and Nassikas, 2017), carbon sequestration (Hawkes *et al*., 2017), soil retention (Borrelli *et al*., 2017), water supply (Sun *et al*., 2006) and are the habitat of many species of plants and animals. However, antrophongenic activities are the main cause of degradation of almost half of the world surface in the last three centuries. That has caused the loss of lots of our precious natural resources. Twenty-five nations have practically degraded 100 % of their forests, and another 29 nations have degraded 10% of their forest areas (Millennium Ecosystem Assessment, 2005).

Temperate forests represent a key element in the carbon cycle (Pan *et al*., 2011). They are important carbon dioxide sinks (Ma *et al*., 2017), offsetting the emissions produced by the world population (FAO, 2018). Temperate forests store 14 % of the planet's carbon (Pan *et al*., 2011). However, projections of global environmental change show that temperate forests show high vulnerability (Gonzalez *et al*., 2010). This vulnerability can change the productivity of forests by modifying net carbon sequestration rates (Peters *et al*., 2013).

Temperate forests of Mexico occupy 17 % of the national territory, represented by 32 million hectares. In this region, the greatest association of pine and oak forests in the world occurs (González *et al*., 2012). Around 23 different species of pines and close to 200 species of oaks live in the ecoregion of Sierra Madre Occidental (Navar, 2009). However, 40 thousand hectares of forests get on average lost annually. This region has the highest deforestation rate in the world (Velázquez *et al*., 2002; Mas *et al*., 2004).

The study of the land use/land cover changes (LULCC) has become a fundamental research topic, since the change in land use/land cover (LULC) affects forest ecosystems and their biodiversity (Gharun *et al*., 2017). The LULCC, produced by anthropogenic activities have significantly altered the ecosystems biodiversity and services (Butler and Laurance, 2008; Miles and Kapos, 2008; Miranda-Aragón, 2013). The dynamics of LULCC directly affect the landscape patterns, the biogeochemical cycles, the ecosysistems structure and function (Scheffer *et al*., 2001). Recently, the analysis of the spatio-temporal patterns has been the objective of several research studies (Huang *et al*., 2009; Manjarrez-Domínguez *et al*., 2015; Vázquez-Quintero *et al*., 2016). The models of LULCC commonly employed, quantify deforested surfaces, measuring the degree of change in the ecosystem (Lapola *et al*., 2011). Regression methods suchs as the logistic regression have been employed to generate models of LULCC. These models suppose that the relationship between the LULCC and the variables that produce it is a logistic function; however, it has been demonstrated that this relationship is too general (Mas, 2010; Mas, 2014). The dynamics and complexity of the ecosystem requires a more complete evaluation of LULCC. The spatial modeling is a technique contemplating alternative scenarios of LULCC, which could contribute to better explain the key processes influencing LULCC (Pijanowski *et al.,* 2002; Eastman *et al*., 2005; Torrens, 2006; Perez-Vega *et al*., 2012). Thus, one of the main functions of the LULCC models is the establishment of scenarios, with the aim of changing policies and inadequate practices for the sustainable management of natural resources (DeFries *et al*., 2007; Berberoğlu *et al*., 2016).

Several approaches to establish LULCC scenarios have been developed and tested to generate scenarios of LULCC. Ferrerira *et al*. (2012) generated deforestation scenarios to 2050 in the central Brazilian savanna biome finding the possible increase of 13.5 % in deforested areas. Kamusoko *et al*. (2011) evaluated three scenarios (optimistic, pessimistic and business-as-usual) in the Luangprabang province, Lao People's Democratic Republic, finding decreases in forest areas in the pessimistic and business-as-usual scenarios and an increase in forest areas in the optimistic scenario under a strict regulatory policy. Gago-Silva *et al*. (2017) used a combination of Bayesian methods and Weights of Evidence to model the probability of change in a western part of Switzerland. Galford *et al*. (2015) used Bayesian Weights of Evidence for policiy scenarios from 2010 a 2050 evaluating plans for agriculture and forest in Democratic Republic of Congo.

The models to establish reference scenarios of changes in LULCC are based on: systems of equations, statistic models, experts, evolutionary and cellular models, even though there have been efforts to combine plataforms in a multiagent system (Mas *et al*., 2014; Stan *et al*., 2017). The statistical models employ spatial statistics and regression, in comparison with the expert models, which allow the expert knowledge to lead the model path (Parker *et al*., 2003; Soares-Filho *et al*., 2013). The evolutionary or cellular models are very competent to determine the ecologycal alteration; however, they just provide information about the causality or the decision-making (Parker *et al*., 2003).

The generation of LULCC scenarios for the forest region of the state of Chihuahua, Mexico is necessary because of the higher temperate forest deforestation rates in the country. The generation of the LULCC scenario shows two important aspects: expert knowledge and knowledge based on data. Expert knowledge is useful to establish methodological processes according to the needs of the user (Gounaridis *et al*., 2018). Knowledge based on data, helps to understand the general behavior between the factors of change of land use in a spatial way (Camacho-Olmedo *et al*., 2018). Most studies are based on knowledge of the data (Peagelow and Olmedo, 2005; Kityuttachai *et al*., 2013), however, few allow the inclusion of both (Sohares-Filho *et al*., 2006; Olmedo *et al*., 2018).

The Dinamica Environment for Geoprocessing Objects (Dinamica-EGO) is a flexible open platform, which allows analyzing distribution, abundance and spatio-temporal dynamic of the landscape (Soares-Filho *et al*., 2002; Lima *et al*., 2013). The model incorporated to Dinamica-EGO employs cellular automata to simulate the changes happening in a grid, estimating the transition probability, as well as the direction of changes based in stocastic processes (Rutherford *et al*., 2008; Arsanjani *et al*., 2011). Dinamica-EGO allows users to incorporate expert knowledge into the overall statistical analysis based on the spatial data set (Mas *et al*., 2014). In addition, Dinamica-EGO incorporates the possibility of modifying landscape metrics in the calibration procedure to generate the simulation (Mas *et al*., 2012). In a comparative evaluation of approaches to modeling LULCC, two key advantages over Dinamica-EGO were emphasized: 1) incorporation of the Patcher and Expander functions. The first function generates new patches in the landscape and the second expands the previously formed patches, 2) Dinamica-EGO allows the incorporation of multiresolution validation by means of the Fuzzy Similarity Index.

The aim of the present study was (a) to evaluate the change dynamics in the period from 1990 to 2017; (b) to simulate the changes of LULCC for the year 2050 and (c) to elaborate a discussion about the impacts of different scenarios, which could happen in the future in a forest region of the state of Chihuahua, Mexico. Specifically, three scenarios, pessimistic, optimistic and stationary state. The model will identify where the different types of LULCC could hapen. This will allow that future studies could determine changes in carbon sequestration in both, on the surface extension and quantity.

MATERIALS AND METHODS

Study Area

The study area is located in the western part of the state of Chihuahua, Mexico. It is part of the 'Sierra Tarahumara' and have a surface area of 497,159 ha. Its extreme coordinates are 108° 00´ W, 29° 00´N and 107° 10´ W, 27° 30´ N (Figure 1). It is one of the regions of temperate forests, which has experimented the greatest disturbances in the past years in the state of Chihuahua (Herrera, 2002). It belongs to the most extensive forest areas in North America. It is immersed within a complex orography composed of large canyons and deep canyons, which results in a mixture of temperate and tropical ecosystems. It is characterized by its high biodiversity and number of endemic species, estimating the presence of around 4000 species of plants. Also, it is recognized by the International Union for the Conservation of Nature as one of the megacenters of plant diversity (Felger *et al*., 1995). The main land uses in the area include: pine forests, oak forests, pine-oak and oak-pine forest associations, agriculture, and grassland communities. The economic activities in the region are forestry, extensive livestock and rainfed agriculture (INEGI, 2003).

Data Source

For the analysis of the LULCC, three scenes of the Landsat sensor (Path 33, Row 41), with a spatial resolution of 30 m, were used. The scenes corresponded to the years 1990, 2005 and 2017 and they were acquired from clear sky days and each of them taken during the same month to reduce the temporal variation. The scenes were downloaded from the United States

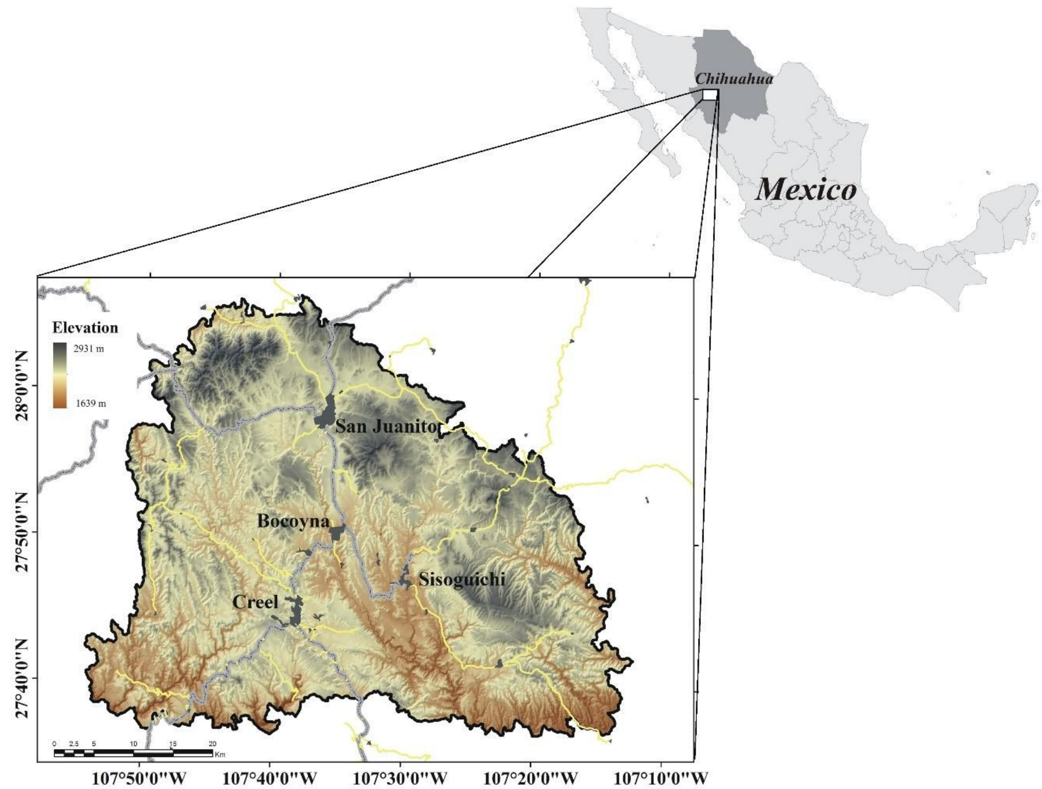

Figure 1. Location and elevations of the study area.

Geological Survey (USGS, 2018). The characteristics of each scene can be seen in Table 1. The scenes were radiometrically corrected. The radiometric correction was carried out with the QGis software 2.8 through the SemiAutomatic Classification plugin (Congedo, 2013).

Integration and Composition of Bands

Once the scenes were corrected, they were integrated into a layer stack. False color composites for the Landsat TM5 were then generated, with a combination of the bands 5, 4 and 3. Band 5 corresponds to the infrared channel (1.55-1.75 μm), band 4 to the near infrared (0.76-0.90 μm) and the band 3 to the red channel (0.63-0.69 μm). This combination was applied to the scenes of 1990 and 2005. Regarding the scene of 2017, the combination for Landsat OLI8 was applied and corresponded to the bands 6, 5 and 4, where band 6 corresponds to the medium infrared channel (1.55-1.65 μm), band 5 to the near infrared channel (0.85-0.88 μm) and band 4 to the red channel (0.64-0.67μm) (Lillesan and Kiefer, 2000).

Land Use and Land Cover Classification

The Suport Vector Machine (SVM) classification was applied to the 1990, 2005 and 2017 images through the software R (R Core Team, 2016) with the package "*caret*" (Kuhn *et al*., 2018) to obtain LULC information. The SVM classifier is a supervised technique of nonparametric statistical methods (Mountrakis and Ogle, 2011). The SVM classification has been used in several research studies in the past (Kavzoglu and Colkesen 2009; Otukei and Blashke 2010; Shao and Lunetta, 2012). For the supervised classification, five classes

Table 1. Scenes characteristics.

Sensor	Date	Characteristics
Landsat TM 5	1990	7 spectral bands, 30 m resolution
Landsat TM 5	2005	7 spectral bands, 30 m resolution
Landsat OLI 8	2017	8 spectral bands, 30 m resolution; 1 panchromatic band 15 m resolution

TM= Thematic Mapper.
OLI= Operational Land Imager.

of land use were defined; 1) primary forest, 2) secondary forest, 3) human settlements, 4) areas without vegetation and (5) water bodies (Table 2).

Modeling and Spatial Simulation with Dinamica-Ego

The LULCC scenarios were made based on the historical trends of change in forest cover during 1990-2017 of the supervised classifications using Dinamica-EGO (Sohares-Filho *et al*., 2002). The historical trends of LULCC is based on the transition matrix (Monteiro *et al*., 2018). Dinamica-EGO uses the algorithm of cellular automata, and the method Weights of Evidence (Camacho-Olmedo *et al*., 2018). For the simulation of deforestation, the following steps were undertaken: 1) selection of change drivers as well as transitions, 2) exploratory analysis of the drivers of deforestation, 3) simulation and 4) validation. These four steps are described in the following sections.

Selection of Variables and Transitions

The selection of the set of exploratory variables to simulate the LULCC is essential for the modeling success (Miranda-Aragón *et al*., 2012; Perez-Vega *et al*., 2016). In this study, 19 variables were used; 17 static and two dynamic variables. Static variables remain constant during model execution. Dynamic variables change during the execution of the model and they are continuously updated in each iteration (Camacho-Olmedo *et al*., 2018). The set of variables used is shown in Table 3.

Table 2. Land use/land cover types determined through the supervised classification method.

Land use and land cover	Acronym	Description
Primary forest	PF	Forest fully covered with canopy
Secondary forest	SF	Forest partially covered with canopy
Human settlements	HS	Residential areas
Areas without vegetation	AWV	Areas without vegetation, agriculture areas or induced grasslands
Water bodies	WB	Water bodies

Table 3. Variables feeding the deforestation model.

No	Variable type	Name	Unit	Acronym
1	Density	Density of main roads	m^2/Km^2	Denmr
2		Density of secondary roads	m^2/Km^2	Densr
3		Density of main streams	m^2/Km^2	Denms
4		Density of secondary streams	m^2/Km^2	Denss
5		Density of rural settlements	m^2/Km^2	Denrs
6	Proximity	Distance to sawmills	m	Diss
7		Distance to water bodies	m	Diswb
8		Distance to main roads	m	Dismr
9		Distance to secondary roads	m	Dissr
10		Distance to main streams	m	Disms
11		Distance to secondary streams	m	Disss
12		Distance to rural settlements	m	Disrs
13		Distance to urban settlements	m	Disus
14		Distance to mines	m	Dism
15		Distance to areas without apparent vegetation	m	Disawav
17	Topographic	Altitude	m	Alt
18		Slope	°	Slop
19		Topographic position index	Dimensionless	TPI

The transition refers to the total amount of LULCC that occurred in the simulation period. In this study, the transitions of interest were: a) primary forest to secondary forest, b) primary forest to areas without apparent vegetation, c) primary forest to urban areas and d) secondary forest to areas without apparent vegetation (Table 4).

Exploratory Analysis of the Data

When we modeled LULCC dynamics, Weights of Evidence (WoE) were applied to project transition probabilities. Regarding deforestation, degradation or any other type of change, we previously know about the location of favorable conditions for LULCC. The influence of static and dynamic variables and the elaboration of the LULC maps was performed with WoE in the Dinamica-EGO software (Soares-Filho *et al*., 2010).

Positive values of WoE represent an attraction between a transition of land use and a specific variable. The greater the value of W+, the greater the probability of transition. Negative values of W- indicate low probabilities of transition instead (Maeda *et al*., 2010). By using the WoE values of the variables used in the analysis of LULCC, the Dinamica-EGO model calculates the transition probability of each pixel to change. Thus, the pixels are assigned with a probability value for a given transition and probability maps are generated for the transitions of interest (Soares-Filho *et al*., 2009 and 2010; Mas and Flamenco, 2011).

Given that the basic hypothesis of the WoE technique is that the driving variables must be independent, for this study the correlation between the variables was tested through the Cramer Coefficient (*V*) showed in Equation 2.

Table 4. Transitions of land use/land cover.

		To				
		PF	SF	HS	AWV	WB
From	PF		✓	✓	✓	
	SF				✓	
	HS					
	AWV					
	WB					

AWV=Areas without vegetation,
SF= Secondary forest,
HS= Human settlements,
WB=Water bodies,
PF= Primary forest

$$V = \sqrt{\frac{\chi^2}{\Gamma \ldots M}} \quad (2)$$

where: χ^2 = is the chi-square statistic of the contingency between two variables, Γ = denotes the sum of the values of contingency, M = is the minimum of *n*-1 or *m*-1, where *n* denotes the number of rows and *m* the number of columns. Bonham-Carter (1994) mentioned that values lower than 0.5 for the Cramer Coefficient (*V*) suggest independence, while values higher than 0.5 involve a greater association (Almeida *et al*., 2003; Teixerira *et al.*, 2009).

Simulation of land use and land cover changes

Three types of scenarios were used for 2050; they were called pessimistic, optimistic and stationary. For the three scenarios, the modeling base was the period 1990-2017. The transition matrix of 1990 and 2017 were used to estimate the possible change in forestry coverage in the future, taking 2017 as the beginning year and 2050 as the final year. In the pessimistic scenario, the transition probability matrix and the change function (*patcher* and *expander*) were modified, increasing the deforestation and fragmentation rates between 1990 and 2017. This was done based on the hypothesis that the development of road infrastructure, urban expansion, fires, uncontrolled exploitation, among others, will produce strong spatial changes of land use. For the optimistic scenario, the state and national forest development plans were considered. Such plans promote the protection and conservation of forest resources (CONAFOR, 2001). For this scenario, the conservation and promotion of strategies to protect forests were represented by reducing the transition matrix value, as well as the patcher and expander change functions. Regarding the stationary scenario, transitions or change functions were not modified. In this case, it is assumed that the trend will be the same as the one between 1990 and 2017.

Validation

To evaluate the model performance, we used a Fuzzy Similarity Index (FSI), where the representation of a pixel is influenced by itself and its neighborhood (Ximenes *et al*., 2011; Yanai *et al*., 2011; Chadid *et al*., 2015). The FSI employed in this study was developed by Hagen (2003), modified by Soares-Filho (2014) and implemented in Dinamica-EGO. The FSI verifies the agreement between the observed and the simulated land use and land cover datasets by obtaining the number of coincident cells within increasing window sizes of a neighborhood (Costanza, 1989; Soares-Filho, 2017). The validation process was carried out by comparing a simulated map and a reference map. The simulation of the 2017 LULCC map was generated. To generate the simulation of 2017, the transition matrix was used between 1990 and 2005. The comparison through the FSI allowed to evaluate the areas of coincidence of change and no change between the real and simulated map of 2017. Finally, the general procedure used in this study is outlined in the flowchart depicted in Figure 2.

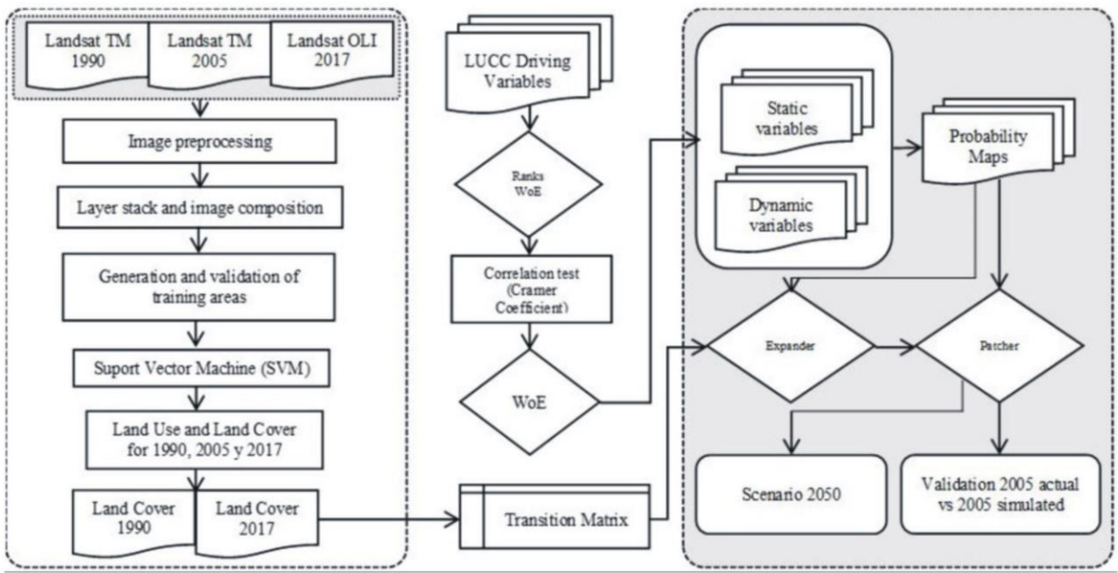

Figure 2. Flowchart of the methodological procedure followed to produce the proposed scenarios. Abbreviations: TM: Tematic Mapper, OLI: Operational Land Imager, WoE: Weights of Evidence, LUCC: Land use and cover change.

RESULTS AND DISCUSSION

Detection of Land Use/Land Cover Changes

Results from the analysis of LULCC show a considerable gain for secondary forest. The forest cover of the primary forest was reduced from 55.8 % of the study area in 1990 to 37.7 % in 2017. The areas without vegetation increased their area from 4.11 % to 4.87 % during 1990-2017 (Table 5). Regarding human settlements and water bodies, they showed a positive trend with an increase from 0.03 % and 0.01 in 1990 to 0.1 % and 0.03 in 2017, respectively. In general, the primary forest was the land use that experimented a negative trend. The rest of the land uses showed surface gains. The rate of change obtained indicate that the secondary forest, the human settlements and the water bodies were the land uses with the greatest transformation rates, with 8.03, 12.58 and 27.48, respectively, for the period of 1990-2017 and with 10.68, 15.96 and 12.3, respectively, from 2005 to 2017. Figure 3 shows the area occupied by the land uses studied. Likewise, it shows the rate of change of these land use/land cover for the periods 1990-2005 and 2005-2017. The calculated global precision, based on the Kappa Index, presented values of 80 %, 85 % and 84 % for 1990, 2005 and 2017, respectively. Table 6 shows the land use/land cover change dynamics. The primary forest lost the greatest surface area (28,406 ha) during 1990-2005, increasing the surface lost to 63,546 ha during 2005-2017. In contrast, the secondary forest showed the largest increases in area with 87,800 ha in the period 1990-2017.

Transition matrix

Table 5. Area occupied for five types of land uses during 1990, 2005 and 2017, and rate of change for the periods 1990-2005 and 2005-2017.

Land Use	Occupied area (%)			Exchange rates	
	1990	2005	2017	1990-2005	2005-2017
AWV	4.11	4.79	4.85	8.33	8.43
SF	40.05	45.05	57.72	8.03	10.68
HS	0.03	0.05	0.10	12.58	15.96
WB	0.01	0.02	0.03	27.48	12.36
PF	55.80	50.08	37.30	6.41	6.21

AWV=Areas without vegetation.
SF= Secondary forest.
HS= Human settlements.
WB=Water bodies.
PF= Primary forest.

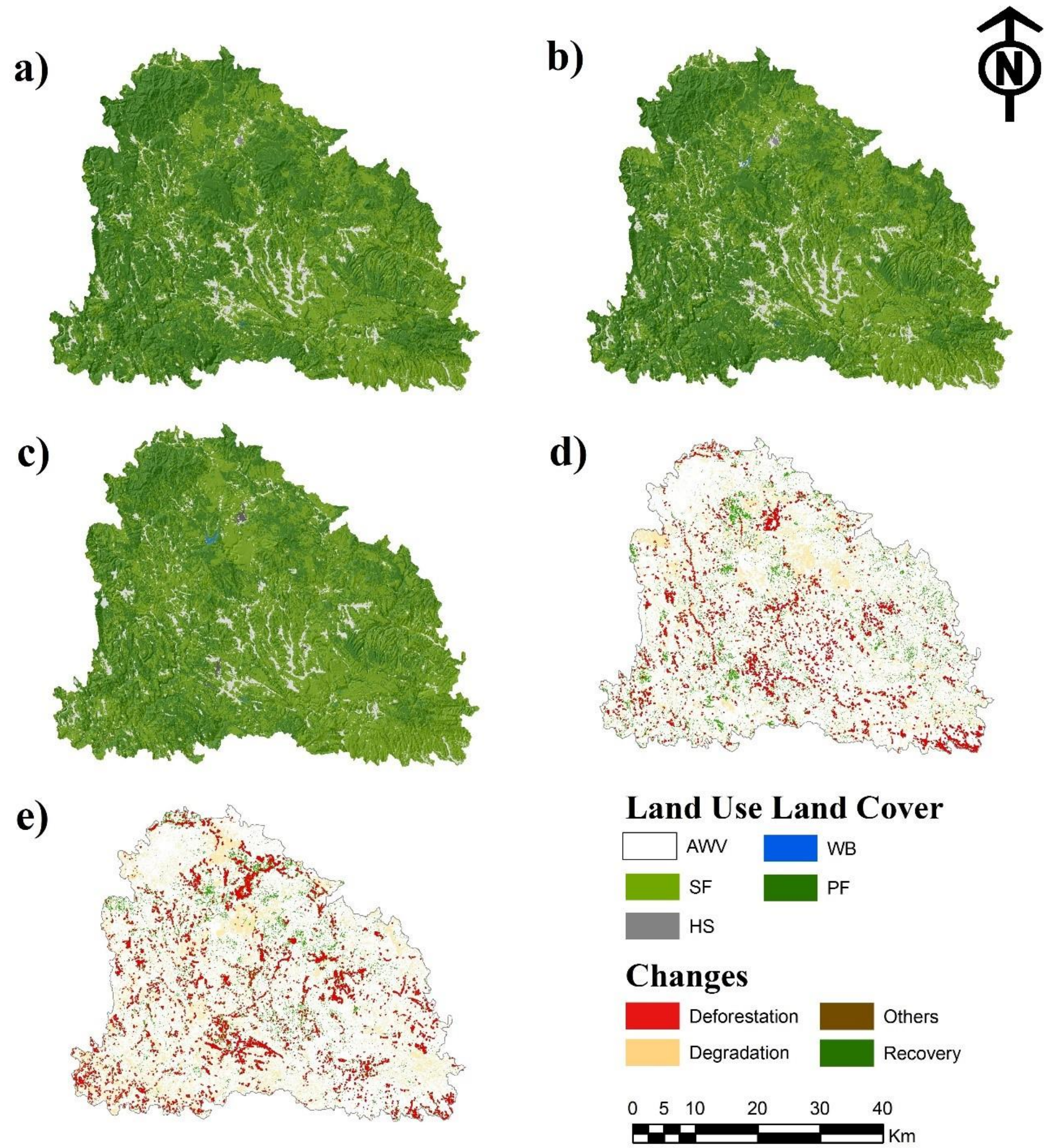

Figure 3. Land use/land cover of 1990 (a), 2005 (b), 2017 (c), changes during 1990-2005 (d) and changes during 2005-2017. Abbreviations: AWV: areas without vegetation, SF: secondary forest, WB: water bodies, HS: human settlements and PF: primary forest.

Table 6. Land use/land cover change dynamics.

Land use	1990-2005	2005-2017	Difference	Type of change	1990-2005	2005-2017
AWV	3384.40	273.34	3657.74	Deforestation	3120.40	7283.95
SF	24826.78	62973.88	87800.66	Degradation	54455.78	73904.27
HS	117.74	249.31	367.05	Other	76.22	219.41
WB	76.79	50.15	126.94	Recovery	27128.71	20204.13
PF	-28406.49	-63546.18	-91952.66	--	--	--

AWV=Areas without vegetation.
SF= Secondary forest.
HS= Human settlements.
WB=Water bodies.
PF= Primary forest.

The transition probabilities of LULCC for the periods 1990-2005 and 2005-2017 are shown in Table 7. The diagonal of the matrix represents the permanence probability, i.e. the probability of a LULC type to remain unchanged. The areas without vegetation showed a 90 % probability of transition from 1990 to 2005, lowering it to 62 % from 2005 to 2017. The areas of primary forest presented a negative trend with a 71 % probability of permanence in the period 1990 to 2005, and changing it to 61 % for the period 2005-2017.

Weights of Evidence (WOE) Analysis

The WoE of the 19 variables were analyzed to eliminate those values that were above 0.5, based on the Cramer Coefficient (*V*). The distance to urban locations showed positive values of WoE from 1000 to 9000 m distance and from 42,000 to 47,000 m indicating an influence for cover change from secondary forest to area without vegetation. The distance to rural localities showed positive values of WoE in distances from 0 to 700 m. The topographic position index showed positive values in the ranges of -150 to -60 and 120 to 240. The distance to sawmills indicates that deforestation appears from 0 to 16,000 m with respect to the process of change between secondary forest to areas without vegetation. The transition from primary forest to area without vegetation is likely to occur in distances to the main roads between 13,000 and 21,000 m. The density of main streams such as rivers and creeks had an influence in densities from 0.039 to 0.079 m^2/ km^2.

Table 7. Transition matrix of probability for land use/land cover change (1990-2005, 2005-2017, 1990-2017).

	Periodo	AWV	PF	HS	WB	PF
	1990-2005	**0.9000**	0.0250	0.0250	0.0250	0.0250
AWV	2005-2017	**0.6250**	0.3504	0.0108	0.0029	0.0109
	1990-2017	**0.6615**	0.3124	0.0120	0.0035	0.0106
	1990-2005	0.0222	**0.7516**	0.0008	0.0005	0.2248
SF	2005-2017	0.0557	**0.8116**	0.0004	0.0000	0.1323
	1990-2017	0.0654	**0.7945**	0.0012	0.0006	0.1384
	1990-2005	0.0452	0.0645	**0.8806**	0.0000	0.0097
HS	2005-2017	0.0557	0.2479	**0.6959**	0.0000	0.0004
	1990-2017	0.0651	0.0774	**0.8575**	0.0000	0.0000
	1990-2005	0.0000	0.1254	0.0000	**0.8553**	0.0193
WB	2005-2017	0.0095	0.1684	0.0000	**0.8030**	0.0191
	1990-2017	0.0000	0.1868	0.0000	**0.7957**	0.0175
	1990-2005	0.0020	0.2865	0.0000	0.0000	**0.7115**
PF	2005-2017	0.0056	0.3798	0.0003	0.0000	**0.6144**
	1990-2017	0.0071	0.4419	0.0002	0.0000	**0.5508**

AWV=Areas without vegetation, SF= Secondary forest, HS= Human settlements, WB=Water bodies, PF= Primary forest.
Bold letters the probability of permanence on the same class.

In the transition from primary forest to secondary forest, the variable altitude showed positive values of WoE in the range of 1,200-1,300 m, suggesting that most of the changes occur in this range. The slope showed that the process of change between primary forest and secondary forest is located on slopes of 45-60 and 60-75 degrees. The transition from primary forest to human settlements was influenced by the distance to secondary streams from 500 to 1000 meters. The distance to sawmills presented an influence from 0 to 6,000 meters. The distance to mines showed that the attraction to change occurs between 2000 and 10,000 m.

Model Validation

The model validation was carried with the simulated and the true land use classification of 2017. The FSI was applied for neighborhoods from 1×1 to 7×7 pixels. The minimum value reported for FSI was 49 % in 1×1 pixel, while in 7×7 pixels the value of FSI was 91 %. These results indicate that the real and simulated land use changes agree from 49 % to 91 %. Simulation starts with 49 % and adjusts to 91 %, reaching a similarity adjustment value at a distance of 210 m. These results agree with that obtained by Ximenes *et al*. (2011). According to Soares-Filho (2017), and similar studies (Carlson *et al*., 2012; de Rezende *et al*., 2015; Elz *et al*., 2015), for the resolution and the number of transitions considered in the model, the values obtained for the FSI suggest that the models are good and can be used in the simulation of LULCC scenarios. Figure 4 represents the FSI in relation to the size of the window.

Scenarios

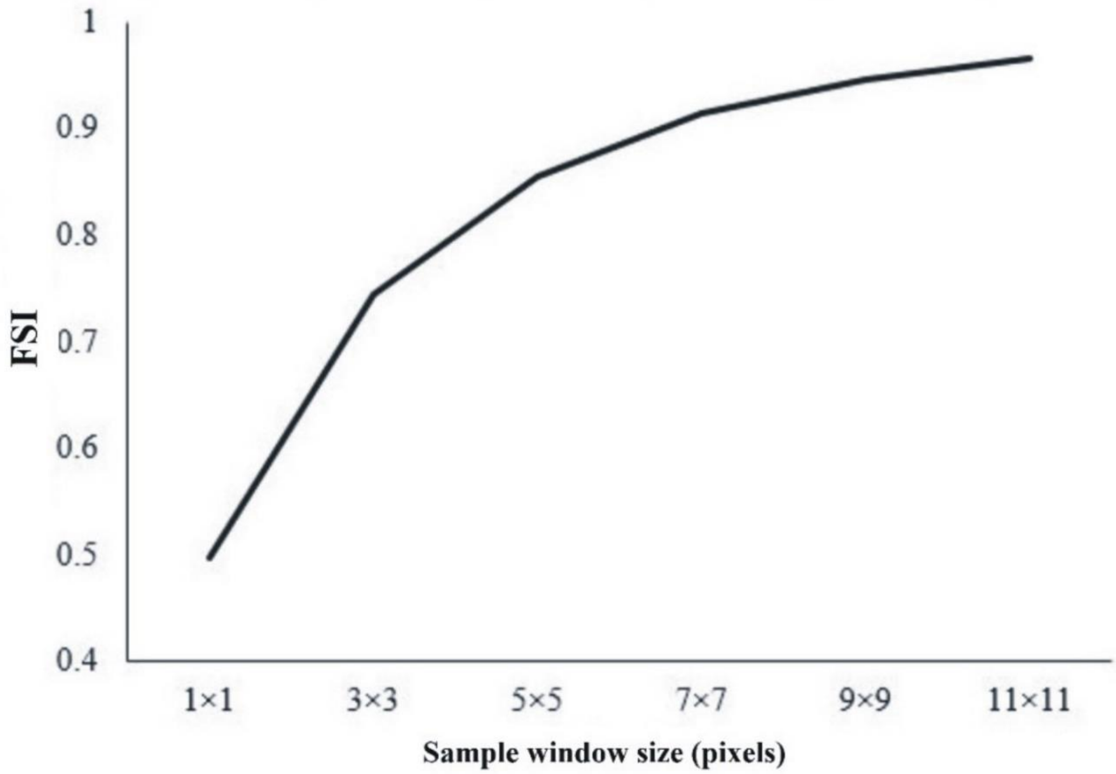

Figure 4. Variation of the FSI as a function of different window sizes.

The LULCC based on the transitions between 1990 and 2017 for the stationary, optimistic and pessimistic scenarios are presented in Table 8. Figure 5 shows the LULC classification of 2017 and the stationary, optimistic and pessimistic scenarios for 2050, after the model calibration.

In the stationary scenario the area without vegetation would increase from 4.8 % in 2017 to 5.27 %. Likewise, the secondary forest would increase from 57.7 % (2017) to 73 %. For this scenario, the changes in human settlement and water bodies would not increase or reduce their area. Conversely, the rate of change of primary forest and secondary forest were the greatest between 2017 and 2050. Regarding the optimistic scenario, it showed reductions in areas of primary forest; however, in lower magnitudes than for the stationary and pessimistic scenarios. For the pessimistic scenario, the Markov matrix was modified considering a greater pressure on the forest ecosystem. The area without vegetation showed a positive trend, with 4.8 % in 2017 and an increase to almost 8 % in 2050. The secondary forest would go from 57.7% to 85.6 % in 2050. Finally, the primary forest would reduce its area to an 8 % and isolated forest areas would appear. The rate of change for this scenario were the ones that showed the highest values. The LULCC dynamics projected for 2050 for the three scenarios (stationary, optimistic, pessimistic) is presented in Table 9.

In this study, scenarios of LULCC for 2017 and 2050 were generated for a temperate forest region of Chihuahua Mexico. The scenarios were developed in Dinamica-EGO. Results were consistent with the results described by Maeda *et al*. (2011). For the generation of transitions and simulation of scenarios, LULC of 1990, 2005 and 2017 were determined. In general, proximity to sources with

Table 8. Percentage of surface area occupied by five land use/land cover types and rate of change for 2017-2050 based on three scenarios.

Land Use	Occupied surface area (%)				Change rate		
	2017	2050s	2050o	2050p	2017-2050s	2017-2050o	2017-2050p
AWV	4.848	5.275	5.017	7.695	3.40	3.23	4.96
SF	57.716	73.721	61.863	83.628	3.99	3.35	4.53
HS	0.105	0.105	0.105	0.105	3.13	3.13	3.13
WB	0.031	0.031	0.031	0.031	3.12	3.13	3.12
PF	37.300	20.868	32.983	8.541	1.75	2.76	0.72

AWV=Areas without vegetation,
SF= Secondary forest,
HS= Human settlements,
WB=Water bodies,
PF= Primary forest

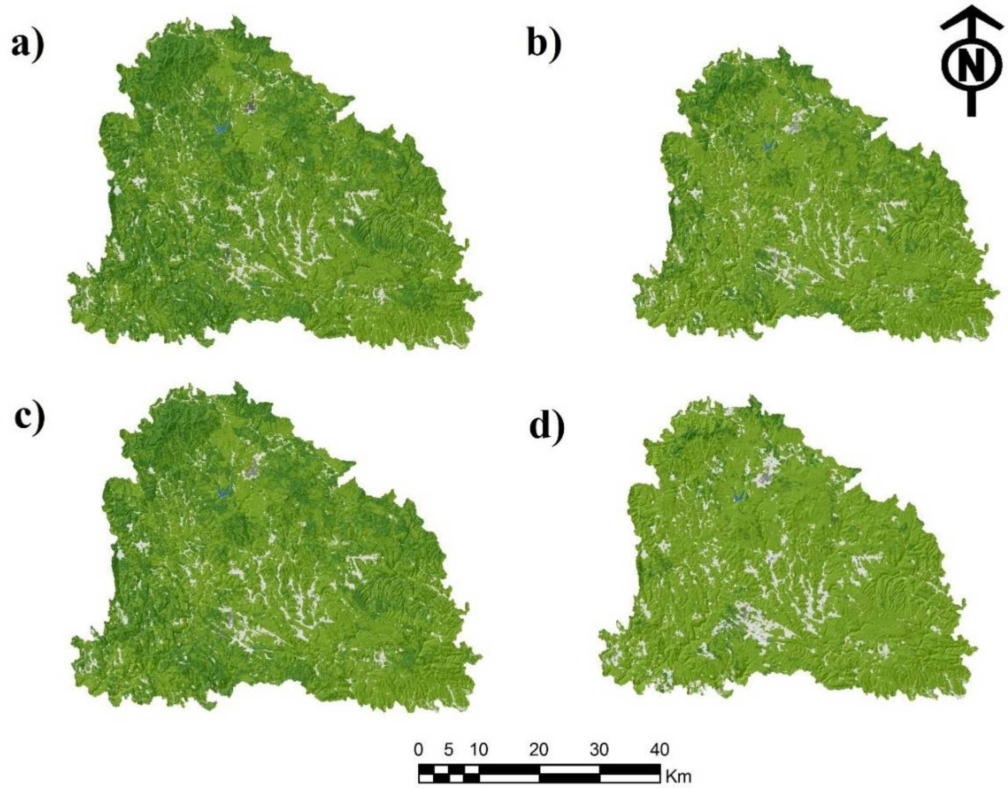

Figure 5. a) Land use/land cover of 2017 and simulated land use/land cover projected for the year 2050 as a result of the b) Stationary, c) Pessimistic and d) Optimistic scenarios. ■ areas without vegetation, ■ secondary forest, ■water bodies, ■ human Settlements and ■primary forest.

Table 9. Land use/land cover change dynamics (ha) under three projected scenarios.

Land use	$2017\text{-}2050_s$	$2017\text{-}2050_o$	$2017\text{-}2050_p$
AWV	2121.97	840.44	14150.28
SF	79565.57	20617.02	128818.00
HS	0.87	1.25	1.05
WB	0.46	0.32	0.10
PF	-81688.00	-21459.04	-142969.31

AWV=Areas without vegetation.
SF= Secondary forest.
HS= Human settlements.
WB=Water bodies.
PF= Primary forest.

anthropogenic activity as well as topography were important factors influencing the change in forest cover. The exchange between primary forest and secondary forest represented the main transition between 1990 and 2017. This transition produced the greatest impact, in agreement with the results reported by Perez-Vega *et al*. (2016). Such transition was influenced by the altitude, slope, and density of water streams, in agreement with the results of Armenteras *et al*. (2006) and Chadid *et al*. (2015). The transition from primary to secondary forest could be attributed to the reduction in pine vegetation, where shrubs would become dominant. A consequence of the reduction of primary forest is the migration of fauna, which deals with the dispersal of the seeds of large-crowned trees (Lehouck *et al*., 2009). Other consequences include the change of lands to livestock production systems (Maeda *et al*., 2010) and the presence of areas with high solar incidence and low coverage, which are prone to fires (de Rezende *et al*., 2015). Another reason for the reduction of primary forest is the proximity to urban rural localities and roads, which is in agreement with the results reported by Aguiar *et al*. (2007) and Osorio *et al*. (2015). The proximity to urban and rural communities indicates the possible extraction of wood for export and also facilitates the expansion of the agricultural or grazing frontier (Chadid *et al*., 2015). This can be verified by the number of sawmills in the study area. The process of deforestation/degradation is strongly related to this cause. In the forested areas of Chihuahua, the rural localities are in a high degree of marginalization (González *et al*., 2012) where there exist agricultural incentives PEF 2025 (CONAFOR, 2001), causing the possible increase of the areas without vegetation. Another reason for the degradation may be the distance to the main roads and the topographic position.

The results obtained for the different scenarios showed differences among the surfaces of land use. The stationary scenario resulted in a considerable change in the primary forest, mainly. This scenario considers that the transition values among land use coverages will continue. The long-term impacts of the deforestation/degradation process include increased reservoir sedimentation and decreased flows in the dry season (Gingrich, 1993). Although the optimistic scenario showed increases in non-forested areas, this scenario was the one that showed the greatest resistance for the transitions from primary forest to any other LULC. This scenario considers the strict application of the regulation of forest resources, in agreement with the general trend in the protection of forest ecosystems to degradation (UN, 2015) and the projections of the PEF 2025 (PEF, 2001). The pessimistic scenario showed the greatest losses in the coverage of the primary forest. In addition, the increase in areas without vegetation, which is mainly associated to cropping and the proximity to water currents, is one of the main outputs of the pessimistic scenario, which agrees with the study by Elz *et al*. (2015). The increase in agricultural areas resulting from this scenario may benefit the inhabitants economically; however, the expansion of this type of land use/land cover could lead to a greater demand of water for irrigation purposes, which could potentially impact water resources (Maeda *et al*., 2010).

Population growth (Barni *et al*., 2015), the market demand and the lack of technification for wood processing cause the opening of land and the extraction of wood for self-consumption. Taking these aspects into account, the simulation of changes in forest cover indicates pressure on forest resources, which is consistent with that found by Kamusko *et al*., (2011). As a consequence, forest degradation could lead to soil loss (Quan *et al*., 2011), loss in biodiversity (Falcucci *et al*., 2007) and landscape connectivity (Tambosi *et al*., 2014), habitat fragmentation (Nagendra *et al*., 2004), the presence of invasive species (Mas *et al*., 2012), among others.

The LULCC model of this study incorporated the Markov chains, Cellular Automata and WoE methods. Several transitions were simulated as in the studies by Soares-Filho *et al*., (2010), Ferreira *et al*. (2013) and Elz *et al*. (2015). The validation was carried out based on the FSI, as it was also performed in previous research (Ximenes *et al*., 2011). The result of this analysis, where the three aforementioned methods are combined, highlighted the variables driving the process of degradation/deforestation, as well as the manipulation based on the knowledge of the transition probabilities, being more suitable for the simulation of LULCC (Mas and Flamenco, 2011). The transition probability matrices revealed that the primary forest has a negative trend in its occupied area, suggestsing that degradation will continue over this land use, this area of primary forest changed to secondary forest. Although the other transitions did not produce important changes in the spatial configuration of the landscape, but their cumulative long-term effect could negatively impact the functioning of the ecosystems and their biodiversity (Pompa, 2008).

In this study, we focused on hypothetical scenarios where the pressure of forest resources was controlled by changing the transition probability. However, it is necessary to study scenarios where market demand (Merry *et al.,* 2009) or illegal timber extraction (Chadid *et al*., 2015) is considered. The wood clandestinage corresponds to 30 % in some forest management units of Chihuahua (Silva, 2009).

The scenarios are not exact projections of the future state of the environment (Feng and Liu, 2016). However, it is an alternative means of supporting forest managers, which can serve as a valuable tool for studying political decisions (Kolb and Galicia, 2018). That would lead to a better knowledge of forest exploitation and protection. Managers can take into account the proposed scenarios and take decisions based on the one with the most promising results.

Due to the distribution of economic information (municipality based) and the lack of information from georeferenced illicit extractions, we believe an approach such as agent-based models would help to improve the study and address these issues. Finally, the model did not consider climatic variations such as precipitation and temperature, which can affect patterns and dynamics in recovery zones. That should be implemented in future studies.

CONCLUSIONS AND RECOMMENDATIONS

The use of scenarios as a methodology to study LULCC has been studied in depth at different scales and in different areas. However, several improvements can be implemented. This study presents an approach that integrates expert knowledge, and geospatial technologies such as geographic information systems and spatial simulation. The developed scenarios were based on the application of the forestry law (non-spatially) as well as the state of the landscape, and not only on the extrapolation of past trends. In addition, the scenarios are spatially explicit, which allow identifying the spatial pattern of change and the possible critical areas of change in forest cover. Finally, this study contributes to the understanding of the future fragmentation of the forest cover. Therefore, the current decisions in the field of forest management and land use/land cover influence the future of our forests and can probably be represented in one of the three proposed scenarios.

REFERENCES

Aguiar A.P.D., G. Câmara, and M.I.S. Escada, 2007. Spatial statistical analysis of land-use determinants in the Brazilian Amazonia: Exploring intra-regional heterogeneity. Ecol. Model. 209:169-188.

Almeida C.M., M. Batty, A.M.V. Monteiro, G. Camara, B.S. Soares-Filho, G.C. Cerqueira, and C.L. Pennachin, 2003. Stochastic cellular automata modelling of urban land use dynamics: Empirical development and estimation. Computers, Environment and Urban Systems 27:481-509.

Armenteras, D., G. Rudas, N. Rodriguez, S. Sua, and M. Romero 2006. Patterns and causes of deforestation in the Colombian Amazon. Ecological Indicators 6:353-368.

Arsanjani J.J., M. Helbich, W. Kainz, and A.D. Boloorani. 2013. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. International Journal of Applied Earth Observation and Geoinformation, 21:265-275.

Barni P.E., P.M. Fearnside, and P.M.L. de Alencastro Graça 2015. Simulating deforestation and carbon loss in Amazonia: impacts in Brazil's Roraima state from reconstructing Highway BR-319 (Manaus-Porto Velho). Environmental management, 55:259-278.

Berberoğlu, S., Akın, A., and Clarke, K.C. 2016. Cellular automata modeling approaches to forecast urban growth for adana, Turkey: A comparative approach. Landscape and Urban Planning 153:11-27.

Bonham-Carter G. 1994. Geographic Information Systems for Geoscientists: Modelling with GIS. New York: Pergamon.

Borrelli P., D.A. Robinson, L.R. Fleischer, E. Lugato, C. Ballabio, C. Alewell, K. Meusburger, S. Modungo, B. Schütt, V. Ferro, V. Bagarello, K. Oost, L. Montanarella, and P. Panagos 2017. An assessment of the global impact of 21st century land use change on soil erosion. Nature communications 8:2013.

Butler, R.A., and W.F. Laurance 2008. New strategies for conserving tropical forests. Trends in Ecology & Evolution 23:469-472.

Carlson K.M., L.M. Curran, D. Ratnasari, A.M. Pittman, B.S. Soares-Filho, G.P. Asner, D. Lawrence, and H.O. Rodrigues 2012. Committed carbon emissions, deforestation, and community land conversion from oil palm plantation expansion in West Kalimantan, Indonesia. Proceedings of the National Academy of Sciences 109:7559-7564

Chadid M.A., L.M. Dávalos, J. Molina, and D. Armenteras 2015. A Bayesian spatial model highlights distinct dynamic in deforestation from coca and pastures in an Andean biodiversity hotspot. Forests, 6:3828-3846

Cominsión Naciaonl Forestal (CONAFOR). 2001. Programa Estratégico Forestal 2025. Available at http://www.conafor.gob.mx:8080/documentos/ver.aspx?articulo=307&grupo=4.

Congedo, L. 2017. Semi-Automatic Classification Plugin for QGIS, Technical Report, Sapienza University, ACC Dar Project: Rome, Italy. Available at https://media.readthedocs.org/pdf/semiautomaticclassificationmanual-v3/latest/semiautomaticclassificationmanual-v3.pdf

Costanza R. 1989. Model goodness of fit: a multiple resolution procedure. Ecological Modelling 47:199–215.

de Rezende C.L., A. Uezu, F.R. Scarano, and D.S.D. Araujo 2015. Atlantic Forest spontaneous regeneration at landscape scale. Biodiversity and conservation 24: 2255-2272.

DeFries R., A. Hansen, B.L. Turner, R. Reid, J. Liu. 2007. Land use change around protected areas: management to balance human needs and ecological function. Ecological Applications 17:1031-1038.

Eastman J.R., L.A. Solorzano, and M.E. Van Fossen 2005. Transition potential modeling for land-cover change In: GIS, spatial analysis, and modeling, eds. Maguire D Batty JM, Goodchild MF. ESRI Press, California.

Elz, I., K. Tansey, S.E Page, and M. Trivedi, 2015. Modelling deforestation and land cover transitions of tropical peatlands in Sumatra, Indonesia using remote sensed land cover data sets. Land, 4:670-687

Falcucci, A., L. Maiorano, and L. Boitani 2007. Changes in land-use/land-cover patterns in Italy and their implications for biodiversity conservation. Landscape ecology 22:617-631.

FAO. 2018. The state of world's forests 2018 - Forest pathways to sustainable development. FAO, Rome. Available at http://www.fao.org/publications/sofo/en/ (Accessed 7 October 2018).

Felger, R.S., and M.F. Wilson 1995. Northern Sierra Madre Occidental and its Apachian outliers: a neglected center of biodiversity. Biodiversity and Management of the Madrean Archipelago: The Sky Islands of Southwestern United States and Northwestern Mexico, 36-51.

Ferreira, M.E., L.G. Ferreira Jr., F. Miziara, and B.S. Soares-Filho. 2013. Modeling landscape dynamics in the central Brazilian savanna biome: future scenarios and perspectives for conservation. Journal of Land Use Science 8:403-421.

Fischer, J., and D.B. Lindenmayer 2007. Landscape modification and habitat fragmentation: a synthesis. Global ecology and biogeography 16:265-280.

Gago-Silva, A., N. Ray, and A. Lehmann 2017. Spatial Dynamic Modelling of Future Scenarios of Land Use Change in Vaud and Valais, Western Switzerland. ISPRS International Journal of Geo-Information 6:115

Galford, G.L., B.S. Soares-Filho, L.J. Sonter, and N. Laporte 2015. Will passive protection save Congo forests? PloS one, 10: e0128473.

Gharun, M., M. Possell, T.L. Bell, and M.A. Adams. 2017. Optimisation of fuel reduction burning regimes for carbon, water and vegetation outcomes. Journal of environmental management 203:157-170.

Gingrich, R.W. 1993. The political ecology of deforestation in the Sierra Madre Occidental of Chihuahua (Master's thesis, University of Arizona).

González, B.A. 2012. La Sierra Tarahumara, el bosque y los pueblos originarios: estudio de caso de Chihuahua México. Available at: http://www.fao.org/forestry/17194-0381f923a6bc236aa91ecf614d92e12e0.pdf

Gonzalez, P., R.P. Neilson, J.M. Lenihan, and R.J. Drapek. 2010. Global patterns in the vulnerability of ecosystems to vegetation shifts due to climate change. Global Ecology and Biogeography 19:755-768

Gounaridis, D., I. Chorianopoulos, and S. Koukoulas. 2018. Exploring prospective urban growth trends under different economic outlooks and land-use planning scenarios: the case of Athens. Applied Geography 90:134-144

Hall, R.J., R.S. Skakun, E.J. Arsenault, and B.S. Case 2006. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. Forest ecology and management 225:378-390.

Hawkes, C.V., B.G. Waring, J.D. Rocca, and S.N. Kivlin. 2017. Historical climate controls soil respiration responses to current soil moisture. Proceedings of the National Academy of Sciences 114:6322-6327.

Herrera, A. 2002. Situación actual de los bosques de Chihuahua. Madera y bosques 8.

Houghton, R.A., and A.A. Nassikas, 2017. Global and regional fluxes of carbon from land use and land cover change 1850–2015. Global Biogeochemical Cycles 31:456-472

Huang, B., L.Z hang, and B. Wu. 2009. Spatiotemporal analysis of rural–urban land conversion. International Journal of Geographical Information Science 23:379-398

Instituto Nacional de Estadística, Geografía e Informática (INEGI). 2003. Síntesis de Información Geográfica del Estado de Chihuahua.

Kamusoko, C., K. Oono, A. Nakazawa, Y. Wada, R. Nakada, T. Hosokawa, T. Someya. 2011. Spatial simulation modelling of future forest cover change scenarios in Luangprabang province, Lao PDR. Forests 2:707-729.

Kavzoglu, T, and I. Colkesen 2009. A kernel functions analysis for support vector machines for land cover classification. International Journal of Applied Earth Observation and Geoinformation, 11:352-359.

Kityuttachai, K., N.K. Tripathi, T. Tipdecho, and R. Shrestha 2013. CA-Markov analysis of constrained coastal urban growth modeling: Hua Hin seaside city, Thailand. Sustainability, 5:1480-1500.

Kuhn, M., J. Wing, and S. Weston, 2015. Package ‘caret’. Classification and regression training.

Lapola, D., M.R. Schaldach, J. Alcamo, A. Bondeau, S. Msangi, J.A. Priess, and B.S. Soares-Filho 2011. Impacts of climate change and the end of deforestation on land use in the Brazilian Legal Amazon. Earth Interactions. 15:1-29.

Lehouck, V., T. Spanhove, L. Colson, A. Adringa-Davis, N.J. Cordeiro, and L. Lens 2009. Habitat disturbance reduces seed dispersal of a forest interior tree in a fragmented African cloud forest. Oikos, 118:1023-1034

Lillesand, T.M., and Kiefer R.W. 2000. Remote Sensing and Image Interpretation, 4th ed.; John Wiley and Sons New York, NY, USA.

Lima, T.C., C.M Guilen-Lima, M.S. Oliveira, and B. Soares-Filho 2013. DINAMICA EGO e Land Change Modeler para simulação de desmatamento na Amazônia brasileira: análise comparativa. In: Anais XVI Simpósio Brasileiro de Sensoriamento Remoto: Foz do Iguaçu, INPE, 6379-6386.

Ma, W., G. Jia, and A. Zhang. 2017. Multiple satellite-based analysis reveals complex climate effects of temperate forests and related energy budget. Journal of Geophysical Research: Atmospheres, 122: 3806-3820

Maeda, E.E., B. Clark, P. Pellikka, and M. Siljander 2010. Driving forces of land-use change in the Taita Hills, Kenya. In: 13th AGILE International Conference on Geographic Information Science, 2-5.

Maeda, E.E., De C.M. Almeida, A. de Carvalho Ximenes, A.R. Formaggio, Y.E. Shimabukuro, and P. Pellikka 2011. Dynamic modeling of forest conversion: Simulation of past and future scenarios of rural activities expansion in the fringes of the Xingu National Park, Brazilian Amazon. International Journal of Applied Earth Observation and Geoinformation 13:435-446

Manjarrez-Dominguez, C., A. Pinedo-Alvarez, C. Pinedo-Alvarez, F. Villarreal-Guerrero, and L. Cortes-Palacios 2015. Vegetation landscape analysis due to land use changes on arid lands. Polish Journal of Ecology 63:167-174

Mas, J.F, M. Kolb, T. Houet, M. Paegelow, and M.C. Olmedo. 2010. Una comparación de diferentes enfoques de modelación de cambios de cobertura/uso del suelo. In: Proceedings of the XIV Simposio Internacional SELPER 2010.

Mas J.F., Flamenco, A. 2011. Modelación de los cambios de coberturas/uso del suelo en una región tropical de México. GeoTrópico 5:1-24.

Mas J.F., M. Kolb, M. Paegelow, M.T.C. Olmedo, and T. Houet. 2014. Inductive pattern-based land use/cover change models: A comparison of four software packages. Environmental Modelling & Software 51: 94-111.

Mas J.F., A. Pérez-Vega, K.C. Clarke 2012. Assessing simulated land use/cover maps using similarity and fragmentation indices. Ecological Complexity 11: 38-45.

Mas, J.F., A. Velázquez, J.R. Díaz-Gallegos, R. Mayorga-Saucedo, C. Alcántara G. Bocco, R. Castro, T. Fernandez, and Pérez-Vega A. 2004. Assessing land use/cover changes: a nationwide multidate spatial database for Mexico. International Journal of Applied Earth Observation and Geoinformation 5:249-261.

Merry, F., B.S. Soares-Filho, D. Nepstad, G. Aamacher, and H. Rodrigues 2009. Balancing, conservation and economic sustainability: the future of the amazon timber industry. Environmental Management EUA 44:395–407

Miles, L., Kapos, V. 2008. Reducing greenhouse gas emissions from deforestation and forest degradation: global land-use implications. Science 320:1454-1455.

Millennium Ecosystem Assessment. 2005. Ecosystems and Human Well-Being: Synthesis. Island Press, Washington, DC.

Miranda-Aragón, L., E.J. Treviño-Garza, J. Jiménez-Pérez, O.A. Aguirre-Calderón, M.A. González-Tagle, M. Pompa-García, C.A. Aguirre-Salado 2012. Modeling susceptibility to deforestation of remaining ecosystems in North Central Mexico with logistic regression. Journal of forestry research 23:345-354.

Monteiro Junior, J.J., E. Silva, A.L. De Amorim Reis, and J.P. Mesquita Souza Santos. 2018. Dynamical spatial modeling to simulate the forest scenario in Brazilian dry forest landscapes. Geology, Ecology, and Landscapes, 1:7.

Mountrakisl G. J. Im, and C. Ogole 2006. Support vector machines in remote sensing: A review. ISPRS J. Photogramm. Remote Sens. 2011, 66, 247–259.

Nagendra, H., D.K. Munroe, and J. Southworth. 2004. From pattern to process: landscape fragmentation and the analysis of land use/land cover change. Agriculture, Ecosystems & Environment, 101:111-115.

Navar J. 2009. Allometric equations for tree species and carbon stocks for forests of northwestern Mexico. Forest ecology and Management. 257:427-434.

Olmedo M.T.C. M. Paegelow, J.F. Mas, and F. Escobar 2018. Geomatic Approaches for Modeling Land Change Scenarios. Switzerland: Springer.

Osorio L.P, Mas J.F, Guerra F, Maass M. 2015. Análisis y modelación de los procesos de deforestación: un caso de estudio en la cuenca del río Coyuquilla, Guerrero, México. Investigaciones geográficas (88): 60-74

Otukei J.R., Blaschke T. 2010. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. International Journal of Applied Earth Observation and Geoinformation 12: S27-S31

Paegelow M, Olmedo, MTC. 2005. Possibilities and limits of prospective GIS land cover modelling—a compared case study: Garrotxes (France) and Alta Alpujarra Granadina (Spain). International Journal of Geographical Information Science 19:697-722.

Pan Y, Birdsey RA., Fang J, Houghton R, Kauppi PE, Kurz WA, Oliver L, Shvidenko PA, Lewis SL, Canadell JG, Ciais, P, Jackson RB, Pacala S, McGuire AD, Piao s, Rautiainen A, Sitch s, Hayes D. 2011. A large and persistent carbon sink in the world's forests. Science, 1201609

Parker D, Manson S, Jansen M, Hoffman M, Deadman P. 2003. Multi-Agent systems for the simulation of land-Use and land-Cover change: a review. Annals of the American Association of Geographers 93: 314–337.

Pérez-Vega A, Mas JF, Ligmann-Zielinska A. 2012. Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. Environmental Modelling & Software 29: 11-23

Pérez-Vega A, Rocha Álvarez F, Regil García HH. 2016. Distribución espacial del uso/cubierta del suelo y degradación forestal en la reserva de la biosfera Sierra Gorda de Guanajuato. Acta Universitaria 26(2)

Peters EB, Wythers KR, Bradford JB, Reich PB. 2013. Influence of disturbance on temperate forest productivity. Ecosystems 16(1):95-110

Pijanowski BC, Brown DG, Shellito BA, Manik, GA. 2002. Using neural networks and GIS to forecast land use changes: a land Transformation Model. Computers, Environment and Urban Systems 26:553–575.

Pompa M. 2008. Análisis de la deforestación en ecosistemas montañosos del noroeste de México. Avances en Investigación Agropecuaria, 12(2).

Quan B, Römkens MJM, Li R, Wang F, Chen J. 2011. Effect of land use and land cover change on soil erosion and the spatio-temporal variation in Liupan Mountain Region, southern Ningxia, China. Frontiers of Environmental Science & Engineering in China 5(4): 564-572.

R Core Team. 2016. R: A language and environment forstatistical computing. Vienna: R Foundation for Statistical Computing.

Rutherford GN, Bebi P, Edwards PJ, Zimmermann NE. 2008. Assessing land-use statistics to model land cover change in a mountainous landscape in the European Alps. Ecological modelling 212(3-4):460-471

Scheffer M, Carpenter S, Foley JA, Folke C, Walker B. 2001. Catastrophic shifts in ecosystems. Nature 413:591.

Shao Y, Lunetta RS. 2012. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. ISPRS Journal of Photogrammetry and Remote Sensing 70:78-87.

Silva RS. 2009. Estudio Regional Forestal “Unidad de Manejo Forestal San Juanito A.C. UMAFOR San Juanito A. C.

Soares-Filho B, Rodrigues H, Follador M. 2013. A hybrid analytical-heuristic method for calibrating land-use change models. Environmental Modelling & Software 43: 80-87.

Soares-Filho BS, Cerqueira GC, Pennachin CL, 2002. DINAMICA — a stochastic cellular automata model designed to simulate the landscape dynamics in an Amazonian colonization frontier. Ecological. Modelling 154: 217–235.

Soares-Filho BS, Ferreira BM, Filgueira DS; Rodrigues HO, Hissa LBV, Lima LS, Machado RF, Costa WLS. Dinamica project. Remote Sensing Center. Federal University of Minas Gerais (UFMG), Belo Horizonte, MG, Brazil. http://www.csr.ufmg.br/dinamica/. Accessed 24 Jul 2017.

Soares-Filho BS, Moutinhob P, Nepstadb D, Andersond A, Rodriguesa H, Garciaa R, 2010. Role of Brazilian amozon protected areas in climate change mitigation. Proceedings of the National Academy of the United States of America 107(24): 10821-10826;

Soares-Filho BS, Nepstad DC, Curran LM, Cerqueira GC, Garcia RA, Ramos CA, Voll E, Mcdonald A, Lefebvre P, Schlesinger, P. 2006. Modelling conservation in the Amazon basin. Nature 440:520

Soares-Filho BS, Rodrigues H, Costa, W. 2009. Modeling Environmental Dynamics with Dinamica Ego. Available at:

http://www.lapa.ufscar.br/geotecnologias-1/Dinamica_EGO_guidebook.pdf

Stan K D, Sanchez-Azofeifa A. 2017. The Edmonton–Calgary corridor: Simulating future land cover change under potential government intervention. Land Use Policy 63:356-368.

Sun G, Zhou G, Zhang Z, Wei X, McNulty SG, Vose JM. 2006. Potential water yield reduction due to forestation across China. Journal of Hydrology 328(3-4):548-558.

Tambosi LR, Martensen AC, Ribeiro MC, Metzger JP. 2014. A framework to optimize biodiversity restoration efforts based on habitat amount and landscape connectivity. Restoration Ecology 22(2): 169-177.

Texeira AM, Soares-eFilho BS, Freitas S, Metzger JPW. 2009. Modeling landscape dynamics in the Atlantic Rainforest domain: implications for conservation. Forest Ecology and Management 257:1219e1230.

Torrens PM. 2006. Geosimulation and its application to urban growth modeling. Springer-Verlag: London, U. K.

United Nations (UN) 2015. Transforming our World: The 2030 Agenda for Sustainable Development Available at https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf

United States Geological Survey. Available at: http://glovis.usgs.gov.

Vázquez-Quintero G, Solís-Moreno R, Pompa-García M, Villarreal-Guerrero F, Pinedo-Alvarez C, Pinedo-Alvarez A. 2016. Detection and projection of forest changes by using the Markov Chain Model and cellular automata. Sustainability 8(3):236.

Velázquez A, Mas JF, Díaz Gallegos JR, Mayorga Saucedo R, Alcántara PC, Castro R, Fernández T, Bocco G, Ezcurra E, Palacio JL. 2002. Patrones y tasas de cambio de uso del suelo en México. Gaceta ecológica (62).

Weiskittel AR, Crookston NL, Radtke PJ. 2011. Linking climate, gross primary productivity, and site index across forests of the western United States. Canadian Journal of Forest Research 41(8): 1710-172.

Ximenes AC, Almeida CM, Amaral S, Escada MIS, Aguiar APD. 2011. Spatial dynamic modelling of deforestation in the Amazon. In Cellular Automata-Simplicity Behind Complexity. InTech.

Yanai AM, Fearnside PM, de Alencastro Graça PML, Nogueira EM. 2012. Avoided deforestation in Brazilian Amazonia: simulating the effect of the Juma Sustainable Development Reserve. Forest Ecology and Management 282:78-91.

STUDY III. A MULTIVARIATE GEOMORPHOMETRIC APPROACH TO PRIORITIZE EROSION-PRONE WATERSHEDS

BY:

Jesús A. Prieto-Amparan[1], Alfredo Pinedo-Alvarez[1], Griselda Vázquez-Quintero[2], María C. Valles-Aragón[2], Argelia E. Rascón-Ramos[1], Martin Martinez-Salvador[1] and Federico Villarreal-Guerrero[1]*.

[1]Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada km 1., C.P. 31453. Chihuahua, Chih. México.
[2]Facultad de Ciencias Agrotecnológicas, Universidad Autónoma de Chihuahua, Campus I, s/n., C.P. 31350. Chihuahua, Chih., México.

*Corresponding author: Federico Villarreal Guerrero; Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada km 1., C.P. 31453. Chihuahua, Chih. México. fvillarreal@uach.mx, tel. +52-(614)-434-1448., DOI: https://doi.org/10.3390/su11185140

ABSTRACT

A MULTIVARIATE GEOMORPHOMETRIC APPROACH TO PRIORITIZE EROSION-PRONE WATERSHEDS

BY:

M. C. JESÚS ALEJANDRO PRIETO AMPARÁN

Soil erosion is considered one of the main degradation processes in ecosystems located in developing countries. In northern Mexico, one of the most important hydrological regions is the Conchos River Basin (CRB) due to its utilization as a runoff source. However, the CRB is subjected to significant erosion processes due to natural and anthropogenic causes. Thus, classifying the CRB's watersheds based on their erosion susceptibility is of great importance. This study classified and then prioritized the 31 watersheds composing the CRB. For that, multivariate techniques such as principal component analysis (PCA), group analysis (GA), and the ranking methodology known as compound parameter (*Cp*) were used. After a correlation analysis, the values of 26 from 33 geomorphometric parameters estimated from each watershed served for the evaluation. The PCA defined linear-type parameters as the main source of variability among the watersheds. The GA and the *Cp* were effective for grouping the watersheds in five groups, and provided the information for the spatial analysis. The GA methodology best classified the watersheds based on the variance of their parameters. The group with the highest prioritization and erosion susceptibility included watersheds RH24Lf, RH24Lb, RH24Nc, and RH24Jb. These watersheds are potential candidates for the implementation of soil conservation practices.

Key words: prioritization; geomorphometric parameters; compound parameter; geospatial distribution; GA.

RESUMEN

A MULTIVARIATE GEOMORPHOMETRIC APPROACH TO PRIORITIZE EROSION-PRONE WATERSHEDS

POR:

M. C. JESÚS ALEJANDRO PRIETO AMPARÁN

Doctor en Philosophia de Recursos Naturales

Secretaría de Investigación y Posgrado

Facultad de Zootecnia y Ecología

Universidad Autónoma de Chihuahua

Presidente: Dr. Alfredo Pinedo Alvarez

La erosión del suelo se considera uno de los principales procesos de degradación de los ecosistemas, principalmente en los países en vías de desarrollo. En el norte de México, una de las regiones hidrológicas más importantes es la Cuenca del Río Conchos (CRC) debido a su utilidad como fuente de escorrentía. Sin embargo, la CRC está sometida a importantes procesos de erosión por causas naturales y antropogénicas. Por lo tanto, es de gran importancia analizar los procesos de erosión que pertenecen a la CRC. Este estudio clasificó y priorizó las 31 cuencas que componen la CRC en función de su susceptibilidad a la erosión. Para ello, se utilizaron técnicas multivariadas como el análisis de componentes principales (ACP), el análisis de grupos (GA) y la metodología de clasificación conocida como parámetro compuesto (*Pc*). Después de un análisis de correlación, los valores de 26 de 33 parámetros geomorfométricos estimados de cada cuenca sirvieron para la evaluación. El ACP definió los parámetros de tipo lineal como la principal fuente de variabilidad entre las cuencas hidrográficas. La AG y el *Pc* fueron efectivos para agrupar las cuencas en cinco grupos, y proporcionaron la información para el análisis espacial. La metodología de la AG clasificó mejor las cuencas en base a la varianza de sus parámetros. El grupo con la mayor prioridad y susceptibilidad a la erosión incluyó las cuencas hidrográficas RH24Lf, RH24Lb, RH24Nc, y RH24Jb. Estas cuencas son candidatas potenciales para la implementación de prácticas de conservación de suelos.

Palabras clave: Priorizacion; parámetros geomorfometrícos; parámetro compuesto; distribución geoespacial; SIG.

INTRODUCTION

Soil erosion is considered one of the most important degradation processes in the world (Alexakis *et al*., 2013; Gansari *et al*., 2015). The soil resource is limited and its wide use is of utmost importance; it sustains biogeochemical processes and is the habitat for a great diversity of microorganisms (Gajbhiye *et al*., 2015). Sustained soil development, conservation, and restoration is one of today's main challenges for humankind.

Hydric erosive processes affect the fertile soil layer, which is a key factor in the primary production of ecosystems (Gutierrez *et al*., 2004). The production of goods and satisfiers for the population such as wood, food, fiber, fodder, water, and recreational areas, among others, in addition to industrial expansion and the need for infrastructure facilities, have increased land-use/land-cover changes, increasing the pressure over the soil (Biswas *et al*., 1999). This has caused experts to pay more attention to the growing trend of soil erosion and the importance of water and soil conservation for achieving sustainable development.

Integrated watershed management is an alternative for soil management (Adhami and Sadeghi, 2016; Malik *et al*., 2019; Robinne *et al*., 2019). Watersheds are one of the spatial units that are used for the planning and management of soil resources (Khosravi *et al*., 2018). Management implies the characterization of the ecosystems inside the watershed and the understanding of the relationships between uplands, lowlands, land use/land cover, geomorphic processes, slope, and soil (Chen *et al*., 2011; Rahaman *et al*., 2015). In watershed management, erosion control is one of the main components (Gajbhiye *et al*., 2015). Thus, the hydrological planning and monitoring of a watershed is important for the development of environmental policies (Sharma *et al*., 2014). In this sense, the quantification of the watersheds' characteristics is fundamental to understanding their dynamics and degradation levels. This knowledge serves to define and implement strategies to prevent soil erosive processes and promote the conservation and restoration of watersheds (Keesstra *et al*., 2016).

Morphometry is used in the analysis of the watershed configuration (Clarke, 1996). Such methodology was developed by Horton (1945) and then modified by Strahler (1964), and provides information on the behavior of the basin (Singh *et al*., 2014). It is an important tool for identifying and prioritizing eroded watersheds (Nautyyal, 1994).

Nevertheless, monitoring soil erosion in situ is costly and time consuming in large watersheds. Thus, the analysis of geomorphometry is often carried out based on geographic information systems (GIS, Shrimali *et al*., 2001; Thakker and Dhiman, 2007; Sharma *et al*., 2010; Viramontes-Olivas *et al*., 2008; Tilahun *et al*., 2014). On a spatial scale, geomorphometric parameters, i.e., the Gravelius compactness coefficient (Zavoianu, 1985) and elongation ratio (Shumm, 1956), among others, are important to know the hydrological configuration of watersheds. The relationships among these parameters are useful for developing hydrological models, which allow prioritizing watersheds based on their condition, such as erosion susceptibility. To determine the aforementioned relationships, statistical methods, such as multivariate techniques, have been widely used worldwide (Saha *et al*., 2012; Sharma *et al*., 2013). For instance, Gavit et al. (2016) characterized 13 geomorphometric parameters in 11 watersheds located in the Godavari river in Maharashta, India. Youssef *et al*. (2011) estimated the erosion risk by using remote sensing technology, GIS, and geomorphometric parameters in 11 watersheds located in Sinai, Egypt. Makwana and Tiwari (2016) used seven geomorphometric parameters to characterize 19 watersheds in the region of Gujarat, India. Sharma *et al*. (2014) applied the multivariate technique of principal component analysis (PCA) to 13 geomorphometric parameters from eight watersheds located in the Madhya district, India. Meshram and Sharma (2017) and Farhan *et al*. (2017) used PCA to analyze the geomorphometric parameters of a group of watersheds located in the Shakkar Catchment River in India and Jordan.

A large number of shape, relief, and hydrological parameters are associated with watershed geomorphometry (Bach *et al*., 2003; Sanderson *et al*., 2012). The statistical techniques of PCA and group analysis (GA) (Miranda *et al*., 1996; Castillo-Rodriguez *et al*., 2010; Bateyneh and Zumlot, 2012; Oketola *et al*., 2013; Tritsch *et al*., 2016; Prieto-Amparán *et al*., 2019), as well as multivariate analysis of variance (MANOVA) and the ranking methodology known as compound parameter (*Cp*) (Altaf *et al*., 2014) have been widely used in recent years for the analysis of environmental data from watersheds. These techniques assist with analyzing the spatial variability of watersheds, their structure, as well as the relationships existing among them.

The most important basin in the state of Chihuahua, as a runoff source (Montero-Martínez and Ibáñez-Hernandez, 2017), is the Conchos River Basin (CRB). Yet, this basin has experienced water stress conditions during the past years. According to Ordoñez (2017), approximately 8000 km^2 (11.8 %) of the basin high lands present strong erosion problems, which could impact waterflow and water quality. In these high lands, deforestation and land-use/land-cover changes had contributed to reduce the amount of infiltrated water, impacting on groundwater availability (Montero-Martínez and Ibáñez-Hernandez, 2017). In the low lands of the basin, agriculture consumes 90 % of the water harvested in the basin. Other consumers include the industrial and domestic sectors (Mahlknecht *et al*., 2008; Montero-Martínez and Ibáñez-Hernandez, 2017). In addition, the international water trade between Mexico and the U.S. from 1944 commands Mexico to deliver annually from this basin a total of 432 × 106 m^3 of water to the U.S. (Sánchez, 2006). Therefore, specific knowledge about water management and the status of the basin's soil erosion is required to implement strategies to solve water-related problems and to promote the sustainable development of the region.

The objective of this study was: (1) to describe the behavior of the 31 watersheds located along the CRB in the state of Chihuahua, Mexico, based on the values of their geomorphometric parameters; (2) to spatially classify the 31 watersheds into groups; and (3) to prioritize the groups of basins according to their erosion susceptibility. For the grouping, multivariate techniques and the compound parameter (*Cp*) were used and their efficacy compared.

MATERIALS AND METHODS

Study Area

The CRB is located in the state of Chihuahua and Durango, Mexico, and is part of the 24th Rio Bravo-Conchos Hydrological Region (CONAGUA, 2001) (Figure 1). The basin has an area of 67,800 km^2 (Nuñez Lopez *et al*., 2014), with an altitudinal distribution ranging from 841 m to 2348 m (Viramontes-Olivas *et al*., 2008). It presents a diversity of climates ranging from temperate in the upper, semi-arid in the middle, and arid in the lower parts of the basin (Aboites-Aguilar, 2002). The physiography of the upper basin belongs to the mountainous zone made up of temperate forests dominated by species of pines and oaks. The middle part of the Altiplano or central valleys is made up of transition zones where oaks and bushes are present. Regarding the lower part, it belongs to the arid region and is made up of shrublands and grasslands (INEGI, 2014). The basin has a precipitation regime from June to September, with July and August being the wettest months, and fluctuating from 200 mm to 700 mm (Nuñez Lopez *et al*., 2014).

Data

Data of the CRB was obtained from the online GIS source of CONABIO (2019). Likewise, the data of the 31 watersheds composing the CRB (Figure 2) were obtained from the Watershed Water Flows Simulator (SIATL, 2019). The relief and hydrology type parameters were estimated by processing the necessary data from a Digital Terrain Model (DTM), with a resolution of 15×15 m, downloaded from INEGI (2019). The values of the basic parameters from the watersheds were obtained by using the Hydrology tool (Hamdy *et al*., 2016) of

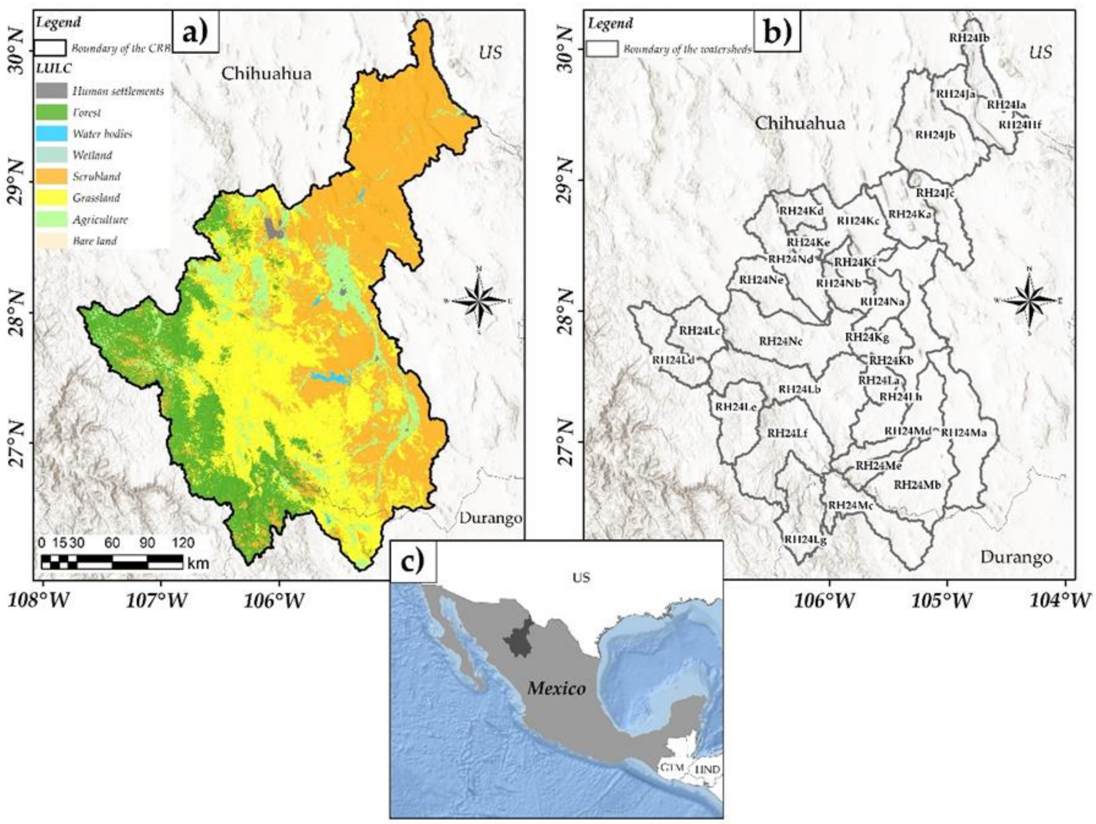

Figure 3. (a) Land-use/land-cover types of the Conchos River Basin (CRB), (b) Delimitation of the 31 watersheds of the CRB, (c) Location of the CRB in Mexico.

ArcMap© 10.3 (ESRI, Redlands, CA, USA; https://wwwesri.com/en-us/home). The values of the shape, relief, and linear type parameters were calculated from the equations listed in Table 1.

Basic parameters include the area (*A*), perimeter (*P*), watershed length (Lb^2), stream order (*u*), main channel length (*Lc*), all channel lengths (*Lu*), and contour length (*Li*). The area and perimeter are the most important parameters of the watersheds to understand their hydrological design and reflect the volume of water that can be discharged in a rainfall event (Patel *et al*., 2013).

Shape parameters include the Gravelius compactness coefficient (*Cc*), elongation ratio (*Re*), shape factor (*Rf*), elongation index (*Ia*), unit shape factor (*Ru*), and circularity ratio (*Rc*). The Gravelius compactness coefficient is the relationship between the perimeter of the watershed and the perimeter of a circle of area equal to that of the watershed (Zavoianu, 1985). Low values of the Gravelius compactness coefficient, shape factor, and elongation index indicate a more elongated watershed, while high values correspond to a less elongated watershed. Watersheds with less elongated shapes have a shorter maximum flow duration, while elongated watersheds correspond to watersheds with longer flow duration (Zavoianu, 1985). The elongation ratio is the diameter of a circle with an area equal to that of the watershed divided by the length of the watershed (Meshram *et al*., 2017). Those watersheds with values close to or greater than the unit correspond to flat watersheds, while values that move away from the unit are watersheds with pronounced relief (Shumm, 1956). The shape factor is the relationship between the area and the length of the watershed (Horton, 1945). The elongation index is also a relationship between the length and width of the

Table 1. Geomorphometric parameters.

Geomorphometric Parameter	Equation	References
	Basic Parameters [1]	
Area (A)	A = Watershed surface area [km^2]	Horton [15]
Perimeter (P)	P = Watershed perimeter [km]	Horton [15]
Length (Lb^2)	Lb^2 = Watershed length [km]	
Stream order (u)	u = Stream order [unitless]	Strahler [16]
Main Channel Length (Lc)	Lc = Main flow channel length [km]	
All Channel Lengths (Lu)	Lu = Length of all the flow channels in the watershed [km]	Horton [15]
Contour Length (Li)	Li = Contour lines' length [km]	
Number of Flow Channels (Nu)	Nu = Number of flow channels [unitless]	
Number of First-Order Flow Channels (No_1)	No_1 = Number of total first-order flow channels in the watershed [unitless]	
Maximum Height (H_{max})	H_{max} = Watershed maximum height [m]	
Minimum Height (H_{min})	H_{min} = Watershed minimum height [m]	
Medium Height (H_{med})	H_{med} = Watershed medium height [m]	
	Shape Parameters	
Gravelius Compactness Coefficient (Cc)	$Cc = P/2\sqrt{\pi A}$	Zavoianu [24]
Elongation Ratio (Re)	$Re = 1.1284\ (\sqrt{A}/Lc)$	Schumm [25]
Shape Factor (Rf)	$Rf = A/Lb^2$	Horton [15]
Elongation Index (Ia)	$Ia = Lb^2/W$ where: W = watershed width (Km)	Horton [15]
Unit Shape Factor (R_U)	$R_U = Lb^2/A^{0.5}$	Horton [15]
Circularity Ratio (Rc)	$Rc = 4\pi A/P^2$	Miller [56]
	Relief Parameters	
Mean Watershed Slope (J)	$J = (\Sigma Li\ E/A) \times 100$ where: E = equidistance among contour lines (km)	Horton [15]
Massivity Coefficient ($tg\alpha$)	$tg\alpha = H_{med}/A$	

Relief Relationship(Rh)	$Rh = H_{max}/Lb$	Schumm [25]
Relative Relief (Rr)	$Rr = H_{max}/P$	Schumm [25]
Orographic Coefficient (Co)	$Co = Hmed \times tg\alpha$	
Linear parameter		
Drainage Density (Dd)	$Dd = \Sigma Lu/A$	Horton [15]
Mean Slope of the Main Channel (j)	$j = (H_{max} - H_{min})/L_c \times 100$	Horton [15]
Mean Distance (Am)	$Am = L_c/(\sqrt{A})$	Horton [15]
Sinuosity of the Main Flow Channel (Scp)	$Scp = L_c/Lb^2$	Mueller [57]
Kirpich Concentration Time (TcK)	$TcK = 0.066(Lb^2/j)^{0.77}$	Kirpich [58]
Average Peak Flow (Qp)	$Qp = 43A^{0.522}$	Sen [59]
Texture Ratio (T)	$T = N_u/P$	Horton [15]
Rivers Frequency (Fu)	$Fu = N_u/A$	Horton [15]
Resistance Number (Rn)	$Rn = H_{max} \times Dd$	Schumm [25]
General Flow Length (Lo)	$Lo = {}^1/_2 \times Dd$	Schumm [25]
Drainage Intensity (Di)	$Di = Fu/Dd$	Faniran [60]

[1] The basic parameters were calculated with the GIS software.

watershed (Horton, 1945). The unit shape factor is the relationship between the length and the area of the watershed. Values less than 2 indicate that they have weak flood discharge periods, while values greater than 2 indicate that their flood discharge is strong (Makwana *et al*., 2016). The circularity ratio is defined as the relationship between the area and the perimeter of the watershed. If the circularity of the watershed is low, then discharge will be slow, and the susceptibility to erosion will be low (Patel *et al*., 2013).

Relief parameters include the mean watershed slope (*J*), massivity coefficient (*tgα*), relief ratio (*Rh*), relative relief (*Rr*), and orographic coefficient (*Co*). The mean watershed slope indicates the degree of the terrain roughness. As the slope increases, the watershed will be prone to erosion. The massivity coefficient represents the relationship between the mean watershed height and its surface area, which is expressed as a percentage. Small values of the massivity coefficient correspond to very mountainous watersheds, while large values correspond to watersheds from moderately mountainous to flat. The relief ratio is directly related to the slope of the currents and the watershed surface, being an indicator of hydrological processes and erosion. The relief ratio, similar to the relative relief, has a direct correlation with the watershed erosion processes (Shumm, 1956).

Linear parameters include drainage density (*Dd*), mean slope of the main channel (*j*), mean distance (*Am*), sinuosity of the main channel (*Scp*), Kirpich concentration time (*Tc*), average peak flow (*Qp*), texture ratio (*T*), rivers frequency (*Fu*), resistance number (*Rn*), general flow length (*Lo*), and drainage intensity (*Di*). High drainage density values indicate a high current density, and therefore a rapid response to precipitation events (Agarwal *et al*., 2009). The mean slope of the main channel indicates the slope of the longest channel in the watershed. The high values of this parameter indicate that sediment flow and entrainment will quickly exit the watershed (Horton, 1945). The sinuosity of the main channel indicates the velocity of flow movement in the channels. The lowest values of sinuosity correspond to channels where the flow travels at greater speed, whereas the channels with the highest values of sinuosity indicate the slow movement of the flow (Muller *et al*., 1968). The Kirpich concentration time is the time when a drop of water falls at the furthest point until it reaches the exit point (Kirpich, 1940). Average peak flow is defined as the mean maximum amount of water passing through a specific section (Sen, 2008). The texture ratio corresponds to the relationship between the total number of streams and the watershed perimeter. Rivers frequency represents the total number of streams of all orders per unit area (Horton, 1945). The resistance number is used to measure the flood potential of rivers. It has a direct relationship with erosion, where increasing its value would represent an increment in erosion susceptibility (Horton, 1945).

Watershed's Description and Classification

Prior to the watersheds' classification, their geomorphometric parameters were correlated (Dillon and Goldstein, 1984; Adhami and Sadeghi, 2016). Correlation indicates when part of the information contained in a set of geomorphometric parameters is also contained in the remaining ones (Meshram and Sharma, 2017). That served to reduce the number of parameters included in the subsequent analysis.

To describe the variability of the geomorphometric parameters, principal component analysis (PCA) was performed. The technique of PCA reduces the data dimensionality, simplifies the dataset, and makes it easier to explain through a set of new principal components (PCs, Jollifie, 2002; Yidiana *et al*., 2010). The first principal component is the linear combination of the original geomorphometric parameters that contributes the most to the total variance; the second principal component, not correlated to the first, contributes the most to the residual variance, and so on until the total variance is analyzed. The PCs are orthogonal variables that could be obtained by multiplying the original variables, which are correlated, with coefficients similar to the ones described in Equation (1):

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + \cdots + a_{im}x_{mj} \tag{5}$$

where: *z* represents the coefficient of the component, a represents the weight of the component, *x* represents the measured value of the variable, *i* corresponds to the component number, *j* represents the sample number, and *m* represents the total number of variables.

The PCA was applied to the values of geomorphometric parameters to calculate the correlation matrix and to obtain the PCs (Bro, 2014). Both the correlation analysis and the PCA were performed in the SAS© 9.1.3 software (The SAS Institute Inc., Cary, NC, USA, https://www.sas.com/en_us/home.html). Then, the PCs were mapped for interpretation.

For the classification, the first method used for the watersheds was a hierarchical group analysis, which was based on the Ward criterion (Ward, 1936). The Ward criterion was applied to the GA by using the square Euclidean distance to explore the clustering of the 31 watersheds. The definition of the watershed groups was performed based on R^2 (Eder, 1994). Finally, a multivariate analysis of variance (MANOVA) was used to analyze whether significant multivariate differences exist between the groups based on the values of their geomorphometric parameters (Johnson, 2007).

The second classification method considered in this work was the compound parameter (*Cp*). Previous research has employed this approach for sustainable watershed planning and management (Altaf *et al*., 2014). Linear and shape parameters have been commonly used for this method, whereas the relief and basic parameters were additionally included in this study. Linear geomorphometric parameters have a direct relationship with erosion susceptibility, where high values are more likely to result where high erosion probabilities are present (Nooka Ratman, 2005; Patel *et al*., 2013). Thus, for watershed classification, the highest value of linear parameters was ranked as 1, the second highest was ranked as 2, and so on. Conversely, shape parameters have an inverse relationship with erosion (Nooka Ratman, 2005; Patel *et al*., 2013), and low values are related to high susceptibility to erosion and vice versa. Then, the lowest value of the shape parameters was ranked as 1, the next lowest value was ranked as 2, and so on. Regarding the relief and basic parameters, the highest value was ranked as 1, the second highest value was ranked as 2, and so on. After this procedure was completed, the ranked values from each watershed were summed and then averaged to produce the *Cp* of each watershed. This average represents the collective impact of all the parameters, and is calculated according to Equation 2 (Altaf *et al*., 2014):

$$C_P = \frac{1}{n}\sum_{i=1}^{n} R \qquad (6)$$

where: *Cp* is the compound parameter, *R* is the ranked value of each parameter, and *n* is the number of parameters.

Based on the *Cp*, the highest priority was assigned to the watersheds with the lowest *Cp* value, the second priority was assigned to the next higher *Cp* value, and so on. Then, the *Cp* was classified into five categories or groups of erosion susceptibility. The categories were defined as: Very High (Group 5), High (Group 4), Moderate (Group 3), Low (Group 2), and Very Low priority (Group 1), similar to classifications made in previous studies (Fajardo *et al*., 2014).

Comparison of the Classification Methods

To compare the grouping methods, the following procedure was followed.

First, an ANOVA was carried out for each geomorphometric parameter (independent variable), and separated for each grouping method. The source of variation or class was considered to be the group. Such analyses served to determine possible significant differences among the groups within each grouping method. After the ANOVA analyses were completed, the grouping method that achieved the highest number of significant P-values ($\alpha<0.05$) was considered the most effective for grouping the watersheds of the CRB.

RESULTS AND DISCUSSION

Characterization of Conchos River Basin Watersheds´s

The values of the basic, shape, relief, and linear-type parameters from the 31 watersheds used in this study are shown in Table 1. Watershed RH24Ia has the smallest area and perimeter with 78.62 km^2 and 50.59 km, respectively. Meanwhile, the watershed with the largest area and perimeter is RH24Lb, with 5428.29 km^2 and 640.77 km, respectively. The watershed with the highest stream order is RH24Kb, while watershed RH24Lb has the longest main channel. First-order streams do not have tributaries, and their flow depends on the secondary surface contributions that converge to them (Patel *et al*., 2013). The watershed with the highest number of channels of order one is RH24Ia, whereas the watershed with the lowest number of channels is RH24Jb. The watershed lengths vary from 11.17 km to 133 km. Watershed RH24Kb presents the largest elongation ratio value, indicating that it is the flattest, while watershed RH24Ib has the lowest value for this parameter, indicating that it has the steepest slope (Shumm *et al*., 1956). The values of sinuosity of the main channel vary from 0.05 km to 4.1 km. The watershed with the shortest Kirpich concentration time is RH24Kb, while watershed RH24Lb has the longest. The lowest average peak flow value corresponds to watershed RH24Ia, whereas watershed RH24Lb presents the highest. The texture ratio values are between 4.98–75.1, which are considered as moderate to high values; the low values correspond to watershed RH24Na, while the highest value belongs to watershed RH24Lg, so the former is not susceptible to erosion, while the latter is.The watershed with the lowest resistance value is RH24Na, while the watershed with the highest value is RH24Lg.

Correlations and Principal Component Analyses

The data matrix of the 31 watersheds and the 33 parameters, which included the basic shape, linear, and relief type parameters, were analyzed through a correlation analysis. A set of parameters showing high correlations were identified. From each pair of highly correlated parameters, only one parameter was chosen, and the rest were eliminated to reduce the data dimensionality. After the reduction, the final number of geomorphometric parameters was 26, as listed in Table 2.

The PCA was performed on the 26 parameters and showed that the first five principal components were the most important for explaining the data variance (Figure 2). The most important parameter was selected according to its contribution to the principal component, as it is shown by the values of the eigenvectors (in bold) of the correlation matrix (Table 2). The first five principal components accounted for 88.44 % of the total variance of the dataset. The linear parameters (hydrology) are the ones mainly explaining the CRB behavior (Table 3).

Watershed's Classification Based on Group Analysis

Figure 3 shows the dendrogram, resulting from the group analysis (GA) of the 31 watersheds. Five groups were identified based on their basic, relief, shape, and hydrology-type parameters, and considering a value of $R^2 = 0.84$. Group 1, with seven watersheds, has the largest amount of low values for the shape, relief, and linear-type parameters, as can be verified in Table 4. Group 2, with eight watersheds, presents the lowest average values of drainage density, sinuosity of the main channel, and general flow length. Group 3, also with eight watersheds,

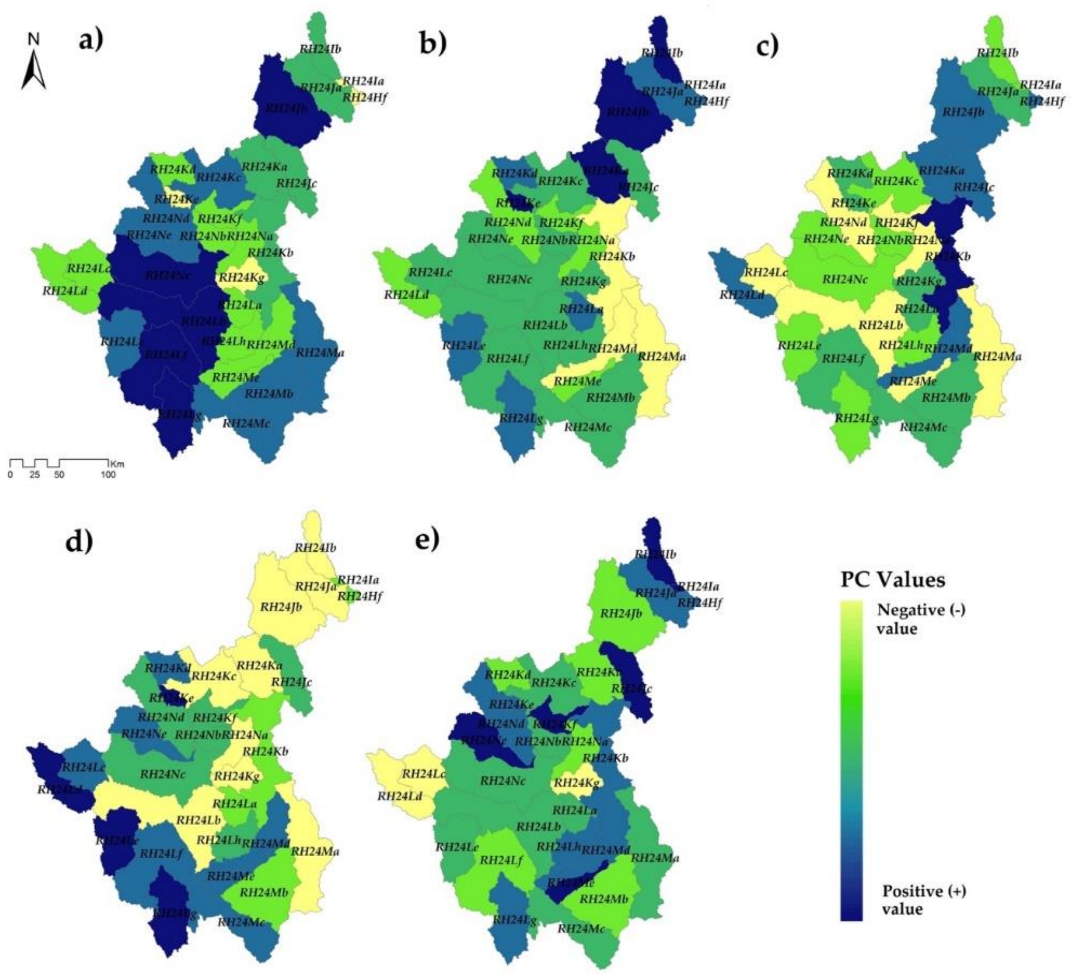

Figure 2. Classification of 31 watersheds based on the values of five principal components resulting from a principal component analysis. Conchos River Basin, Chihuahua, Mexico. (a) PC1, (b) PC2, (c) PC3, (d) PC4, (e) PC5. PC1 = Principal component 1, PC2 = Principal component 2, PC3 = Principal component 3, PC4 = Principal component 4, PC5 = Principal component 5.

Table 2. Eigenvectors of the correlation matrix.

GP	PC1	PC2	PC3	PC4	PC5
Lc	0.2678	0.0223	−0.2753	−0.1330	0.0095
Lb^2	0.2345	−0.2721	−0.0327	−0.0059	0.0880
Li	0.2872	0.0343	0.0149	0.1182	−0.1631
Lu	**0.3014**	0.0344	0.1560	−0.1236	−0.0398
H_{max}	0.2170	−0.1034	−0.0543	0.3656	−0.1169
H_{min}	0.0664	−0.0941	−0.1183	**0.4424**	−0.3506
Cg	0.1308	−0.2425	0.0677	0.0028	0.2681
Re	−0.0340	−0.1992	**0.4039**	0.0072	0.0603
Rf	0.2180	0.1309	0.2002	−0.1898	−0.3142
Ia	0.0510	−0.3141	−0.1533	0.1096	**0.3442**
J	0.1269	0.1467	−0.0457	**0.4034**	−0.1466
tgα	0.2759	−0.0721	0.1407	−0.2389	−0.0850
Dd	0.0720	**0.3174**	0.1766	−0.0248	**0.3345**
j	−0.0484	−0.2118	0.4035	0.0593	0.0789
a	0.1271	0.1028	**−0.4225**	−0.0973	0.1850
TcK	0.2483	0.0091	−0.2925	−0.1728	−0.0094
Scp	0.0694	0.2821	−0.2272	−0.1822	0.0258
Qp	**0.3031**	−0.1104	0.0908	−0.0989	−0.1442
T	0.2642	0.1539	0.1430	0.0413	−0.0230
Ru	0.0641	**−0.3193**	−0.1582	0.1261	**0.3286**
Fu	0.1737	0.2478	0.1291	0.1961	0.2488

Rn	0.2305	0.1724	0.0819	0.2424	0.1779
Rh	−0.2207	0.2512	0.0133	0.1126	−0.0453
Rr	−0.2415	0.1661	−0.0986	0.2273	0.0458
Lo	0.0720	**0.3174**	0.1766	−0.0248	**0.3345**
Di	0.1911	0.0795	0.0241	0.2992	0.1179

GP = Geomorphometric parameter.
PC1 = Principal component 1, PC2 = Principal component 2, PC3 = Principal component 3, PC4 = Principal component 4, PC5 = Principal component 5.
Lc = Length of main channel, Lb^2 = Length of watershed, *Li* = Length of contour lines, *Lu* = Length of channels, *Hmin* = Minimum height, *Hmax* = Maximum height, *Cc* = Gravelius compactness coefficient, *Re* = Elongation ratio, *Rf* = Form factor, *Ia* = Elongation index, *J* = Mean slope of watershed, *tgα* = Mass coefficient, *Dd* = Drainage density, *j* = mean slope of the main channel, *a* = Medium distance, *TcK* = Kirpich concentration time, *Scp* = Sinuosity of the main channel, *Qp* = Average peak flow, *T* = Texture ratio, *Ru* = Unit shape factor, *Fu* = River frequency, *Rn* = Resistance number, *Rh* = Relief ratio, *Rr* = Relative relief, *Lo* = General flow length, *Di* = Drainage intensity.
Bold letters indicate the dominant coefficient.

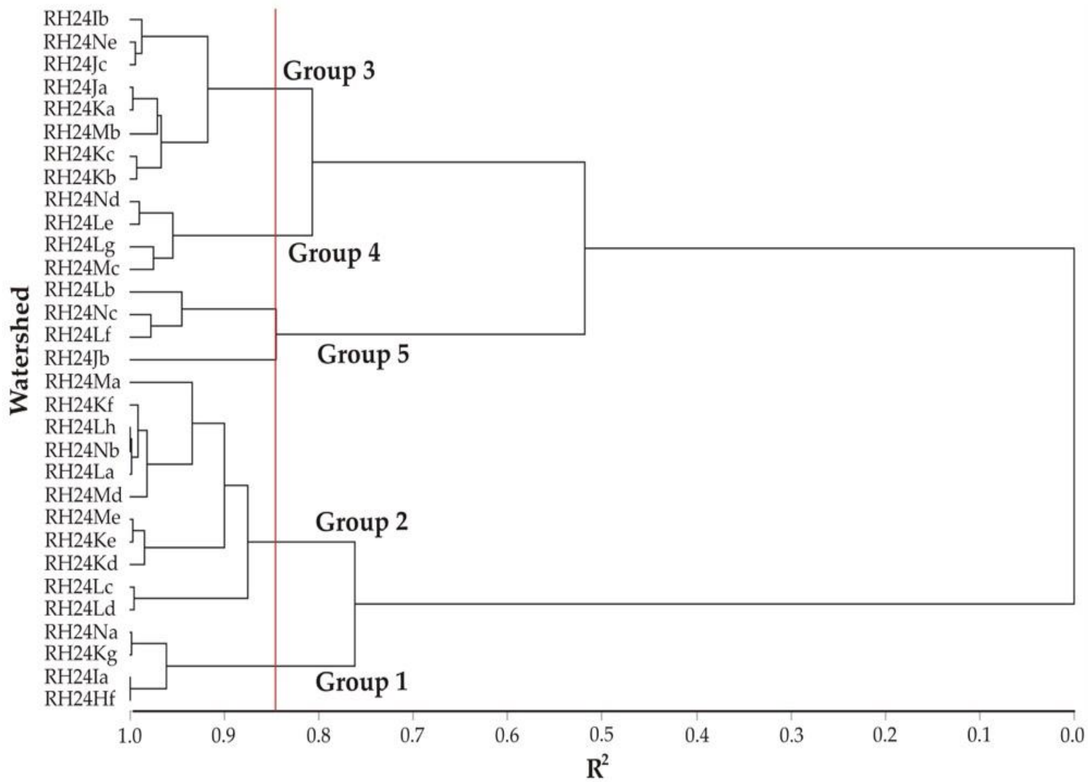

Figure 3. Dendrogram classifying 31 watersheds by group analysis. Conchos River Basin, Chihuahua, Mexico. The red line was drawn to define the groups.

showed the highest values of elongation ratio, drainage density, mean slope of the main channel, and general flow length. Group 4, with four watersheds, has the highest values in maximum and minimum height, mean slope of the watershed, mean distance, unit shape factor, resistance number, and drainage intensity. In this group, the lowest values correspond to the elongation ratio. Group 5, also with four watersheds, presents the largest amount of high values for the basic, shape, relief, and linear-type parameters. The multivariate analysis of variance (MANOVA) showed significant differences among the groups of watersheds, showing a value of Wilks' lambda equal to 0.0025, with a value of $P<0.0001$.

The geospatial distribution of the groups is shown in Figure 4. Group 1 shows a homogeneous pattern in its distribution, which is concentrated in the central part of the study area. In contrast, Group 2 shows a dispersed distribution, mainly at the edges of the CRB. Group 3 is distributed in the northern, central, and southern parts, and is represented by small clusters of two or three watersheds. Group 4 corresponds to watersheds spatially dispersed over the basin. The watersheds of these groups are isolated from the other watersheds of the same group. Group 5 shows watersheds clustered in the southern part of the basin, except for the watershed RH24Jb, which is located in the northern region. The GA showed a clustered geospatial pattern for Groups 1, 3, and 5, who share characteristics in their parameters and space.

Watershed's Classification Based on the Compound Parameter (*Cp*)

Considering the 26 geomorphometric parameters selected after the correlation analysis was performed, the value of the compound parameter (*Cp*)

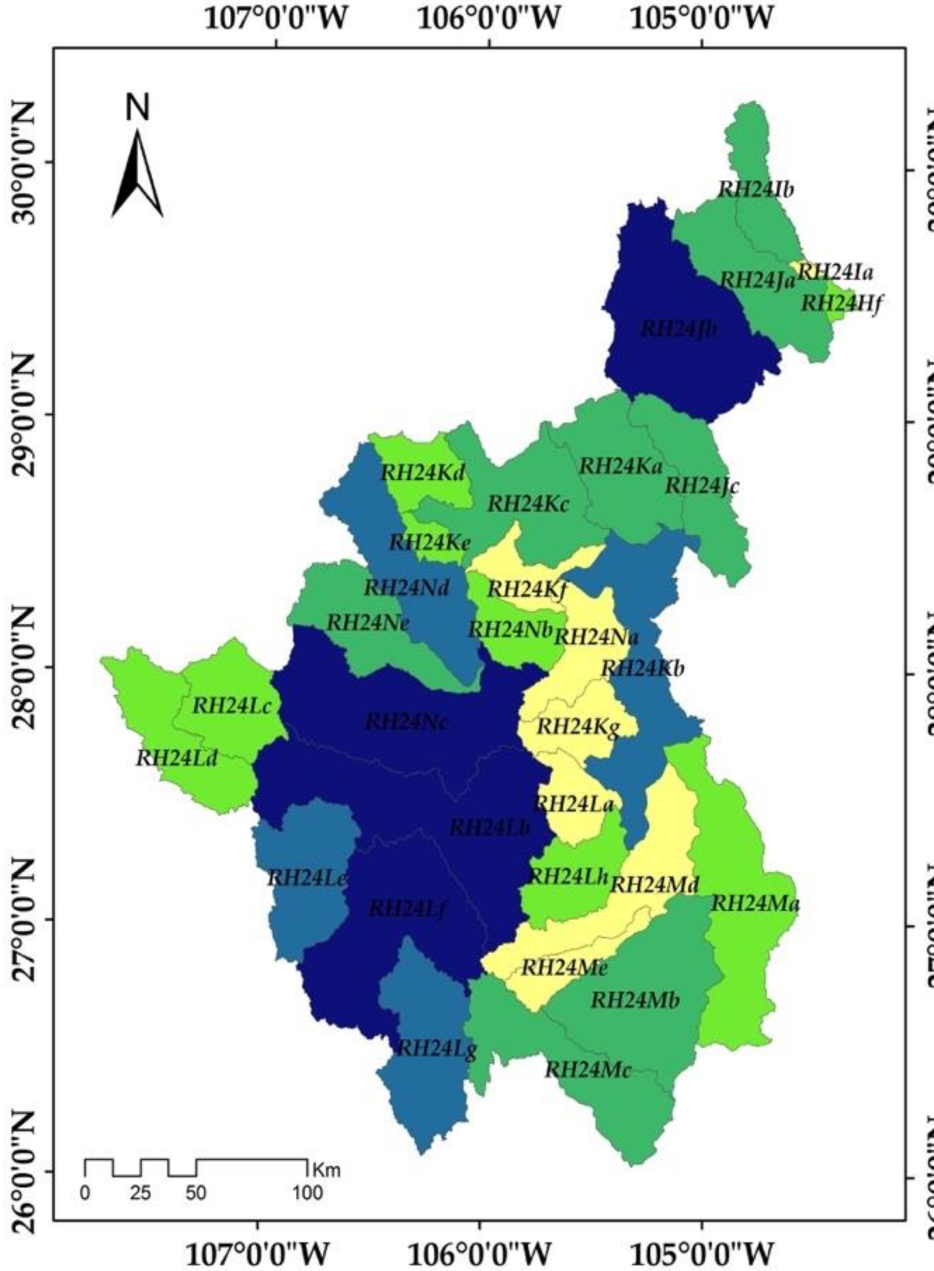

Figure 4. Geospatial distribution of watershed groups classified by group analysis. Very High (●), High (●), Moderate (●), Low (●), Very Low (●). Conchos River Basin, Chihuahua, Mexico.

was calculated for the 31 watersheds of the CRB (Figure 5a). The watershed RH24Lg received the highest priority (1), followed by the watershed RH24Le (2). The watershed with the lowest priority (31) was watershed RH24Kg. A high priority is an indicator of a high degree of erosion susceptibility in the watershed.

The resulting *Cp* map (Figure 5a) was reclassified into the following five categories: Very High, High, Moderate, Low, and Very Low (Figure 5b). The spatial distribution of the groups was reclassified by natural breaks (Fajardo *et al*., 2014). The watersheds classified as Very High show a homogeneous pattern in their distribution in the southwestern part of the watershed. Meanwhile, the watershed RH24Jb is isolated in the northwestern part. The High, Moderate, and Low classes show a dispersed distribution, with at least two of their watersheds clustered in space. The Very Low class shows a homogeneous distribution in space in the southeastern part of the study area, with only one dispersed watershed (RH24Hf).

Comparison of the Classification Methods

Regarding the GA classification method, the results from the ANOVA analyses performed on 14 geomorphometric parameters, out of 26, detected significant differences among the groups defined by the method. In the case of the *Cp* classification method, the ANOVA analysis of only two parameters detected significant differences among the groups defined by this classification method (Table 3).

The prioritization of watersheds, based on susceptibility to erosion, has been carried out in different regions of the world (Paul *et al*., 2012; Darabi *et al.*, 2014) using different prioritization methods (Rawat *et al.,* 2014). This study

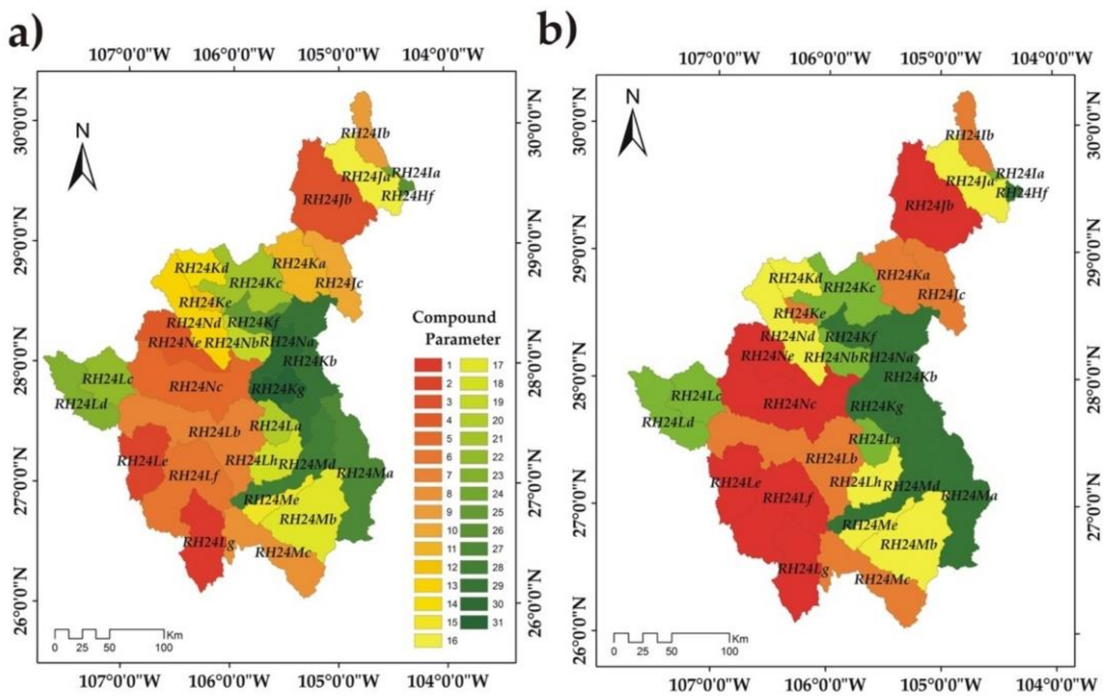

Figure 5. (a) Watersheds and their compound parameter (*Cp*). (b) Geospatial distribution of watershed groups by *Cp* reclassification. Very High 1 (●), High 2 (●), Moderate 3 (●), Low 4 (●), Very Low 5 (●). Conchos River Basin, Chihuahua, Mexico.

Table 3. Average values of the geomorphometric parameters by group.

G_{id}	*Lb*^{2}	*Lc*	*Li*	*Lu*	Hmin	Hmax	*Cc*	*Re*	*Rf*	*Ia*	*J*	*tgα*	*Dd*
1	46.12	31.31	371.43	1038.50	1495	995	3.22	0.74	20.03	1.69	6.43	0.49	1.97
2	80.16	65.56	1820.63	2660.91	2489	1415	3.70	0.91	23.28	3.11	13.62	0.79	1.97
3	102.47	75.16	2765.79	6206.70	2294	1060	3.95	1.99	34.73	2.41	12.07	1.48	2.59
4	146.56	88.94	5556.08	6201.08	2970	1545	3.86	0.40	30.30	3.10	21.45	1.17	2.36
5	192.52	99.49	8176.13	12,155.08	2680	1230	4.07	0.44	51.66	2.20	17.31	2.45	2.58

G_{id}	*j*	*a*	*TcK*	*Scp*	*Qp*	*T*	*Ru*	*Fu*	*Rn*	*Rh*	*Rr*	*Lo*	*Di*
1	1.42	1.92	11.74	1.51	194.51	8.83	1.24	2.56	2673.67	67.01	13.48	0.99	1.30
2	2.44	2.17	18.48	1.37	313.28	16.48	1.69	3.03	4835.28	47.81	11.26	0.98	1.51
3	3.70	2.18	22.19	1.69	417.90	27.78	1.50	3.86	5915.17	34.49	6.97	1.30	1.51
4	0.98	2.87	31.18	1.68	433.61	37.17	1.73	4.87	7040.26	34.58	8.63	1.18	2.04
5	0.82	2.77	42.44	1.94	589.36	49.69	1.44	4.99	6806.71	28.86	5.50	1.29	1.95

G_{id} = Group identification, 1 = Group 1 (Very low erosion susceptibility), 2 = Group 2 (Low erosion susceptibility), 3 = Group 3 (Moderate erosion susceptibility), 4 = Group 4 (High erosion susceptibility), 5 = Group 5 (Very high erosion susceptibility).

Lc = Length of main channel, *Lb*^{2} = Length of watershed, *Li* = Length of contour lines, *Lu* = Length of channels, *Hmin* = Minimum height, *Hmax* = Maximum height, *Cc* = Gravelius compactness coefficient, *Re* = Elongation elation, *Rf* = Form factor, *Ia* = Elongation index, *J* = Mean slope of watershed, *tgα* = Mass coefficient, *Dd* = Drainage density, *j* = Mean slope of main channel, *a* = Medium distance, *TcK* = Kirpich concentration time, *Scp* = Sinuosity of main channel, *Qp* = Average peak flow, *T* = Texture ratio, *Ru* = Unitary shape factor, *Fu* = River frequency, *Rn* = Resistance number, *Rh* = Relief ratio, *Rr* = Relative relief, *Lo* = Length of general flow, *Di* = Drainage intensity.

contributes to the lack of knowledge regarding the susceptibility to erosion in northern Mexico. This was assessed by implementing two methods for prioritization based on the analysis of a set of 33 parameters, which differ from other studies (Biswas *et al.,* 1999; Suji *et al.,* 2015; Malik *et al.,* 2019). The inclusion of several geomorphometric parameters and their relationships within several connected watersheds enriched the study of their erosion susceptibility. In this sense, multivariate techniques have proved to be appropriate methods for establishing priorities, reducing the dimensionality of the dataset by losing the least amount of information (Prieto-Amparan *et al.,* 2018).

This study integrated a multivariate analysis of several geomorphometric parameters that served to identify those watersheds, which may be prone to erosion. That was possible by evaluating the behavior of such geomorphometric parameters and representing them in a geospatial basis (Singh *et al.,* 2009; Yunus *et al.,* 2014). Their relationships provided significant information about the main sources of variability among the studied watersheds (Rawat *et al.,* 2014). Previous research studies have reported that topography, geomorphology, and land use/land cover are the most important factors in the watershed susceptibility to erosion (Adediran *et al.,* 2004; Javed *et al.,* 2009; Kompani-Zare *et al.,* 2011; Welde, 2016). In this study, the factors with the greatest influence on the hydrological behavior of watersheds and their erosion susceptibility were the average peak flow and the all channel lengths, as it has also been found in previous studies (Subyani *et al,*. 2012; Adhami and Sadeghi, 2016).

The PCA is considered a statistical exploratory technique, whose results have helped explain the distribution of environmental attributes (Johnson *et al*., 2007). Results from the PCA were useful to identify the sources of variance, which were mainly represented by the dominant parameters influencing the data structure. Then, the basin's hydrological configuration was explained by those geomorphometric parameters explaining the greatest portion of the variance among the watersheds. The PCA results from this study are consistent with the observations made by Meshram and Sharma (2017) and Farhan *et al*. (2017).

From the PCA analysis, PC1 and PC2 are mainly influenced by linear geomorphometric parameters. Some of the linear parameters with an influence on PC1 are the average peak flow. This is shown in Figure 3b, where the lowest PC1 coefficients correspond to the watersheds with the lowest mean slope values of the small channels. Regarding PC2, drainage density is one of the linear parameters with an influence. Watersheds with low drainage density indicate the presence of permeable surface material, good vegetation cover, and low relief (Harlin and Wijeyawickrema, 1985; Luo, 2000). The map of PC2 (Figure 3b) was highly influenced by drainage density, since the watersheds with low values of this parameter are located in the south–central part of the study area and grouped in such a map.

PC3 and PC4 are influenced by linear parameters such as mean distance and shape parameters such as elongation ratio, as well as topographic parameters such as minimum height and mean slope. These factors are associated with the main channel, relief, and slopes, among others. In Figure 3d, the watersheds with the greatest heights and slopes correspond to the watersheds located in a mountainous zone, while the watersheds with the lowest elevations and slopes (Rai *et al.,* 2017a) correspond to the arid and semi-arid zones of the state of Chihuahua. Regarding PC5, it is mainly influenced by the elongation index (shape parameter) and the general flow length (linear parameter). The high values of the elongation index correspond to enlarged watersheds, which are related to high drainage densities. A watershed with a high drainage density implies a quick hydrological response to rainfall events, while non-enlarged watersheds correspond to fan-shaped watersheds, which are characterized by short channels (Rai *et al.,* 2017b).

Group analysis (GA) was one of the methodologies used in this study to group and then prioritize the watersheds. It was useful to relate watersheds that share the same characteristics based on their geomorphometric parameters. The groups delineated by the analysis have a unique combination in terms of their geomorphometric attributes (Tritsch *et al.,* 2016). The groups of watersheds follow a territorial pattern. Group 1 includes watersheds located in the zones with the least slopes, where the predominant economic activity is agriculture. Group 2 and 3 belong to watersheds located in transition zones because of their moderate slopes. Groups 4 and 5 are watersheds with rugged topography, with vegetation of shrublands and oak forests.

The compound parameter (*Cp*) was the second methodology employed to prioritize the watersheds. The value of *Cp* was calculated for each of the 31 watersheds composing the CRB (Figure 5a). Based on the value of *Cp,* watershed RH24Hf received the highest priority (1), followed by the watershed RH24Le (2). By comparing the results from GA and *Cp*, Group 4 was identically integrated by the two methodologies. This group is characterized by watersheds having the highest average values of maximum and minimum height, elongation ratio, elongation index, mean watershed slope, slope of the main channel, and unit shape factor. The high values of these parameters correspond to watersheds with a high erosion susceptibility (Rai *et al*., 2017). Conversely, Group 5 was formed by watersheds having the highest values of main channel length, watershed length, contour length, all channel lengths, Gravelius compactness coefficient, shape factor, massivity coefficient, mean distance, Kirpich concentration time, sinuosity of the main channel, average peak flow, texture ratio, river frequency and resistance number. This coincide with high values of *Cp*, which correspond to watersheds distributed in the southwestern zone of the study area and may have a low erosion susceptibility (Patel *et al*., 2013).

The two prioritization schemes used in this study gave similar results according to the spatial distribution of watershed groups. The prioritization of watersheds, obtained through GA and *Cp*, highlighted those watersheds with potential for the implementation of soil and water conservation practices. Based on the ANOVA analyses performed to statistically compare the GA and *Cp* methodologies, the former resulted in more effectively classifying the watersheds, since it permitted better differentiating the watershed groups.

Results from the GA show that erosion susceptibility is strongly related to linear parameters (surface hydrology) for southwestern watersheds, where steep slopes of both the watershed and the main channel influence soil erosion (Tucker and Bras, 1998). Watersheds RH24Lg, RH24Le, RH24Lf, RH24Mc, RH24Lb, RH24Nc, and RH24Ne have the steepest slopes, making them more prone to erosion (Youssef *et al*., 2011).

One of the advantages of using the watershed as a territorial unit is the analysis of multiple geomorphometric parameters, which are related to the watershed's hydrological configuration, topography, and shape. Most of the watershed surface attributes depend on local topographic conditions (Biswas *et al*., 1999). In this study, the Basin's altitudinal gradient, a surface attribute, assists in exhibiting the contrasts among watersheds groups, while showing a homogeneous geographic distribution within them. The linear and shape-type parameters are important because of their influence on soil erosion.

The description and spatial grouping of the 31 watersheds through their 26 parameters using multivariate techniques proved to be useful to understand the main factors that control the variance in the CRB. Prioritization through the two types of grouping was also effective in detecting those watersheds susceptible to erosion. The proposed methodology for prioritizing watersheds on a geospatial basis is a feasible approach for identifying watersheds that are susceptible to erosion. However, prioritization with parameters that are based on shape, linear, and relief of the watersheds may not be sufficient. Thus, the incorporation of information regarding the activities on the territory of the CRB would help to improve the efficacy of the classification of watersheds based on their erosion susceptibility. Therefore, future research could include socioeconomic attributes that contribute to soil loss, such as agriculture (Miranda *et al*., 1996). Despite the limitations of this study, the contribution of this work represents an advance in the identification of the watersheds that are most susceptible to erosion in the CRB. This in turn contributes to land-use planning, which may help mitigate soil degradation processes.

CONCLUSIONS AND RECOMMENDATIONS

The application of GA and *Cp* methodologies allowed integrating a large set of geomorphometric parameters, which served to classify watersheds according to their characteristics.

GA more effectively clustered the watersheds of the Conchos River Basin than *Cp*, since the groups formed by GA were more differentiated based on the analysis of the watersheds' geomorphometric parameters. The results of GA show that watersheds RH24Lf, RH24Lb, RH24Nc, and RH24Jb might be subjected to strong erosion processes, and are potential candidates to be subjected to soil conservation practices.

The present study demonstrates the usefulness of integrating GIS and multivariate techniques to prioritize watersheds based on their erosion susceptibility. Such an integration approach showed the spatial relationships of the different geomorphometric parameters analyzed. Although the present study permitted a definition of watershed groups according to the values of their geomorphometric parameters and their relation with erosion susceptibility, the integration of additional variables in the analysis may provide a more insightful classification and thus a more reliable watershed prioritization. Such variables could include land use/land cover, soil type, lithology, geomorphology, and socioeconomic activities, among others.

REFERENCES

Aboites-Aguilar, L. 2002. Hacia una historia del río Conchos; CIESAS, Juárez: Chihuahua, México.

Adediran, A.O., I. Parcharidis, M. Poscolieri, and K. Pavlopoulos. 2004. Computer-assisted discrimination of morphological units on north-central Crete (Greece) by applying multivariate statistics to local relief gradients. Geomorphology 58:357–370.

Adhami, M., and S.H. Sadeghi. 2016. Sub-watershed prioritization based on sediment yield using game theory. J. Hydrol. 541:977–987.

Agarwal, K.K., R. Bali, M.G. Kumar, P. Srivastava, and P.V. Singh. 2009. Active tectonics in and around Kimin-Ziro area, Lower Subansiri District, Arunachal Pradesh, NE India. Z. Geomorphol. 53:109–120.

Alexakis, D.D., D.G. Hadjimitsis, and A. Agapiou. 2013. Integrated use of remote sensing, GIS and precipitation data for the assessment of soil erosion rate in the catchment area of “Yialias” in Cyprus. Atmos. Res. 131:108–124.

Altaf, S., G. Meraj, and S. Ahmad Romshoo. 2014. Morphometry and land cover based multicriteria analysis for assessing the soil erosion susceptibility of the western Himalayan watershed. Environ. Monit. Assess. 186:8391–8412.

Bach, K., M. Schawe, S. Beck, G. Gerold, S.R. Gradstein, and R. Moraes. 2003. Vegetación, suelos y clima en los diferentes pisos altitudinales de un bosque montano de Yungas, Bolivia: Primeros resultados. Ecol. Boliv. 38:3-14.

Batayneh, A., and T. Zumlot. 2012. Multivariate statistical approach to geochemical methods in water quality factor identification; application to the shallow aquifer system of the Yarmouk basin of north Jordan. Res. J. Environ. Earth Sci. 4:756–768.

Biswas, S., S. Sudhakar, and V.R. Desai. 1999. Prioritisation of subwatersheds based on morphometric analysis of drainage basin: A remote sensing and gis approach. J. Indian Soc. Remote Sens. 27:155–166.

Bro, R., and A.K. Smilde. 2014. Principal component analysis. Anal. Methods 6:2812.

Castillo-Rodríguez, M., J. López-Blanco, and E. Muñoz-Salinas. 2010. A geomorphologic GIS-multivariate analysis approach to delineate environmental units, a case study of La Malinche volcano (central México). Appl. Geogr. 30:629–638.

Chen, T., R.Q. Niu, P.X. Li, L.P. Zhang, and B. Du. 2011. Regional soil erosion risk mapping using RUSLE, GIS, and remote sensing: A case study in Miyun Watershed, North China. Environ. Earth Sci. 63:533–541.

Clarke, J.I. 1966. Morphometry from Maps. Essays in Geomorphology; Elsevier Publishing Co.: New York, NY, USA. 235–274.

CONABIO (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad). Portal de Geoinformación. Available online: http://www.conabio.gob.mx/informacion/gis/ (accessed on 29 July 2019).

CONABIO (Comisión Nacional para el Conocimiento y Uso de la Biodiversidad). 2014. La biodiversidad en Chihuahua: Estudio de Estado, 1st ed.; Comisión Nacional para el Conocimiento y Uso de la Biodiversidad: Méx, México. 23–25.

CONAGUA (Comisión Nacional del Agua). 2001. Comisión de cuenca del río Conchos. Diagnóstico Versión 1.0; Gerencia Estatal Chihuahua: Chihuahua, México.

Darabi, H., K. Shahedi, K. Solaimani, and M. Miryaghoubzadeh. 2014. Prioritization of subwatersheds based on flooding conditions using hydrological model, multivariate analysis and remote sensing technique. Water Environ. J. 28:382–392.

Dillon, W.R., and M. Goldstein. 1984. Multivariate Analysis. Methods and Applications; John Wiley & Sons Inc.: New York, NY, USA. 1–608.

Eder, B.K., J.M., Davis, and P. Bloomfield. 1994. An Automated Classification Scheme Designed to Better Elucidate the Dependence of Ozone on Meteorology. J. Appl. Meteorol. 33:1182–1199.

Fajardo, J., J. Lessmann, E. Bonaccorso, C. Devenish, and J. Munoz. 2014. Combined Use of Systematic Conservation Planning, Species Distribution Modelling, and Connectivity Analysis Reveals Severe Conservation Gaps in a Megadiverse Country (Peru). PLoS ONE, 9, e114367.

Faniran, A. 1968. The Index of Drainage Intensity—A Provisional New Drainage Factor. Aust. J. Sci. 31:328–330.

Farhan, Y., A. Anbar, N. Al-Shaikh, and R. Mousa. 2017. Prioritization of semi-arid agricultural watershed using morphometric and principal component analysis, remote sensing, and GIS techniques, the Zerqa River Watershed, Northern Jordan. Agric. Sci. 8:113–148.

Gajbhiye, S., S.K. Sharma, and M.K. Awasthi. 2015. Application of Principal Components Analysis for Interpretation and Grouping of Water Quality Parameters. Int. J. Hybrid Inf. Technol. 8:89–96.

Ganasri, B., H. Ramesh. 2016. Assessment of soil erosion by RUSLE model using remote sensing and GIS—A case study of Nethravathi Basin. Geosci. Front. 7:953–961.

Gavit, B.K., R.C. Purohit, P.K. Singh, M.K. Kothari, and H.K. Jain. 2016. Grouping of Significant Geomorphic Parameters using Multivariate Parameters using Multivariate. Res. J. Recent Sci. 5:32–38.

Gutiérrez, M., E. Carreón-Hernández. 2004. Salinidad en el bajo Rio Conchos: Aportes y tendencias. Terra Latinoam. 22:499–506.

Hamdy, O., S. Zhao, M.A. Salheen, and Y.Y. Eid. 2016. Identifying the Risk Areas and Urban Growth by ArcGIS-Tools. Geosciences. 6:47.

Harlin, J.M., and C. Wijeyawickrema. 1985. Irrigation and groundwater depletion in caddo county, oklahoma. JAWRA J. Am. Water Resour. Assoc. 21:15–22.

Horton, R.E. 1945. Erosional development of streams and their drainage basins: A hydrophysical approach to quantitative morphology. Geol. Soc. Am. Bull. 56:275–370.

INEGI (Instituto Nacional de Estadística, Geografía e Informática). 2014. Guía para la Interpretación de la Cartografía uso del Suelo y Vegetación, Escala 1:250,000 Serie V; INEGI: Aguascalientes, México. 30. Available online: https://www.inegi.org.mx/contenidos/temas/mapas/usosuelo/metadatos/guia_interusosuelov.pdf (accessed on 31 July 2019).

INEGI (Instituto Nacional de Estadística, Geografía e Informática). Sistema de descarga del Continuo de Elevaciones Mexicano. Available online: https://www.inegi.org.mx/app/geo2/elevacionesmex/ (accessed on 29 July 2019).

Javed, A., M.Y. Khanday, and R. 2009. Ahmed, Prioritization of sub-watersheds based on morphometric and land use analysis using remote sensing and GIS techniques. J. Indian Soc. Remot 37:261–274.

Johnson, R.A., and D.W. Wichern. 2007. Applied Multivariate Statistical Analysis, 6th ed.; Pearson Prentice-Hall Inc.: Upper Saddle River, NJ, USA. 296–301.

Jolliffe, I.T. 2002. Principal Component Analysis, 2nd ed.; John Wiley & Sons Inc.: Charlottesville, VA, USA. 150–166.

Keesstra, S., P. Pereira, A. Novara, E.C. Brevik, C. Azorin-Molina, L. Parras-Alcántara, A. Jordán, and A. Cerdà. 2016. Effects of soil management techniques on soil water erosion in apricot orchards. Sci. Total Environ. 551:357–366.

Khosravi, K., B.T. Pham, K. Chapi, A. Shirzadi, H. Shahabi, I. Revhaug, I. Prakash, and D.T. Bui. 2018. A comparative assessment of decision trees algorithms for flash flood susceptibility modeling at Haraz watershed, northern Iran. Sci. Total Environ. 627:744–755.

Kirpich, Z.P. 1940. Time of concentration of small agricultural watersheds. Civ. Eng. 10:362.

Kompani-Zare, M., M. Soufi, H. Hamzehzarghani, and M. Dehghani. 2011. The effect of some watershed, soil characteristics and morphometric factors on the relationship between the gully volume and length in Fars Province, Iran. Catena 86:150–159.

Luo, W. 2000. Quantifying groundwater-sapping landforms with a hypsometric technique. J. Geophys. Res. Space Phys. 105:1685–1694.

Mahlknecht, J., A. Horst, R. Aravena, G. Hernández-Limón, and G. Hernández-Limón. 2008. Groundwater geochemistry of the Chihuahua City region in the Rio Conchos Basin (northern Mexico) and implications for water resources management. Hydrol. Process. 22:4736–4751.

Makwana, J., and M.K. Tiwari. 2016. Prioritization of agricultural sub-watersheds in semiarid middle region of Gujarat using remote sensing and GIS. Environ. Earth Sci. 75:137.

Malik, A., A. Kumar, D.P. Kushwaha, O. Kisi, S.Q. Salih, N. Al-Ansari, and Z.M. Yaseen. 2019. The Implementation of a Hybrid Model for Hilly Sub-Watershed Prioritization Using Morphometric Variables: Case Study in India. Water 11:1138.

Meshram, S.G., and S.K. Sharma. 2017. Prioritization of watershed through morphometric parameters: A PCA-based approach. Appl. Water. Sci. 7:1505–1519.

Miller, V.C. 1953. A Quantitative Geomorphic Study of Drainage Basin Characteristics in the Clinch Mountain Area, Virginia and Tennessee; Office of Naval Research (USA): Arlington, VA, USA. 1–51.

Miranda, J., E. Andrade, A. López-Suárez, R. Ledesma, T.Cahill, and P. Wakabayashi. 1996. A receptor model for atmospheric aerosols from a southwestern site in Mexico City. Atmos. Environ. 30:3471–3479.

Montero-Martínez, M.J., and Ó.F. Ibáñez-Hernández. 2017. La cuenca del río Conchos: Una mirada desde las ciencias ante el cambio climático, 1st ed.; Instituto Mexicano de la Tecnología del Agua: Morelos, Mexico. 10.

Mueller, J.E. 1968. An introduction to the hydraulic and topographic sinuosity indexes 1. Ann. Assoc. Am. Geogr. 58:371–385.

Nautiyal, M.D. 1994. Morphometric analysis of a drainage basin using aerial photographs: A case study of Khairkuli basin District Deharadun. J. Indian Soc. Remote Sens. 22:251–262.

Nooka Ratnam, K. 2005. Check Dam positioning by prioritization ofmicro watershed using SYI model and morphometric analysis—Remote sensing and GIS perspective. J. Indian Soc. Remote Sens. 33:25–38.

Núñez López, D., E.J. Treviño Garza, V.M. Reyes Gómez, C.A. Muñoz Robles, O.A. Aguirre Calderón, and J. Jiménez Pérez. 2014. Uso de modelos de regresión para interpolar espacialmente la precipitación media mensual en la cuenca del río Conchos. Rev. Mex. Cienc. Agrícolas 5:201–213.

Oketola, A.A., S.M. Adekolurejo, and O. Osibanjo. 2013. Water Quality Assessment of River Ogun Using Multivariate Statistical Techniques. J. Environ. Prot. 4:466–479.

Ordoñez, R. 2017. Modelación de la erosión hídrica en la Cuenca Alta del Río Conchos, Chihuahua. Master's Thesis, Universidad Autónoma de Chihuahua, Chih, Mexico.

Patel, D.P., C.A. Gajjar, and P.K. Srivastava. 2013. Prioritization of Malesari mini-watersheds through morphometric analysis: A remote sensing and GIS perspective. Environ. Earth Sci. 69:2643–2656.

Paul, I.I., and E.N. Bayode. 2012. Watershed Characteristics and Their Implication for Hydrologic Response in the Upper Sokoto Basin, Nigeria. J. Geogr. Geol. 4:147–155.

Prieto-Amparán, J.A., B.A. Rocha-Gutiérrez, M.D.L. Ballinas-Casarrubias, M.C. Valles-Aragón, M.D.R. Peralta-Pérez, and A. Pinedo-Alvarez. 2018. Multivariate and Spatial Analysis of Physicochemical Parameters in an Irrigation District, Chihuahua, Mexico. Water. 10:1037.

Prieto-Amparán, J.A., E. Santellano-Estrada, F. Villarreal-Guerrero, M. Martinez-Salvador, A. Pinedo-Alvarez, G. Vázquez-Quintero, M.C. Valles-Aragón, and C. Manjarrez-Domínguez. 2019. Spatial Analysis of Temperate Forest Structure: A Geostatistical Approach to Natural Forest Potential. Forests 10:168.

Rahaman, S.A., S.A. Ajeez, S. Aruchamy, and R. Jegankumar. 2015. Prioritization of Sub Watershed Based on Morphometric Characteristics Using Fuzzy Analytical Hierarchy Process and Geographical Information System—A Study of Kallar Watershed, Tamil Nadu. Aquat. Procedia 4:1322–1330.

Rai, P.K., P.K. Chaubey, K. Mohan, and P. Singh. 2017a. Geoinformatics for assessing the inferences of quantitative drainage morphometry of the Narmada Basin in India. Appl. Geomat. 9:167–189.

Rai, P.K., V.N. Mishra, and K. Mohan. 2017b. A study of morphometric evaluation of the Son basin, India using geospatial approach. Remote Sens. Appl. Soc. Environ. 7:9–20.

Rawat, K.S., V.K. Tripathi, and A.K. Mishra. 2014. Sediment yield index mapping and prioritization of Madia subwatershed, Sagar District of Madhya Pradesh (India). Arab. J. Geosci. 7:3131–3145.

Robinne, F.N., K.D. Bladon, U. Silins, M.B. Emelko, M.D. Flannigan, M.A. Parisien, X. Wang, S.W. Kienzle, and D.P. Dupont. 2019. A Regional-Scale Index for Assessing the Exposure of Drinking-Water Sources to Wildfires. Forests 10:384.

Saha, S., and S. Mukhopadhyay. 2012. Determination of Significant Geomorphic Parameters of Kunur River Basin through Principal Component Analysis. J. Eng. Comput. Appl. Sci. 4:232–243.

Sánchez, A. 1944 Water Treaty between Mexico and the United States: Present situation and future potential. Front. Norte 2006, 18:125–144.

Sanderson, M., M. Santini, R. Valentini, and E. Pope. 2012. Relationships between Forests and Weather. EC Directorate General of the Environment; Met Office: Devon, UK. Available online: http://ec.europa.eu/environment/forests/pdf/EU_Forests_annex1.pdf (accessed on 29 July 2019).

Schumm, S.A. 1956. Evolution of drainage systems and slopes in badlands at perth amboy, new jersey. GSA Bull. 67:597.

Şen, Z. 2008. Modified hydrograph method for arid regions. Hydrol. Process. 22:356–365.

Sharma, S.K., G.S. Rajput, S. Tignath, and R.P. Pandey. 2010. Morphometric analysis of a watershed using GIS. J. Indian Water Res. Soc. 30:33–39.

Sharma, S.K., S. Gajbhiye, and S. Tignath, 2014. Application of principal component analysis in grouping geomorphic parameters of a watershed for hydrological modeling. Appl. Water Sci. 5:89–96.

Sharma, S.K., S. Gajbhiye, R.K. Nema, and S. Tignath. 2014. Assessing vulnerability to soil erosion of a watershed of tons River basin In Madhya Pradesh using Remote sensing and GIS. Int. J. Environ. Res. Dev. 4:153–164.

Sharma, S.K., S. Tignath, S. Gajbhiye, and R. Patil. 2013. Application of principal component analysis in grouping geomorphic parameters of Uttela watershed for hydrological modeling. Int. J. Remote Sens. Geosci. 2:63–70.

Shrimali, S., S. Aggarwal, and J. Samra. 2001. Prioritizing erosion-prone areas in hills using remote sensing and GIS—A case study of the Sukhna Lake catchment, Northern India. Int. J. Appl. Earth Obs. Geoinf. 3:54–60.

SIATL. Simulador de Flujos de Agua de Cuencas Hidrográficas. Available online: http://antares.inegi.org.mx/analisis/red_hidro/siatl/ (accessed on 29 July 2019).

Singh, P., and A. Gupta, M. Singh. 2014. Hydrological inferences from watershed analysis for water resource management using remote sensing and GIS techniques. Egypt. J. Remote Sens. Space Sci. 17:111–121.

Singh, P.K., V. Kumar, R.C. Purohit, M. Kothari, and P.K. Dashora. 2009, Application of principal component analysis in grouping geomorphic parameters for hydrologic modeling. Water Resour. Manag 23:325.

Strahler, A.N. 1964. Quantitative geomorphology of drainage basins and channel networks. In Handbook of Applied Hydrology, 1st ed.; Chow, V.T., Ed.; McGraw-Hill: NY, USA. 1:4–39.

Subyani, A.M., M.H. Qari, and M.I. Matsah, 2012. Digital elevation model and multivariate statistical analysis of morphometric parameters of some wadis, western Saudi Arabia. Arab J. Geosci. 5:147–157.

Suji, V.R., R.V. Sheeja, S. Karuppasamy, V.R. Suji, R.V. Sheeja, and S. Karuppasamy. 2015. Prioritization using Morphometric Analysis and Land use/Land Cover Parameters for Vazhichal Watershed using Remote Sensing and GIS Techniques. Int. J. Innov. Res. Sci. Tech. 1:2349–6010.

Thakker, A.K., and SP. Dhiman. 2007. Morphometric analysis and prioritization of miniwatersheds in Mohr watershed, Gujarat using Remote sensing and GIS techniques. J. Indian Soc. Remote Sens. 35:313–321.

Tilahun, S.A., C.D. Guzman, A.D. Zegeye, E.K. Ayana, A.S. Collick, B. Yitaferu, and T.S. Steenhuis. 2014. Spatial and temporal patterns of soil erosion in the semi-humid Ethiopian highlands: A case study of Debre Mawi watershed. In Nile River Basin, 1st ed.; Assefa, M., Melesse, W., Abtew Shimelis, G., Eds.; Springer: Cham, Switzerland. 149–163.

Tritsch, I., P. Sist, I.D.S. Narvaes, L. Mazzei, L. Blanc, C. Bourgoin, G. Cornu, and V. Gond. 2016. Multiple Patterns of Forest Disturbance and Logging Shape Forest Landscapes in Paragominas, Brazil. Forests, 7:315.

Tucker, G.E., and R.L. Bras. 1998. Hillslope processes, drainage density, and landscape morphology. Water Resour. Res. 34:2751–2764.

Viramontes-Olivas, O.A., L.F. Escoboza-García, C. Pinedo-Alvarez, A. Pinedo-Alvarez, V. Reyes Gómez, J. Román Calleros, and A. Pérez-Márquez. 2008. Morfometría de la cuenca del río San Pedro, Conchos, Chihuahua. Tecnociencia Chihuah. 1:21–31.

Viramontes-Olivas, O.A., V.M. Reyes-Gómez, L.F. Escoboza-García, J.A. Román-Calleros, A. Pérez-Márquez, C. Pinedo-Álvarez, P.B. Sánchez-Fernández, and N. Miranda Ojeda. 2008. Hidrología de los suelos de las cuencas media y baja del Río Conchos, Chihuahua, México. Rev. Latinoam. Recur. Nat. 4:31–45.

Ward, J.H. 1963. Hierarchical Grouping to Optimize an Objective Function. J. Am. Stat. Assoc. 58:236–244.

Welde, K. 2016. Identification and prioritization of subwatersheds for land and water management in Tekeze dam watershed, Northern Ethiopia. Int. Soil Water Conserv. Res. 4:30–38.

Yidana, S.M., B. Banoeng-Yakubo, and T.M. Akabzaa. 2010. Analysis of groundwater quality using multivariate an spatial analyses in the Keta basin, Ghana. J. Afr. Earth Sci. 58, 220–234.

Youssef, A.M., B. Pradhan, and A.M. Hassan. 2011. Flash flood risk estimation along the St. Katherine road, southern Sinai, Egypt using GIS based morphometry and satellite imagery. Environ. Earth Sci. 62:611–623.

Yunus, A.P., T. Oguchi, and Y.S. Hayakawa. 2014. Morphometric Analysis of Drainage Basins in the Western Arabian Peninsula Using Multivariate Statistics. Int. J. Geosci 5:527–539.

Zavoianu, I. 1985. Morphometry of Drainage Bassins, 2nd ed.; Editura Academiei: Bucharest, Romania; Elsevier Science Publishers: Amsterdam, The Netherlands. 9–235.

STUDY IV. SPATIAL ANALYSIS OF TEMPERATE FOREST STRUCTURE: A GEOSTATISTICAL APPROACH TO NATURAL FOREST POTENTIAL

BY:

Jesús A. Prieto-Amparan[1], Eduardo Santellano-Estrada[1], Federico Villarreal-Guerrero[1], Martin Martinez-Salvador[1], Alfredo Pinedo-Alvarez[1], Griselda Vázquez-Quintero[2], María Cecilia Valles-Aragón[2], Carlos Manjarrez-Domínguez[2]*

[1]Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada km 1., C.P. 31453. Chihuahua, Chih. México.
[2]Facultad de Ciencias Agrotecnológicas, Universidad Autónoma de Chihuahua, s/n., C.P. 31350. Chihuahua, Chih., México.

*Corresponding author: Carlos Baduel Manjarrez Dominguez; Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada km 1., C.P. 31453. Chihuahua, Chih. México. manjarrez.carlos@gmail.com, tel. +52-(614)-4-34-03-63. DOI: https://doi.org/10.3390/f10020168.

ABSTRACT

SPATIAL ANALYSIS OF TEMPERATE FOREST STRUCTURE: A GEOSTATISTICAL APPROACH TO NATURAL FOREST POTENTIAL

BY:

M. C. JESUS ALEJANDRO PRIETO AMPARAN

Forest ecosystems represent an important means of ecosystem services; they are key as carbon sinks, water collectors, soil stabilizers, suppliers of great biological diversity, among other benefits. In addition, regionalization based on forest conditions provides a valuable approach to understanding and analyzing spatial patterns, which is useful as a tool for implementation of forest ecosystem protection and conservation programs. In this research, the structure of a temperate forest in the western Sierra Madre region of Mexico was analyzed and characterized. The study unit was the watershed and the analysis used a geospatial approach combined with multivariate techniques such as: Principal Component Analysis, Cluster Analysis (CA), Discriminant Analysis (DA) and Multivariate Analysis of Variance. We evaluated the relationships between spectral satellite data, thematic maps and structural forest variables. A total of 345 watersheds were grouped based on these variables. The grouping of watersheds under low, medium and high production conditions was carried out with CA, defining 3 groups. The validation of the grouping was performed through DA, estimating errors with the restitution method, as well as with the cross-validation method. Significant differences were found among the groups. The grouping of watersheds provides observable evidence of the variability of the forest condition throughout the area. This study allows the identification of forest areas with

different levels of productivity and can help to identify levels of vulnerability and ecological fragility in natural forests in temperate zones.

Keywords: Multivariate analysis; watershed; Landsat; classification; GIS; forest structure; regionalization; spatial clustering.

RESUMEN

SPATIAL ANALYSIS OF TEMPERATE FOREST STRUCTURE: A GEOSTATISTICAL APPROACH TO NATURAL FOREST POTENTIAL

POR:

M. C. JESÚS ALEJANDRO PRIETO AMPARÁN

Doctor en Philosophia de Recursos Naturales

Secretaría de Investigación y Posgrado

Facultad de Zootecnia y Ecología

Universidad Autónoma de Chihuahua

Presidente: Dr. Alfredo Pinedo Alvarez

Los ecosistemas forestales representan un medio importante para el sustento de servicios ecosistemicos; son clave como sumideros de carbono, recolectores de agua, estabilizadores de suelos, proveedores de gran diversidad biológica, entre otros beneficios. Además, la regionalización basada en las condiciones de los bosques proporciona un enfoque valioso para comprender y analizar los patrones espaciales, que es útil como herramienta para la implementación de programas de protección y conservación de los ecosistemas forestales y sus servicios ecosistemicos que susttentan. En esta investigación se analizó y caracterizó la estructura de un bosque templado en la región occidental de la Sierra Madre de México. La unidad de estudio fue la cuenca hidrográfica y el análisis utilizó un enfoque geoespacial combinado con técnicas multivariantes tales como el análisis de componentes principales (ACP), análisis de grupos (AG), análisis discriminante (AD) y análisis multivariado de la varianza (AMVA). Se evaluaron las relaciones entre los datos de de imagenes espectrales, mapas temáticos y variables forestales estructurales. Un total de 345 cuencas fueron agrupadas con base en estas variables. La agrupación de cuencas bajo condiciones de baja, media y alta producción se realizó con el AG, definiendo 3 grupos. La validación de la agrupación se realizó a través de AD, estimando los errores con el método de restitución, así como con el método de validación cruzada. Se encontraron diferencias significativas entre los grupos. La agrupación de las cuencas hidrográficas proporciona evidencia observable de la variabilidad de la condición del bosque en toda el área. Este estudio permite identificar áreas forestales con diferentes niveles de productividad y puede ayudar a identificar los niveles de vulnerabilidad y fragilidad ecológica en los bosques naturales de las zonas templadas.

Palabras clave: Analisis multivariado; cuenca; Landsat; clasificación; SIG; estructura del bosque; regionalización; agrupamiento espacial.

INTRODUCTION

Forest ecosystems provide essential benefits for humanity, including the protection of biodiversity, climate regulation, carbon sinks, among others (Randolph *et al*., 2005; Bolton *et al*., 2017). The assessment of changes in forest ecosystems and the understanding of their causes are of great concern. Factors that can alter the forest structure such as fire, pests, human activities related to the settlement or opening of agricultural land and the extraction of resources, generate loss of biomass that influences biogeochemical cycles (Hilker *et al*., 2009). Climatic characteristics determine the health of the world´s forest ecosystems (Sáenz *et al*., 2010). With a continental area of approximately two million square kilometers, Mexico is one of the richest countries in biodiversity (Groombridge and Jenkins, 2000). The temperate forests of Mexico cover approximately 32 million hectares, equivalent to 18 % of the territory. In these ecosystems we can find a great diversity of associations between pines and oaks that are not present in another part of the world (González *et al*., 2012). These forests are mainly composed of conifers and broad-leaved trees, where pine trees (*Pinus spp*.), Oyameles (*Abies spp*.), Pinabetes (*Picea spp*., *Pseudotsuga spp*.) and Oak (*Quercus spp*.) are the dominant species in their composition and structure, predominating in the temperate zones of the main mountainous regions of the country (Challenger, 1998).

However, Mexico is also the country with the highest rates of deforestation worldwide (Velázquez *et al*., 2002). Many of its native ecosystems are gradually being reduced to small remnants of the original vegetation (Aguilar *et al*., 2015). To ensure the sustainability of forests, we need to characterize their structure

(Jupp, 1997). This implies the understanding of the different variables associated with the structure of forests (Ludwig *et al*., 1997). Forest structural variables, such as volume or biomass, average diameter and height, are important data needed to assess forest resources (Achard *et al*, 2006; Bonan, 2008). Due to the extensive surface of temperate forests that are distributed in Mexico, together with a very rugged orography, only a portion of the forest can be sampled (Chambers *et al*., 2007). Under these conditions, it is important to determine the spatial and temporal distribution of the forest structure. These actions allow classifying the forest coverage according to its natural potential (Zhu *et al*., 2012).

The forest condition is defined in this study as the related set of forest characteristics determined by the number of trees per hectare, their spectral response, their distance from anthropogenic activities, among others. A good condition would present a high number of trees per hectare, a low spectral response and high distance from anthropogenic activities. At the regional level, a feasible procedure to monitor the forest condition on a regular and continuous basis is based on satellite data, as well as field information (Franklin, 2001). Satellite image data represent a source of key information in the monitoring of the forest condition (Wang *et al*., 2013). The measurement of the forest structure through remote sensing reduces costs by covering large areas of land (Pflugmacher *et al*., 2012). In addition, the analyzed data serve to determine the health status of the forests (Ingram *et al*., 2005). There is a great variety of studies, where satellite images have been used, as well as field data to characterize the structure of the forest (Wolter *et al*., 2009; Powell *et al*., 2010; Soenen *et al*., 2010; Dube and Mutanga, 2015). Many of them explain the relationship of the forest structure with the spectral satellite data. Traditionally, the measurement of the forest structure considers the forest's dasometric variables. However, there are environmental variables that influence the productivity of the forest and represent an important element in forest management (Bach *et al*., 2003; Sanderson et al., 2012). A large number of biophysical, biological, topographic and anthropogenic variables are associated with forest productivity (Vázquez-Quintero *et al*., 2016).

To identify these different forest conditions and variability at the regional level, regionalization or spatial grouping of data is applied. The process of delimitation and classification of areas with homogeneous characteristics according to their environmental properties distributed in the forest landscape allows the definition of environmental units on a spatial scale. In relation to the above, the classification, modeling and interpretation of the monitored data are the most important steps in the evaluation of the forest structure. The spatial identification of regions with homogeneous characteristics often lacks an accepted and clearly articulated theoretical basis. Regions are typically delineated by expert criteria, which is sometimes subjective. The spatial continuity of the resulting regions is rarely managed quantitatively. In this sense, geostatistical multivariate techniques allow us to study the spatial variability of forests, forest structure, topographic and biophysical factors, as well as the relationships that may exist among these variables. Cluster Analysis (CA), Principal Component Analysis (PCA) and Discriminant Analysis (DA) have been used in recent years to analyze environmental variables and extract significant information (Pasher and King, 2010; Bhuiyan *et al*., 2011; Batayneh and Zumlot, 2012; Oketola *et al*., 2013). The use of these techniques allows defining new variables or groups that provide information on the spatial variability of forests, (Singh *et al*., 2004). There are few studies that integrate variables in the forest structure using remote sensing, biophysical, topographic or proximity (roads, communities) data and their interaction (Kalabokidis *et al*., 2007; Verdú *et al*., 2012; Tritsch *et al*., 2016). Such studies have been carried out mainly in developing countries like Mexico, where there are large gaps of information that can be reinforced with the integration of different sources.

The objective of this study was to analyze the forest condition through environmental, topographic, proximity and spectral variables, employing multivariate techniques with a geospatial approach in a forest area located in the municipality of ‘Guadalupe y Calvo’, Chihuahua, Mexico.

MATERIALS AND METHODS

Study Area

The study area is located in the municipality of Guadalupe y Calvo, in the state of Chihuahua, Mexico, between parallels 107°00' W–26°30' N and 106°30' W–26°00' N (Figure 1). The property has an area of 113.73 ha, of which 95.13 ha are temperate forests (15 % of the total forest area of the municipality); 42.60 ha are exploitable pine and oak woods (Torres *et al*., 2014). The region is characterized by its high number of endemic species, estimated at around 4000 plant species (Felger and Johnson, 1995). The area is one of the forest regions with the highest biodiversity, which supports various environmental services for the region. It has a unique and wide system of deep canyons, resulting in a mixture of temperate and tropical ecosystems (CONABIO, 2014). The main land uses of the property are coniferous and hardwood forests. Forestry is the main productive activity, representing 42 % of the income for the inhabitants. It includes the extraction of wood and the use of dead wood (Torres, 2012). The second productive activity is livestock for self-consumption, which represents 24 % of the economy (INEGI, 2003).

Forest Data

The data were collected between January and May of 2014 and comprised four thousand 0.1-ha circular ground survey plots (Figure 2). Sample measurements included tree height recorded using digital hypsometer and tree diameter at the breast height (dbh) using a tape. In all plots, the position was derived using GPS. Canopy closure status was determined by recording the type

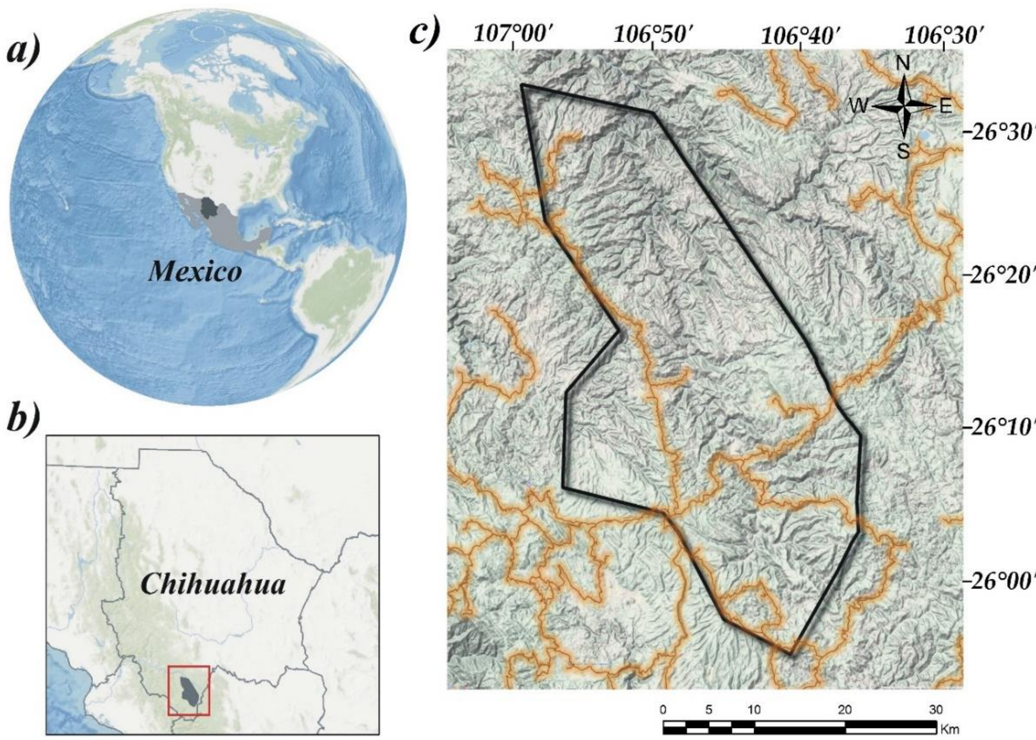

Figure 1. Geographical location of the study area: (a) Mexico, (b) “Ejido Chinatu” location in Chihuahua, Mexico, (c) “Ejido Chinatu”. Boundary of the studied area (—), principal roads (—).

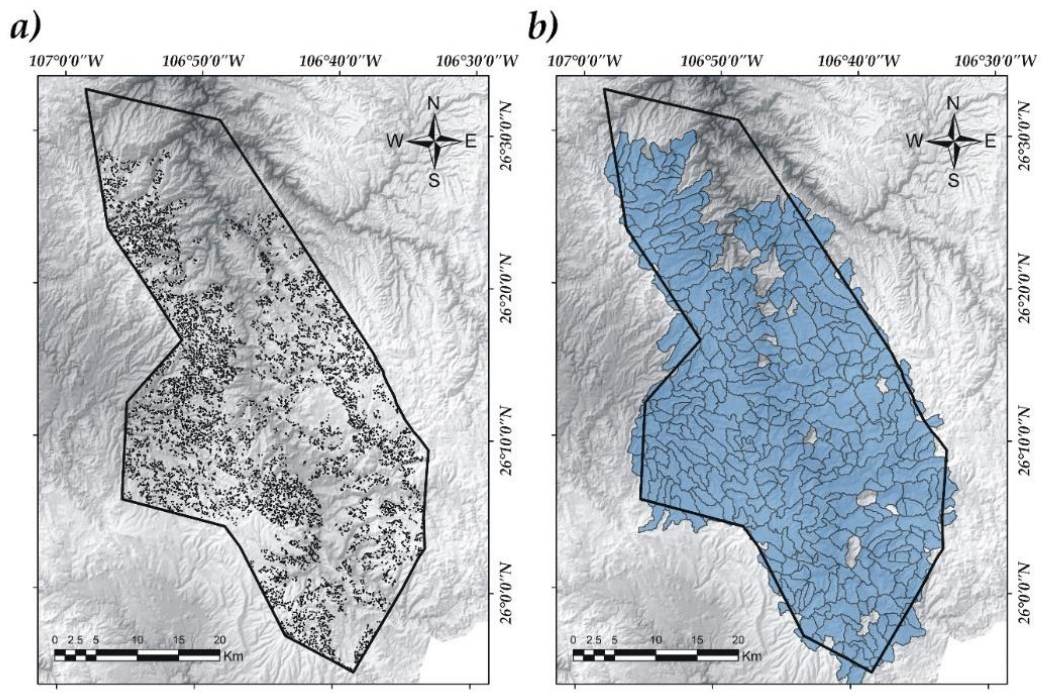

Figure 2. Sampling of structural variables of the forest (a), study unit: watershed (b).

and proportion of shrub stratum vegetation present in each sample plot using a 1-meter square quadrat split into 4 equal quadrants. Plots with more than 50 % dead vegetation on the forest floor were classified as closed canopy. Conversely, those with less than 50 % understory vegetation were classified as open canopy.

Geospatial Data

We used data from the Landsat 8 OLI satellite. The original bands were transformed into physical reflectance tools. The spectral value was extracted using a 3×3 window, to minimize the error (Hall *et al*., 2006). The Normalized Difference Vegetation Index (NDVI) and the Modified Soil Adjusted Vegetation Index 2 (MSAVI2) were generated (Equations 1 and 2).

$$NDVI = \frac{p_{NIR} - p_{red}}{p_{NIR} + p_{red}} \tag{1}$$

$$MSAVI2 = \frac{2p_{NIR} + 1 - \sqrt{(2_{p_{NIR}})^2 - 8(p_{NIR} - p_{red}}}{2} \tag{2}$$

where: p_{NIR}= near infrared, p_{red}= red band

The variables considered for the analysis were recorded in the same reference system (WGS84) and were converted to raster format with the same number of rows and columns using ArcGis 10.3© software (Environmental Systems Research Institute, Redlands, CA, USA.; https://www. esri.com/en-us/home). We used 13 variables for the analysis (Table 1). They were selected based on the relationships found in other studies (Miranda *et al*., 2012; Pérez *et al*., 2012; Bax *et al*., 2016).

The variables were measured in 345 basins, delimited within the study area (Figure 2). The size of the basins varied from 100 to 2100 ha, with an average of 424.02 ha and a standard deviation of 274.07 ha. The geographic information

Table 1. Variables used in the data analysis.

Variable	Unit	Source	Acronym
Average total volume per tree	m^3	Sampling	ATVT
Number of trees per hectare	No	Sampling	NTPH
Basal area per hectare	m^2	Sampling	BAPH
Total volume of trees per hectare	m^3	Sampling	TVTPH
Average quadratic diameter	Cm	Sampling	AQD
Spectral band 3	W/(m^2 sr µm)	USGS	SB3
Spectral band 7	W/(m^2 sr µm)	USGS	SB7
Normalized difference vegetation index	Adimensional	Own source	NDVI
Modified soil-adjusted vegetation index 2	Adimensional	Own source	MSAVI2
Distance to roads	m	Own source	DR
Distance to water bodies	m	Own source	DWB
Slope	Degrees	INEGI	Slope
Mean annual temperature	°C	CONAGUA	MAT

system (GIS) was used to delineate basins. A digital terrain model (INEGI, 2005) was used for the generation of the basins (Maidment and Djokic, 2000). The details of the procedure for the delimitation of watersheds in GIS can be found in the study by Hamdy *et al*. (2016).

Statistical Analyses

To identify the relationships between the set of variables (forest and geospatial), a correlation analysis was used in the SAS© software, version 9.1.3 (SAS, 2006). Subsequently, a PCA was used through the routine PRINCOMP with the SAS software. This analysis helped to reduce the amount of data; it was carried out using Equation 3.

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + . . . + a_{im}x_{mj} \quad (3)$$

where: *z* represents the component coefficient, a represents the weight of the component, *x* represents the measured value of the variable, *i* corresponds to the number of the component, *j* represents the number of the sample and *m* represents the total number of variables.

The basins studied were classified by ascending group analysis (AHC) using the Ward criterion (SAS program, CLUSTER rutine). Ward's criterion was applied to that of AHC using the square euclideanEuclidean distance to explore the grouping of the 345 watersheds. The group definition was determined based on the coefficient of determination (R^2), the pseudo T^2, the Cubic Clustering Criterion (*CCC*) and the pseudo F statistic (Eder *et al*., 1994).

Group formation was evaluated based on the original variables using the Discriminant Analysis (DA) (Equation 4). It was analyzed to determine whether if tthe variables contribute to discriminatione among the groups. Through this analysis, we were able to determine the association of each basin with a group. The classification error of the generated DA was made calculated with the restitution method as well as with the cross- validation method. Finally, the variance was analyzed to verify if there were significant differences between the classification and the independent variables as a whole (MANOVA, Johnson, 1988).

$$f\,(Gi) = k_i + \sum_{j=1}^{n} w_{ij}\, P_{ij} \tag{4}$$

where: $f(Gi)$ represents the number of groups, k_i represent an inherent constant for each group, n number of parameters, and w_{ij} weight of the coefficient assigned by the DA for a given parameter P_{ij}.

RESULTS AND DISCUSSION

In the correlation analysis, we found significant positive correlations among the variables (Table 2). The evaluation shows different relationship conditions for the forest in the study area; this condition allowed us to perform a differential characterization as a measure to explain the forest heterogeneity. The PCA showed that the first four components explain 71 % of the total variability provided by the 13 original variables. The values of the eigenvalue and the accumulated variability explain the proportion of the original variables (Figure 3).

Table 3 shows the characteristics of the first four components and their eigenvectors. The variables that contribute the most to the PC1 (weak coefficients) are Average total volume per tree, Average quadratic diameter, Distance to roads, Slope and Mean annual temperature that represent the forest structural and biophysical variables. The moderate coefficients of PC3 correspond to Number of trees per hectare, Basal area per hectare and Total volume of trees per hectare, which are structural variables of the forest. PC3 has moderate-high coefficients for Average total volume per tree, Average quadratic diameter, NDVI and MSAVI2, which are part of the structural and spectral forest variables. Finally, PC4 has the highest coefficients for the spectral variables Spectral band 3, Spectral band 7 and NDVI. The grouping of the sites based on the variables is presented in Figure 4. The variables can be grouped into 3: Group 1: Mean annual temperature, Distance to water bodies, Average quadratic diameter, Average total volume per tree, Slope, Group 2: Spectral band 7, MSAVI2, Distance to water bodies, Spectral band 3 and Group 3: Total volume of trees per hectare, Basal area per hectare and Number of trees per hectare.

Table 2. Matrix of different variable correlations considered in the analysis.

	ATVT	NTPH	BAPH	TVTPH	DQ	SB3	SB7	NDVI	MSAVI2	DR	DWB	Slope	MAT
ATVT	1.00												
NTPH	−0.41**	1.00											
BAPH	0.10	0.72**	1.00										
TVTPH	0.24**	0.70**	0.89*	1.00									
DQ	0.92**	−0.47**	0.11*	0.18*	1.00								
SB3	0.05	−0.06	0.29*	−0.05	0.03	1.00							
SB7	−0.17*	−0.07	−0.17*	−0.18*	−0.15*	0.35**	1.00						
NDVI	0.18*	−0.05	0.11*	0.12*	0.23**	−0.15*	−0.40**	1.00					
MSAVI2	0.01	−0.10	−0.05	−0.06	0.07	0.21*	0.58**	0.45**	1.00				
DR	0.03	−0.01	0.03	−0.01	0.08	0.11*	0.21**	0.02	0.22**	1.00			
DWB	0.19*	−0.16*	−0.04*	−0.04	0.24*	−0.05	−0.09	0.00	−0.13*	−0.02	1.00		
Slope	0.14*	−0.07	−0.01	0.01	0.10	−0.06	−0.27**	−0.13*	−0.41**	−0.24**	0.22**	1.00	
ANT	0.19*	−0.11*	−0.08	−0.04	0.09	0.03	-0.10	−0.31**	−0.37**	−0.30**	0.47**	0.61**	1.00

ATVT = Average total volume per tree, NTPH = Number of trees per hectare, BAPH = Basal area per hectare, TVTPH = Total volume of trees per hectare, AQD = Average quadratic diameter, SB3 = Spectral band 3, SB7 = Spectral band 7, NDVI = Normalized difference vegetation index, MSAVI2 = Modified soil-adjusted vegetation index 2, DR = Distance to roads, DWB = Distance to water bodies, MAT = Mean annual temperature.

* = Significant P<0.05.

** = Highly significant P<0.0001.

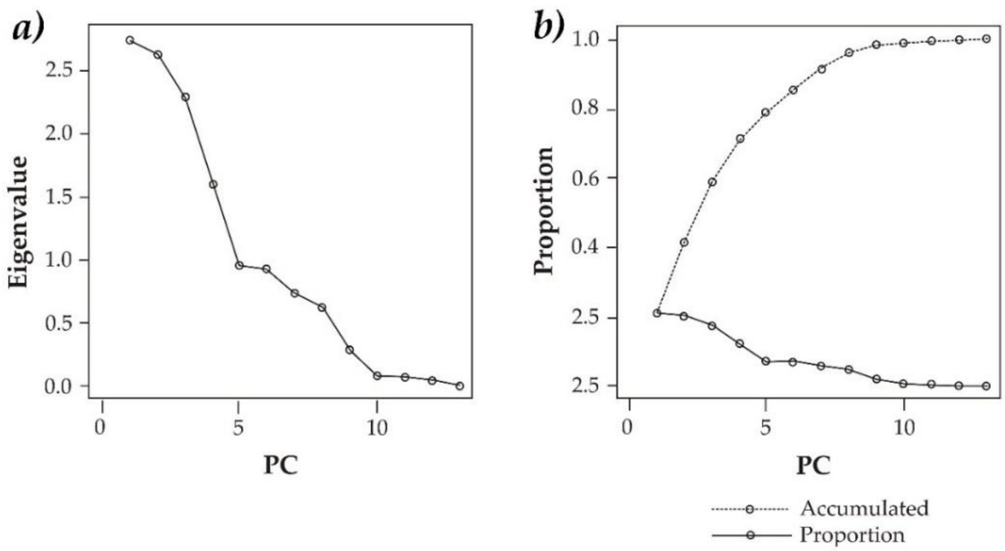

Figure 3. Plot indicating the eigenvalues and their contribution to the total variance (a) and explained variance (b).

Table 3. Eigenvectors of the correlation matrix.

	PC1	PC2	PC3	PC4
ATVT	**0.3520**	0.0774	**0.4682**	0.1101
NTPH	−0.3087	**0.4643**	−0.2010	0.0330
BAPH	−0.1517	**0.5470**	0.1315	0.2141
TVTPH	−0.0915	**0.5614**	0.1527	0.0964
AQD	**0.3326**	0.0453	**0.5076**	0.0861
SB3	−0.1082	−0.0420	0.1070	**0.5630**
SB7	−0.2637	−0.2888	0.0276	**0.5045**
NDVI	−0.0359	0.0715	**0.3801**	**−0.4627**
MSAVI2	−0.2956	−0.2097	**0.3703**	0.0775
DR	−0.1854	−0.0899	0.2187	0.1282
DWB	**0.3217**	0.0119	0.0093	0.1521
Slope	**0.3911**	0.1258	−0.2137	0.0936
MAT	**0.4219**	0.0606	−0.2338	0.2905

ATVT = Average total volume per tree, NTPH = Number of trees per hectare, BAPH = Basal area per hectare, TVTPH = Total volume of trees per hectare, AQD = Average quadratic diameter, SB3 = Spectral band 3, SB7 = Spectral band 7, NDVI = Normalized difference vegetation index, MSAVI2 = Modified soil-adjusted vegetation index 2, DR = Distance to roads, DWB = Distance to water bodies, MAT = Mean annual temperature.
PC=Principal Component
Bold letters indicate the dominant coefficients.

The PCA results indicate that most of the variations in forest structure can be explained by Average total volume per tree, Average quadratic diameter, Distance to roads, Slope and Mean annual temperature. The CA determined three groups; 85.2 % of the total variability is explained through the three determined groups. Other statistical criteria that reinforced this decision were the pseudo T^2, the *CCC* and pseudo F (Figure 5).

Figure 6 presents the grouping of the basins. A line was drawn near the value 0.83 to illustrate the cut-off point that delimits the number of groups established according to the euclidean distance. The line layout can vary according to the statistical and practical criteria defined, increasing or decreasing the number of groups.

The groups were evaluated and were statistically different ($P \leq 0.0001$). Group 1 comprises 133 basins, group 2 comprises 140 basins and group 3 comprises a total of 72 basins (Figure 7).

Group 1 (Bilbao group, n = 133). They have the lowest average of the Average total volume per tree, a moderate value for the number of trees per hectare, the smallest value for the basal area per hectare and for the total volume of trees per hectare. Regarding the average quadratic diameter, this group has an intermediate average value of 21.85 cm. Bands 3 and 7 in this group showed intermediate values between the other groups (0.036 and 0.08). The vegetation indexes showed higher values than group 1 and 2 for NDVI with 0.55 and 0.18 for MSAVI, indicating good coverage with respect to the photosynthetic activity of the vegetation. There are great distances to water bodies. The distance to the roads is reduced compared with the group 2 and 3. They present the lowest slope

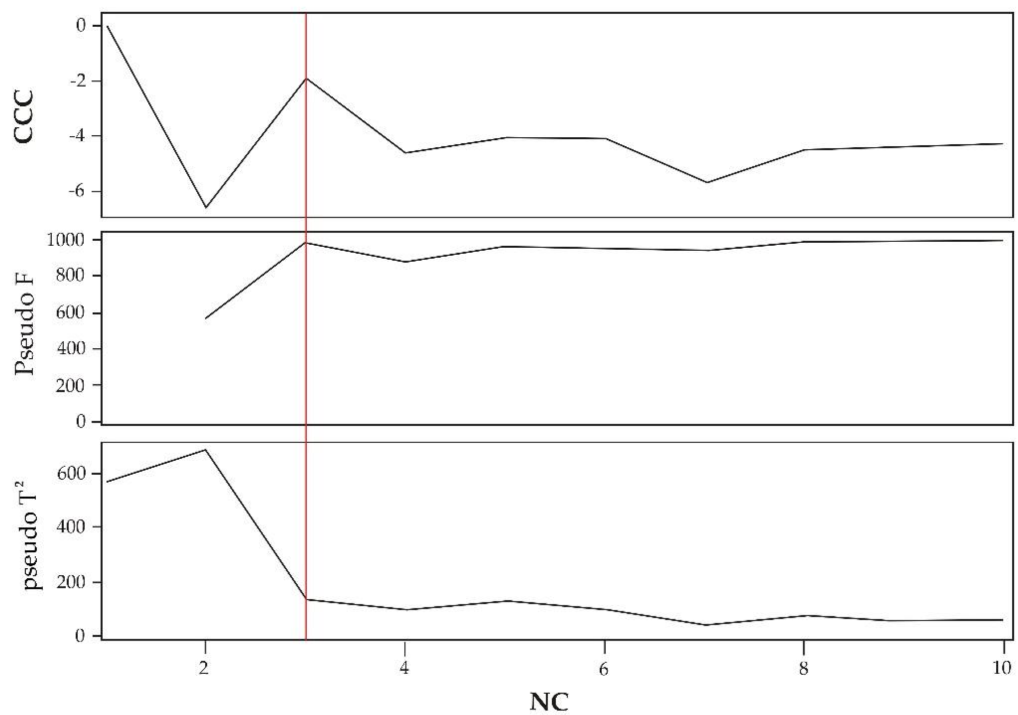

Figure 4. Statistical criteria for determining the number of groups to be established. CCC = Cubic Clustering Criterion. NC = Number of clusters.

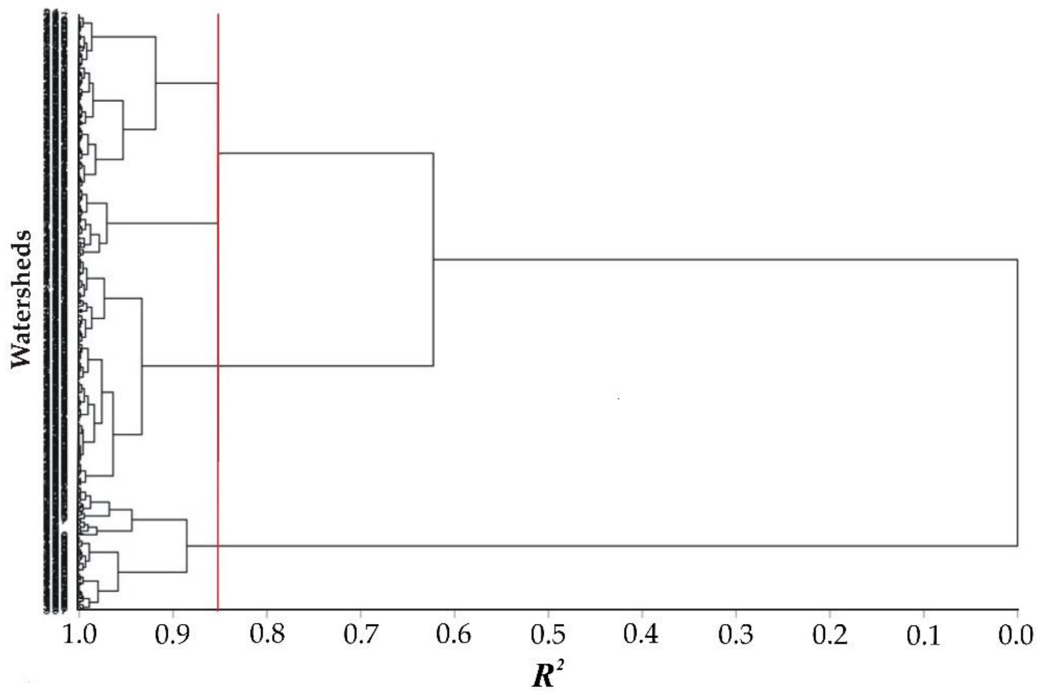

Figure 5. Group tree of watersheds using cluster analysis based on the coefficient of determination (R^2).

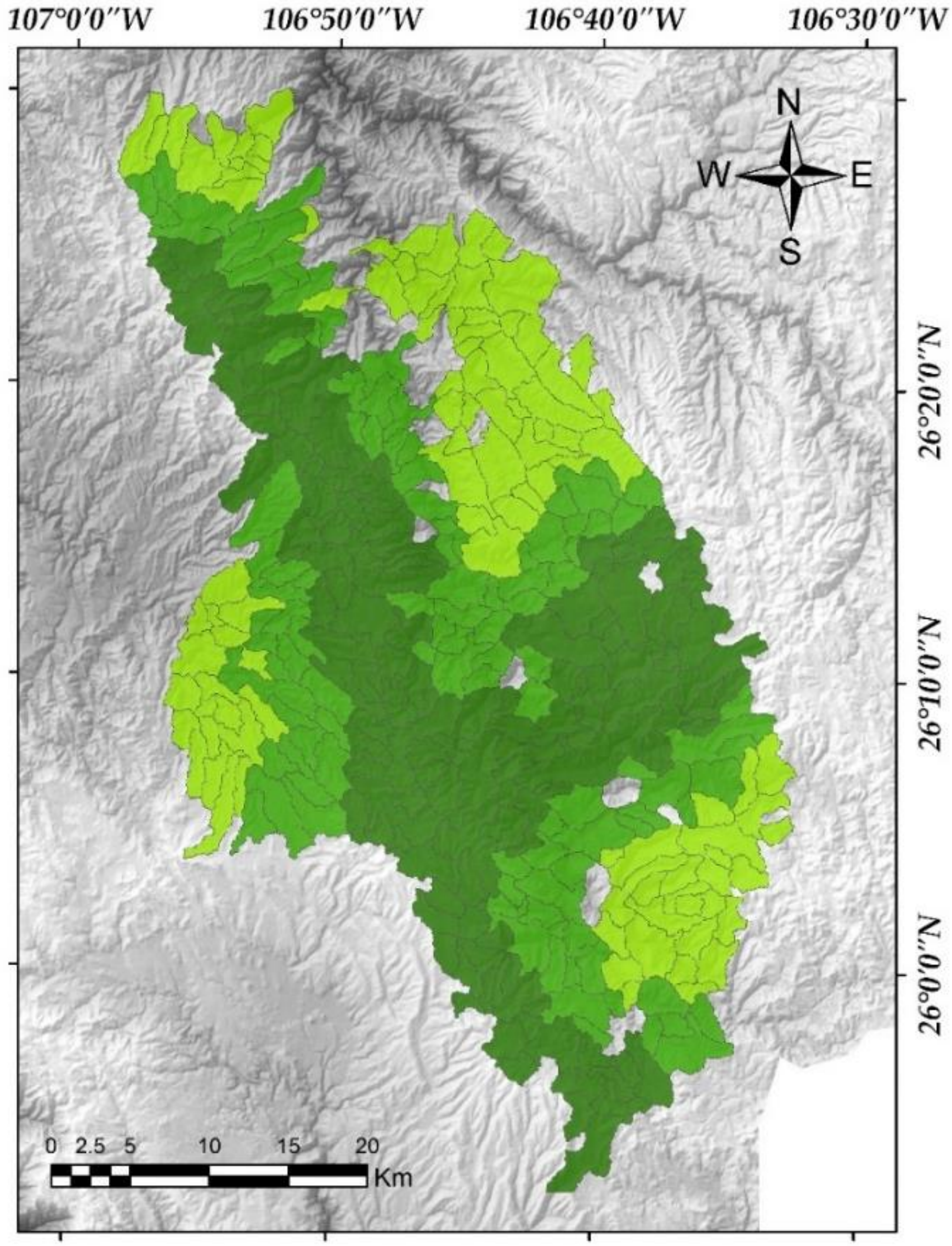

Figure 6. Spatial distribution of the tree clusters defined by group analysis. Group 1(●), Group 2(●), Group 3(●). Name of colors: Bilbao (●), Limeade (●), Inch Worm (●).

averages of the three groups. The mean annual temperature is also the lowest of the groups.

Group 2 (Limeade group, n = 140). Basins are characterized by having the lowest average value of the average total volume per tree (0.28) and average quadratic diameter (0.2157). The highest values are for the number of trees per hectare, the basal area per hectare and the total volume of trees per hectare. Bands 3 and 7 had the highest values (0.039 and 0.081). The NDVI and the MSAVI2 had the smallest values. They have intermediate values towards the distance water bodies and the distance to roads, as well as with the slope and the mean annual temperature. Due to the characteristic values of the average total volume per tree and the average quadratic diameter, a large number of trees with small trunks are distributed in these basins. This group can be characterized as a group with medium forest productive potential, possibly a young forest.

Group 3 (Inch Worm group, n = 72). This group is characterized by having the highest average value of the average total volume per tree (0.33) and average quadratic diameter (0.23). Low values were observed for the number of trees per hectare, the basal area per hectare and the total volume of trees per hectare. Band 3 and band 7 showed the lowest reflectance values, indicating that they have an abundant vegetation cover. In these basins, there are trees with large diameters. The NDVI showed the highest value for this group, which can be attributed to the high photosynthetic activity. The group of basins has a considerable distance to water bodies, indicating that they are remote areas without alteration, supported by the distance to the roads. The slope shows the highest values, which may be an indicator of remote areas with rugged topography. The mean annual temperature was the highest of the groups.

The classification of the basins in the groups was evaluated through the methods of restitution and cross-validation of the DA. A total of 13 watersheds were poorly classified, which resulted in a total error of 0.03. With the method of cross-validation, 18 basins were reported as poorly classified with a total error of 0.04. The group basin map generated in the CA was modified based on the DA (Figure 8). Using the results of the cross-validation, the misclassified observations were changed to those suggested.

Finally, the multivariate analysis of variance (MANOVA) showed that there are significant differences between basin groups with respect to the original variables (Wilk's Lambda = 0.1, Pr> F <0.001), (Table 3).

The multivariate techniques and spatial information maps were useful to interpret the relationship between the multiple variables. The PCA has been previously used to examine and interpret the spatial behavior of the forest (Li *et al*., 2006; Zhang *et al*., 2011). The relationship between variables from field sampling, thematic maps, and spectral maps provided significant information on the condition of the forest. In this study, four components were needed to explain the original data set. The four components showed the behavior of the variables in the basins.

According to Yidina *et al*. (2010), in the analysis of dimensionality reduction of variables, PC1 usually represents the most important mixture of processes. The PCA was useful in the interpretation of qualitative variables. The dominant variables in the first main component were average total

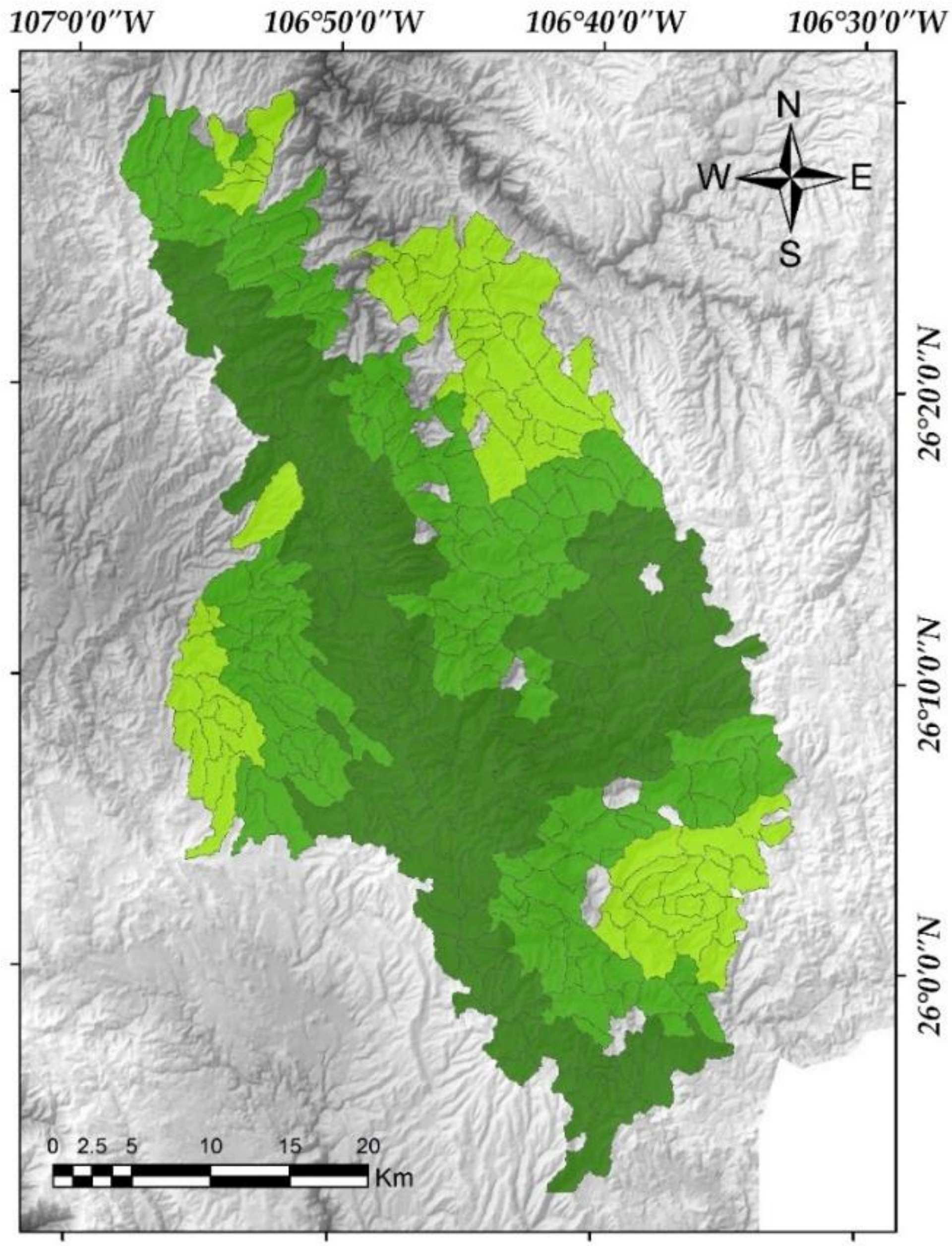

Figure 8. Spatial distribution of the 3 groups modified by the method of cross-validation of the Discriminant Analysis. Group 1(●), Group 2(●), Group 3(●). Name of colors: Bilbao (●), Limeade (●), Inch Worm (●).

Table 4. MANOVA test criteria and approximations of *F*.

Contrasts	Value	*F*-Value	DF	Pr > *F*
All	0.1026	53.84	26	<.0001
1 vs 2 y 3	0.1389	157.31	13	<.0001
2 vs 1 y 3	0.8128	5.85	13	<.0001
3 vs 1 y 2	0.1579	135.37	13	<.0001

DF= degree of freedom,
Pr > *F*= p-value of F statics

volume per tree, average quadratic diameter, corresponding to forest variables. Meanwhile, distance to roads, slope and mean annual temperature, presented a strong relationship in the variation of the forest structure in the 'Ejido Chinatu' community. The above agrees with the study by Castillo-Rodríguez *et al*. (2010). The results of PC1 are consistent with the distribution of vegetation through the altitudinal gradient, where temperature and topography play an important role in the presence of certain species of panaceas (Rzedowski, 1998; Hlásny *et al*., 2017). This is consistent with what was mentioned by Helman *et al*. (2017), where they found that mean annual temperatures influenced forest productivity in Mediterranean Mount Carmel forests in Israel. The values of the coefficients within the PC1 present similarities that imply the variables have a similar influence on the variation of the forest structure. PC2 consisted of the variables number of trees per hectare, basal area per hectare and total volume of trees per hectare had weak and moderate coefficients. PC2 can be interpreted as a component that describes the structure of the forest in relation to the number of trees, the basal area and the volume measured per hectare. PC3 was represented by average total volume per tree, average quadratic diameter and the spectral indices NDVI and MSAVI2. This is consistent with what was mentioned by Kumar et al. (Kumar *et al*., 2018), who related structural variables of the forest such as diameter and height with the NDVI, obtaining strong relationships of R^2 = 0.90. This component, similar to PC1, presents information on the forest structure based on the MSAVI2 spectral indices. As shown in the variable displacement plot (Figure 4), as the average quadratic diameter and the average total volume per tree decrease, the MSAVI2 increases. The MSAVI2 spectral index helps us to understand the reflectance emitted by bare soil. In this case, lower average quadratic diameter and higher MSAVI2 indicate that a large number of trees are small.

In agreement with the spatial distribution of the groups shown in Figure 7 and 8, the grouping by CA presented a homogeneous geographical behavior. Group 1 represents the basins that correspond to the central corridor, where forest conditions for exploitation could be more accessible. The proximity to the roads is an indicator of anthropic presence (Kirby *et al*., 2006), and the moderate slopes for this group of basins make wood extraction more feasible. Espinosa *et al.*, (2016) found that topography is related to species richness in dry forests in southern Ecuador. This is consistent with the distribution of forest variables across the 'Ejido Chinatu', where the increase in slopes ranges from flat to very steep slopes. The moderate spectral values of bands 3 and 7 indicate there is moderate absorption by vegetation and areas that are in a state of recovery. This group shows geographical continuity without being very isolated. Group 2 corresponds to basins that have average values with respect to almost all the variables, considering it as a group of basins that can be under a moderate anthropic intervention (transition zone) (Barber *et al*., 2014). It is divided into several homogeneous groups and does not present cases of isolated basins. Finally, Group 3 contains basins that, due to their spatial distribution, are further away from anthropogenic activities and may not have forest management. The low reflectivity in bands 3 and 7, and the high photosynthetic activity represented by the NDVI values indicate good vegetation conditions (Mancino *et al*., 2014). These basins are also divided into groups and present an isolated basin.

The PCA detects the factors that control the spatial variation of the forest structure. Although the PCA is generally considered as a statistical exploratory technique (Johnson, 1998), it is capable of being incorporated into the results that explain the distribution of a particular landscape (De La Cueva, 2008). GIS-based maps have the ability to visualize the spatial relationships between environmental data and other attributes, as reported by Facchinelli *et al*. (2001). The CA was useful to identify those basins that are similar in terms of their multiple environmental characteristics. Riitters and Coulston (2005) used spatial statistics to identify the geographical concentrations of forests located near deforested or clear areas. Wang *et al*. (2016) used the multivariate techniques PCA and CA for the delimitation of environmental units. Kupfer *et al*. (2012) regionalized 2100 watersheds based on landscape metrics. Trakhtenbrot and Kadmon (2005) used the CA for the identification of sites that represent the diversity of species. Ramachandra *et al*. (2016) analyzed information on landscape metrics and social variables using multivariate techniques in forests in Uttara Kannada District, India.

Based on the results of PCA, CA and DA it was possible to understand the multivariate relationship of the set of variables. At the same time, with the combination of geospatial data and multivariate techniques, it was possible to analyze the spatial variability from a point of view of reducing the dimensionality of the information. The combination of both was useful to examine patterns in common group of variables, allowing us to summarize the multiple relationships of variables in geographic regions to use in forest management analysis.

The identification and delimitation of geographic regions in forests is an active area of ongoing research. Technological and methodological advances allow analyzing data from free sources and field information. This study provides a perspective for the analysis of this information, helping us to enhance its interpretation with a more quantitative approach. Traditional grouping and classification methods are widely used, leading to the delimitation of coherent forest region classifications from a geospatial point of view.

From a methodological point of view, the study provides the identification of forest regions that are possibly more vulnerable to the effects of anthropogenic activities, such as changes in land cover/land use, deforestation, degradation. In addition, the methodology serves to identify the forest geographical areas with potential for conservation. Such areas are of great importance for the development of conservation plans and protection of forest areas, promoted mainly by government agencies.

In particular, the increasingly availability of free and georeferenced data sources, such as those from remote sensing, spectral indicators, digital elevation models and information derived from them, represents a valuable resource for quantitative approaches to zoning. The presentation of geospatial information from forest areas allows users to examine the characteristics of each area, their variability and their level of productivity. However, the detailed presentation of these variations in a broader regional context, where transition zones could be detected, is complex but promising for future studies.

CONCLUSIONS AND RECOMMENDATIONS

The use of GIS associated with geostatistical techniques represents a solid scientific tool for regionalization and grouping of landscape features. The features and regionalization of hydrological basins based on environmental attributes and vegetation structures are key to the planning and environmental management of the territories. The application of this methodology allows the rapid integration of several environmental and biological attributes, which can be grouped according to their characteristics. In this way we can define the productive potentials through the regionalization analysis. Forest regionalization mapping can benefit a wide variety of management, conservation and protection activities. Thus, if a forest pattern has been identified in one region as favorable or problematic, it can be replicated in another region with similar conditions.

Multivariate statistical methods, forest structure data, topographies and satellites can be useful to group and regionalize forest areas. The multivariate geographic information system approach showed the spatial relationships between the different variables. Although the results of the present study provided preliminary conclusions about the characteristics of the basin groups, more studies such as multicriteria evaluation techniques, interpolation methods, among others, are needed to obtain a better understanding of the source of variation in the forest structure.

REFERENCES

Achard, F., D. Mollicone, H.J. Stibig, D. Aksenov, and L. Laestadius, Z. Li, P; Potapov, and A. Yaroshenko. 2006. Areas of rapid forest-cover change in boreal Eurasia. For. Ecol. Manage. 237:322-334.

Aguilar, S.V., M.A. Castillo, V.F. Guerrero, C. Wehenkel, and A.C. Pinedo. 2015. Modeling the potential distribution of Picea chihuahuana Martínez, an endangered species at the Sierra Madre Occidental, Mexico. Forest. 6:692-707.

Bach, K., M. Schawe, S. Beck, G. Gerold, S.R. Gradstein, and R. Moraes. 2003. Vegetación, suelos y clima en los diferentes pisos altitudinales de un bosque montano de Yungas, Bolivia: Primeros resultados. Ecología en Bolivia. 38:3-14.

Barber, C. P., M. A. Cochrane, C. M. Souza Jr, and W. F. Laurance. 2014. Roads, deforestation, and the mitigating effect of protected areas in the Amazon. Biol. Conserv. 177:203-209.

Batayneh, A., and T. Zumlot. 2012. Multivariate statistical approach to geochemical methods in water quality factor identification; application to the shallow aquifer system of the Yarmouk basin of north Jordan. Research Journal of Environmental and Earth Sciences. 4:756–768.

Bax, V., W. Francesconi, and M. Quintero. 2016. Spatial modeling of deforestation processes in the Central Peruvian Amazon. Journal of Natural Conservancy. 29:79-88.

Bhuiyan, M.A., M.A. Rakib, S.B. Dampare, S. Ganyaglo, and S. Suzuki. 2011. Surface water quality assessment in the central part of Bangladesh using multivariate analysis. Ksce J. Civ. Eng. 15:995-1003.

Bolton, D.K., N.C. Coops, T. Hermosillo, M.A. Wulder, and J.C. White. 2017. Assessing variability in postfire forest structure along gradients of productivity in the Canadian boreal using multisource remote sensing. J. Biogeogr. 44:1294-1305.

Bonan, G.B. 2008. Forests and climate change: forcing, feedbacks, and the climate benefits of forests. Science. 320:1444-1449.

Castillo, R., J.L. Blanco, and E.M. Salinas. 2010. A geomorphologic GIS-multivariate analysis approach to delineate environmental units, a case study of La Malinche volcano (central México). Appl. Geogr. 30:629-638.

Challenger, A. 1998. La zona ecológica templada húmeda. Utilización y Conservación de Los Ecosistemas Terrestres de México: Pasado, Presente y Futuro. CONABIO-Instituto de Biología, UNAM Sierra Madre. México. 433–518.

Chambers, J.Q., G.P. Asner, D.C. Morton, L.O. Anderson, S.S. Saatchi, F.D. Espírito-Santo, and C. Souza. 2007. Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. Trends Ecol. Evol. 22:414-423.

CONABIO. 2014. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. La biodiversidad en Chihuahua: Estudio de Estado. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. México.

De La Cueva, and A.V. 2008. Structural attributes of three forest types in central Spain and Landsat ETM+ information evaluated with redundancy analysis. Int. J. Remote Sens. 29:5657-5676.

Dube, T., and O. Mutanga. 2015. Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in uMgeni catchment, South Africa. Isprs J. Photogramm. 101:36-46.

Eder, B.K., J.M. Davis, and P. Bloomfield. 1994. An automated classification scheme designed to better elucidate the dependence of ozone on meteorology. J. Appl. Meteorol. 33:1182-1199.

Espinosa, C. I., M. De la Cruz, A. Jara-Guerrero, E. Gusmán, and A. Escudero. 2016. The effects of individual tree species on species diversity in a tropical dry forest change throughout ontogeny. Ecography. 39:329-337.

Facchinelli, A., E. Sacchi, and L. Mallen, 2001. Multivariate statistical and GIS-based approach to identify heavy metal sources in soils. Environ. Pollut. 114:313-324.

Felger, R.S., and M.B. Johnson. 1995. Trees of the northern Sierra Madre Occidental and sky islands of southwestern North America. 71-83.

Franklin, S.E. 2001. Remote sensing for sustainable forest management. CRC Press.

González, E.M., J.A. Meave, N.R. Marcial, T.T. Aceves, F.G. Hernández, and G.I. Manríquez. 2012. Los bosques de niebla de México: conservación y restauración de su componente arbóreo. Revista Ecosistemas. 21:1-2.

Groombridge, B., and M.D. Jenkins. 2000. Global biodiversity: Earth's living resources in the 21st century. World Conservation Press: Cambridge, USA.

Hall, R. J., R. S. Skakun, E. J. Arsenault, and B. S. Case. 2006. Modeling forest stand structure attributes using Landsat ETM+ data: Application to mapping of aboveground biomass and stand volume. Forest. Ecol. Manag. 225:378-390.

Hamdy, O., S. Zhao, M.A. Salheen, and Y.Y. Eid. 2016. Identifying the Risk Areas and Urban Growth by ArcGIS-Tools. Geosciences. 6:47.

Helman, D., Y. Osem, D. Yakir, and I. M. Lensky. 2017. Relationships between climate, topography, water use and productivity in two key Mediterranean forest types with different water-use strategies. Agr. Forest. Meteorol. 232:319-330.

Hilker, T., M.A. Wulder, N.C. Coops, J. Linke, G. Mcdermid, J.G. Masek, F. Gao, and J.C. White. 2009. A new data fusion model for high spatial-and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. Remote Sens. Environ. 113:1613-1627.

Hlásny, T., J. Trombik, M. Bošeľa, J. Merganič, R. Marušák, V. Šebeň, P. Štěpánekd, J. Kubištaa, and M. Trnka. 2017. Climatic drivers of forest productivity in Central Europe. Agr. Forest Meteorol. 234:258-273.

INEGI. 2003. Síntesis de información geográfica del Estado de Chihuahua. Instituto Nacional de Estadística, Geografía e Informática. Aguascalientes, Ags., Méx.

INEGI. 2005. Instituto Nacional de Estadística, Geografía e Informática. Sistema de Descarga del Continuo de Elevaciones Mexicano. (online) http://mapserver.inegi.gob.mx. (Accessed on July 19, 2018).

Ingram, J.C., T.P. Dawson, and R.J. Whittaker. 2005. Mapping tropical forest structure in southeastern Madagascar using remote sensing and artificial neural networks. Remote Sens. Environ. 94:491-507.

Johnson, D. 1998. Métodos multivariados aplicados al análisis de datos. México: International Thompson Editors.

Johnson, R.A., and D.W. Wichern. 1988. Applied Multivariate Statistical Analysis; Prentice-Hall, Inc.: Upper Saddle River, NJ, USA.

Jupp, D.L., and J. Walker. 1997. Detecting structural and growth changes in woodlands and forests: The challenge for remote sensing and the role of geometric-optical modelling. In The use of remote sensing in the modeling of forest productivity. Springer, Dordrecht. 75-108.

Kalabokidis, K.D., N. Koutsias, P. Konstantinidis, and C. Vasilakos. 2007. Multivariate analysis of landscape wildfire dynamics in a Mediterranean ecosystem of Greece. Area. 39:392-402.

Kirby, K.R., W.F. Laurance, A.K. Albernaz, G. Schroth, P.M. Fearnside, S. Bergen, E.M. Venticinque, and C.D. Costa. 2006. The future of deforestation in the Brazilian Amazon. Futures. 38:432–453.

Kumar, P., H. Sajjad, K. K. Mahanta, R. Ahmed, and V. P. Mandal. 2018. Assessing suitability of allometric models for predicting stem volume of

Anogeissus pendula Edgew in sariska Tiger Reserve, India. Remote Sensing Applications: Society and Environment. 10:47-55.

Kupfer, J.A., P. Gao, and D. Guo. 2012. Regionalization of forest pattern metrics for the continental United States using contiguity constrained clustering and partitioning. Ecol. Inform. 9:11-18.

Li, A., A. Wang, S. Liang, and W. Zhou. 2006. Eco-environmental vulnerability evaluation in mountainous region using remote sensing and GIS—a case study in the upper reaches of Minjiang River, China. Ecol. Model. 192:175-187.

Ludwig, B., P.K. Khanna, R.J. Raison, and K. Jacobsen. 1997. Modelling changes in cation composition of a soil after clearfelling a eucalypt forest in East Gippsland, Australia. Geoderma. 80:95-116.

Maidment, D.R., and D. Djokic. 2000. Hydrologic and Hydraulic Modelling Support: with Geographic Information Systems. ESRI, Inc. 65-84.

Mancino, G., A. Nolè, F. Ripullone, and A. Ferrara. 2014. Landsat TM imagery and NDVI differencing to detect vegetation change: assessing natural forest expansion in Basilicata, southern Italy. IFOREST. 7:75.

Miranda, A.L., E.J. Garza, J. J. Pérez, O.A. Calderón, M.A. Tagle, P.M. García, and A.C. Salado. 2012. Modeling susceptibility to deforestation of remaining ecosystems in North Central Mexico with logistic regression. J. Forestry Res., 23:345-354.

Oketola, A.A., S.M. Adekolurejo, and O. Osibanjo. 2013. Water quality assessment of River Ogun using multivariate statistical techniques. J. Environ. Prot. Ecol. 4.466.

Pasher, J., and D.J. King. 2010. Multivariate forest structure modelling and mapping using high resolution airborne imagery and topographic information. Remote Sens. Environ. 114:1718-1732.

Pérez, V.A., J.F. Mas, and A.L. Zielinska. 2012. Comparing two approaches to land use/cover change modeling and their implications for the assessment of biodiversity loss in a deciduous tropical forest. Environ. Modell. Softw. 29:11-23.

Pflugmacher, D., W. B. Cohen, and R. E. Kennedy. 2012. Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. Remote Sens. Environ. 122:146-165.

Powell, S. L., W. B. Cohen, S. P. Healey, R. E. Kennedy, G. G. Moisen, K. B. Pierce, and J. L. Ohmann. 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. Remote Sens. Environ. 114:1053-1068.

Ramachandra, T. V., B. Setturu, and S. Chandran. 2016. Geospatial analysis of forest fragmentation in Uttara Kannada District, India. Forest Ecosystems. 3:10.

Randolph, J.C., G.M. Green, J. Belmont, T. Burcsu, and D. Welch. 2005. Forest ecosystems and the human dimension. Seeing the forest and the trees: Human–environment interactions in forest ecosystems. Moran, E.F., Ostrom, E. Eds.; MIT Press: Cambridge, Massachusetts, USA. 81–103.

Riitters, K.H., and J.W. Coulston. 2005. Hotspots of perforated forest in the Eastern United States. Enviro. Manage. 35:483-492.

Rzedowski, J. 1998. Diversidad y orígenes de la flora fanerogámica de México. Diversidad biológica de México: orígenes y distribución; Ramamoorthy, T.P., Bye R., Lot A., Fa, J. Eds.; UNAM: Instituto de Biología, México. 129-145

Sáenz, R.C., G.E. Rehfeldt, N.L. Crookston, P. Duval, R.S.T. Amant, J.B. Beaulieu, and A. Richardson. 2010. Spline models of contemporary, 2030, 2060 and 2090 climates for Mexico and their use in understanding climate-change impacts on the vegetation. Climatic Change. 102:595-623.

Sanderson, M., M. Santini, R. Valentini, and E. Pope. 2012. Relationships between forests and weather. EC Directorate General of the Environment, (Accessed on Oct 19, 2018)

SAS. SAS software. Version 9.1.3 Cary: SAS. 2006. Institute Inc.

Singh, K.P., A. Malik, D. Mohan, and S. 2004. Sinha. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. Water Res. 38:3980-3992.

Soenen, S. A., D. R. Peddle, R. J. Hall, C. A. Coburn, and F. G. Hall. 2010. Estimating aboveground forest biomass from canopy reflectance model inversion in mountainous terrain. Remote Sens. Environ. 114:1325-1337.

Torres, A.B., L. Acuña, and J.M. Vergara. 2014. Integrating CBM into land-use based mitigation actions implemented by local communities. Forests. 5:3295-3326.

Torres, R. 2013. Cuenca río Turuachi con atención a áreas críticas de deforestación y degradación forestal. Propuesta de Proyecto enviada por Ejido Chinatú como parte de la Convocatoria a proyectos de campo a través de Alianzas para la preparación a REDD+ del proyecto México de Reducción de Emisiones por Deforestación y Degradación Forestal (MREDD+). Alianza México REDD+: Mexico City, Mexico, 2012.

Trakhtenbrot, A., and R. Kadmon. 2005. Environmental cluster analysis as a tool for selecting complementary networks of conservation sites. Ecol. Appl. 15:335–345.

Tritsch, I., P. Sist, I.D.S. Narvaes, L. Mazzei, L. Blanc, C. Bourgoin, and V. Gond. 2016. Multiple patterns of forest disturbance and logging shape forest landscapes in Paragominas, Brazil. Forests. 7:315.

Vázquez-Quintero, G., R. Solís-Moreno, M. Pompa-García, F. Villarreal-Guerrero, C. Pinedo-Alvarez, and A. Pinedo-Alvarez. 2016. Detection and projection of forest changes by using the Markov Chain Model and Cellular Automata. Sustainability. 8:236.

Velázquez, A., J.F. Mas, J.R.D. Gallegos, R.M. Saucedo, P.C. Alcántara, R. Castro, and J.L. Palacio. 2002. Patrones y tasas de cambio de uso del suelo en México. Gaceta ecológica. 62.

Verdú, F., J. Salas, and C.V. García. 2012.A multivariate analysis of biophysical factors and forest fires in Spain, 1991–2005. Int J. Wildland Fire. 21:498-509.

Wang, H.F., S. Qureshi, B.A. Qureshi, J.X. Qiu, C.R. Friedman, J. Breuste, and X.K. Wang. 2016. A multivariate analysis integrating ecological, socioeconomic and physical characteristics to investigate urban forest cover and plant diversity in Beijing, China. Ecol. Indic. 60:921-929.

Wang, X., G. Shao, H. Chen, B.J. Lewis, G. Qi, D. Yu, and L. Dai. 2013. An Application of remote sensing data in mapping landscape-level forest biomass for monitoring the effectiveness of forest policies in Northeastern China. Environ. Manage. 52:612-620.

Wolter, P. T., P. A. Townsend, and B. R. Sturtevant. 2009. Estimation of forest structural parameters using 5 and 10 meter SPOT-5 satellite data. Remote Sens. Environ. 113:2019-2036.

Yidana, S.M., B. Banoeng-Yakubo, and T.M. Akabzaa. 2010. Analysis of groundwater quality using multivariate an spatial analyses in the Keta basin, Ghana. J. Afr. Earth Sci. 58:220–234.

Zhang, Q. F., F. Q. Wu, L. Wang, L. Yuan, and L. S. Zhao. 2011. Application of PCA integrated with CA and GIS in eco-economic regionalization of Chinese Loess Plateau. Eco. Econ. 70:1051-1056.

Zhu, Z., C.E. Woodcock, and P. Olofsson. 2012. Continuous monitoring of forest disturbance using all available Landsat imagery. Remote Sens. Environ. 122:75-91.

STUDY V. MULTIVARIATE AND SPATIAL ANALYSIS OF PHYSICOCHEMICAL PARAMETERS IN AN IRRIGATION DISTRICT, CHIHUAHUA, MEXICO

BY:

Jesús A. Prieto-Amparan[1], B. A. Rocha-Gutiérrez[2], M. L. Ballinas-Casarrubias[2], M. C. Valles-Aragón[3]*, M. R. Peralta-Pérez[2], Alfredo Pinedo-Alvarez[1]

[1]Facultad de Zootecnia y Ecología, Universidad Autónoma de Chihuahua. Periférico Francisco R. Almada km 1., C.P. 31453. Chihuahua, Chih. México.
[2]Facultad de Ciencias Químicas, Universidad Autónoma de Chihuahua. Circuito Universitario Campus II, s/n., C.P. 31125. Chihuahua, Chih., México
[3]Facultad de Ciencias Agrotecnológicas, Universidad Autónoma de Chihuahua, Campus I, s/n., C.P. 31350. Chihuahua, Chih., México.

*Corresponding author: María Cecilia Valles Aragon; Facultad de Ciencias Agrotecnológicas, Universidad Autónoma de Chihuahua, Campus I, s/n., C.P. 31350. Chihuahua, Chih., México. valles.cecilia@gmail.com, tel. +52-(614)-439-1844, DOI: https://doi.org/10.3390/w10081037.

ABSTRACT

MULTIVARIATE AND SPATIAL ANALYSIS OF PHYSICOCHEMICAL PARAMETERS IN AN IRRIGATION DISTRICT, CHIHUAHUA, MEXICO

BY:

M. C. JESUS ALEJANDRO PRIETO AMPARAN

Water quality is relevant due to the complexity of the interaction of physicochemical and biological parameters. The Irrigation District 005 (ID005) located in Chihuahua, Mexico; it is one of the most important districts in Chihuahua, Mexico; for that reason, it was proposed to investigate the water quality of the site. Water samples were collected in two periods: Summer (S1) and Fall (S2). The samples were taken from 65 wells in S1, and 54 wells in S2. Physicochemical parameters (PhP) such as Arsenic (As), Temperature, Electrical Conductivity (EC), Oxide Reduction Potential (ORP), Hardness, pH, Total Dissolved Solids (TDS), and Turbidity were analyzed. The data were subjected to statistical principal component analysis (PCA), cluster analysis (CA) and spatial variability tests. In both seasons, the TDS exceeded the Mexican maximum permissible level (MPL) (35 % S1, 39 % S2). Turbidity exceeded the MPL in S1 (29 %) and in S2 (12 %). Arsenic was above the MPL for water of agricultural use in 9 % (S1) and 13 % (S2) of the wells. The PCA results suggested that most variations in water quality in S1 were due to As, pH and Temperature, followed by EC, TDS and Hardness; while in S2 to EC, TDS and Hardness, followed by As and pH.

Keywords: spatial distribution; PCA; CA; water quality

RESUMEN

MULTIVARIATE AND SPATIAL ANALYSIS OF PHYSICOCHEMICAL PARAMETERS IN AN IRRIGATION DISTRICT, CHIHUAHUA, MEXICO

POR:

M. C. JESÚS ALEJANDRO PRIETO AMPARÁN

Doctor en Philosophia de Recursos Naturales

Secretaría de Investigación y Posgrado

Facultad de Zootecnia y Ecología

Universidad Autónoma de Chihuahua

Presidente: Dr. Alfredo Pinedo Alvarez

La calidad del agua es relevante debido a la complejidad de la interacción de los parámetros físico-químicos y biológicos. El Distrito de Riego 005 (ID005) ubicado en Chihuahua, México; es uno de los distritos más importantes de Chihuahua, México, por lo que se propuso investigar la calidad del agua del lugar. Las muestras de agua se recogieron en dos períodos: Verano (S1) y otoño (S2). Las muestras se tomaron de 65 pozos en S1 y 54 pozos en S2. Los parametros parámetros físico-químicos (PhP) anaizados fueron: arsénico (As), temperatura (Temp), conductividad eléctrica (CE), potencial de reducción de óxido (PRO), dureza (Dur), pH, sólidos disueltos totales (SDT) y turbidez (Turb). Los datos se sometieron al análisis de componentes principales (ACP), análisis de grupos (AG) y pruebas de variabilidad espacial. En ambas temporadas, el TDS superó el nivel máximo permitido en México (35 % S1, 39 % S2). La turbidez superó el limite máximo permisible (LMP) en S1 (29 %) y en S2 (12 %). El arsénico estaba por encima de la LMP para el agua de uso agrícola en el 9 % (S1) y el 13 % (S2) de los pozos. Los resultados del ACP sugirieron que la mayoría de las variaciones en la calidad del agua en S1 se debían a As, pH y Temp, seguidos por CE, SDT y Dur; mientras que en S2 a CE, SDT y Dur, seguidos por As y pH.

Palabras clave: distribución espacial; ACP; AG; calidad del agua.

ABSTRACT

MULTIVARIATE AND SPATIAL ANALYSIS OF PHYSICOCHEMICAL PARAMETERS IN AN IRRIGATION DISTRICT, CHIHUAHUA, MEXICO

BY:

M. C. JESUS ALEJANDRO PRIETO AMPARAN

Water quality is relevant due to the complexity of the interaction of physicochemical and biological parameters. The Irrigation District 005 (ID005) located in Chihuahua, Mexico; it is one of the most important districts in Chihuahua, Mexico; for that reason it was proposed to investigate the water quality of the site. Water samples were collected in two periods: Summer (S1) and Fall (S2). The samples were taken from 65 wells in S1, and 54 wells in S2. Physicochemical parameters (PhP) such as Arsenic (As), Temperature, Electrical Conductivity (EC), Oxide Reduction Potential (ORP), Hardness, pH, Total Dissolved Solids (TDS), and Turbidity were analyzed. The data were subjected to statistical principal component analysis (PCA), cluster analysis (CA) and spatial variability tests. In both seasons, the TDS exceeded the Mexican maximum permissible level (MPL) (35 % S1, 39 % S2). Turbidity exceeded the MPL in S1 (29 %) and in S2 (12 %). Arsenic was above the MPL for water of agricultural use in 9 % (S1) and 13 % (S2) of the wells. The PCA results suggested that most variations in water quality in S1 were due to As, pH and Temperature, followed by EC, TDS and Hardness; while in S2 to EC, TDS and Hardness, followed by As and pH.

Keywords: spatial distribution; PCA; CA; water quality.

INTRODUCTION

Economic development, industrialization and urbanization, along with population growth, lead to an accelerated water consumption, which has generated concern for fresh water as a scarce resource (Kengnal *et al*., 2015; Varol and Davras, 2015). Water quality is an important factor that affects human health and ecological systems (Qadir *et al*., 2008). In the rural context, groundwater is the support of agricultural irrigation and it is essential for providing additional food security resources (Morris *et al*., 2003) However, food security can be affected by pollutants present in the irrigation water, causing serious clinical and physiological problems to humans when such pollutants get accumulated in large amounts (Sharma *et al*., 2007; Khan *et al*., 2008). In general, water quality for various applications is determined by its physical characteristics, chemical composition, biological parameters and uses (Gupta *et al*., 2009, Kengnal *et al*., 2015). These parameters reflect the inputs from natural sources, including atmosphere, soil and particular geological characteristics of each region, as well, as anthropogenic influence of various activities (Patil *et al*., 2012; Brahman *et al*., 2013; AlSuhaimi *et al*., 2017).

The evaluation of water quality in most countries has become a critical issue in recent years (Varol and Davraz, 2015). Water quality is subject to constant changes due to seasonal and climatic factors (AlSuhaimi *et al*., 2017). Likewise, spatial variations emphasize the need for water monitoring that provides a representative and reliable estimate (Muangthong and Shrestha, 2015). Recently, several approaches have been used in for water quality analysis. Among them, we can find methods based on modelling, monitoring or statistic techniques (Giri and Qiu, 2016). Modelling tools such as Soil and Water Assesment Tool (SWAT) or Agricultural Nonpoint have been employed to evaluate water quality at the watershed scale. The statistic techniques commonly used for the monitoring of water quality include: Ordinary Least Sqares (OLS), Geographic Weighted Regression (GWR) among others. The monitoring techniques provide key information for decision making regarding water quality (Giri and Qiu, 2016). However, in comparison to these approaches, multivariate techniques such as Principal Component Analysis (PCA) and Cluster Analysis (CA) could be used for analyzing big water quality databases without losing important information (Helena *et al*., 2000; Singh *et al*., 2005; Wang *et al*., 2013).

Multivariate techniques and exploratory data analyses are appropriate for the synthesis of data and its interpretation (Singh *et al*., 2005). Classification, modeling and interpretation of the monitored data are the most important steps in the evaluation of water quality (Zhao *et al*., 2007; Brogna *et al*., 2017; Boyacioglu, 2006). The application of multivariate statistical techniques, such as principal component analysis (PCA) or Cluster Analysis (CA), has significantly increased in recent years, especially for the analysis of environmental data and extracting significant information (Bhuiyan *et al*., 2011; Batayneh and Zumlot, 2012). Additionally, these analyses have been reported as effective methods for the characterization and evaluation of water quality parameters (Brahman et al., 2013). PCA and CA are the most common multivariate statistical methods used in environmental studies (Oketola *et al*., 2013).

The PCA is a mathematical technique used to reduce the dimensions of multivariate data and explain the correlation between a large number of variables observed by extracting a smaller number of new variables (i.e. the principal components or PC) (Jackson, 1983; Wunderlin *et al*., 2001; Loska and Wiechula, 2003). The CA helps grouping objects (cases) based on homogeneity and heterogeneity between groups. The clusters characteristics are not known in advance but may be determined in the analysis. Such analysis benefits the interpretation of the data by pointing out associations among the studied variables (Vega *et al*., 1998; Al-Bassam, 2006). The application of different multivariate statistical techniques aids in the interpretation of complex data matrices to better understand the water quality and ecological status of the studied systems Kazi *et al*., 2009; Muangthon and Shrestha, 2015). It also allows the identification of possible factors or sources that influence water systems and offers a valuable tool for the reliable management of water resources, both in quantity and quality (Batayneh and Zumlot, 2012). Previous research has shown that the use of multivariate analysis allows defining new variables that provide information on the variability of environmental data, as well as on the influence of each variable (Chiu, 2000; Medellín-Vázquez, 2003).

Furthermore, interpolation methods have been employed to map the spatial distribution of soil properties (Villatoro *et al*., 2008; Bhunia *et al*., 2011), heavy metals (Xie *et al*., 2011; Yan *et al*., 2015), population characteristics (Navarrete, 2012), precipitation (Wang *et al*., 2014; Núñez *et al*., 2014), radioactive elements (Skeppttröm and Olofsson, 2006; Maroju, 2007), among others. Data interpolation offers the advantage of projecting maps or continuous surfaces from discrete data (Johnston *et al*., 2001). Therefore, spatial interpolation techniques are essential to create a continuous (or predictable) surface from values of sampled points (Wang *et al*., 2014). Interpolation is an efficient method to study the spatial allocation of elements, their inconsistency, reduce the error variance and execution costs (Behera and Shukla, 2015). The interpolation methods are useful for identifying contamination sources, assessing pollution trends and risks (Markus and McBratney, 2001; Rawlins *et al*., 2006). A growing number of studies have shown the need to determine the spatial distribution of pollutants. Spatial data helps to define areas where risks are higher and contribute to making the decision to identify the locations where remediation efforts should be concentrated (Maas *et al*., 2010). However, one of the characteristics of the spatial distribution of pollutants lies in their frequent spatial heterogeneity (Walter, 2002).

Few studies that combine multivariate techniques and interpolation methods have been completed (Chaoyang *et al*., 2009; Fu and Wei, 2013). The selection of the PCA and CA methods is carried out to understand the multivariate relationship of the parameters. Many times, the studies are carried out univariately. In this study, we used these techniques in a descriptive way, so that by visually comparing the results of the group analysis and the interpolations, we understood the spatial distribution of the parameters and even more the spatial distribution of the CPs. The modeling techniques were not considered for the amount of information used necessary to achieve the calibration. The objective of the present study was to analyze eight physicochemical parameters (PhP) in water samples from wells of the Irrigation District 005 (ID005) in Chihuahua, Mexico, perform a data analysis using multivariate techniques to evaluate the PhP contribution in water quality and apply interpolation methods to analyze the spatial variation of the PhP.

MATERIALS AND METHODS

Research Area

The ID005 is located in the south-central region of the State of Chihuahua (Figure 1), among the geographical coordinates 105°4’O - 28°30’N, 105°2’O - 28°30’N, 1054’O - 28°10’N, 105°2’O - 28°10’N. It has an average altitude of 1,156 meters above sea level. The predominant climate is semi-desert, with an average of annual rainfall of 350 mm (Ortega-Gaucin *et al*., 2009). The ID005 is composed by 10 irrigation modules, which are administered by local associations. The district is divided into two constituted irrigation units based on infrastructure characteristics, to facilitate water distribution (Aguirre-Grijalva, 2003). Each unit is managed by a Limited Liability Corporation, that is integrated as follows: 1) The first unit called Conchos is composed by irrigation modules 1 to 5 and 12, which are mainly supplied by water from La Boquilla dam; 2) The second unit, called San Pedro, is integrated by modules 6, 7, 8 and 9, which are supplied by water from the Francisco I. Madero dam, groundwater and, to a lesser extent, by water from the La Boquilla dam (Ortega-Gaucin, 2012).

Sampling

Two different sampling periods were performed in operating wells in the studied area during 2017. The first sampling was performed in Summer (S1) and the second sampling in Fall (S2) (Figure 2), following the standard procedures in NOM-014-SSA1-1993 (SSA, 1993). Two samples of 1 L were collected, one for PhP determination, and another for Arsenic (As) determination in which 2 mL of nitric acid (HNO3) were added for its preservation for. The samples were transported in coolers, taken to the laboratory and stored at 4 °C until their

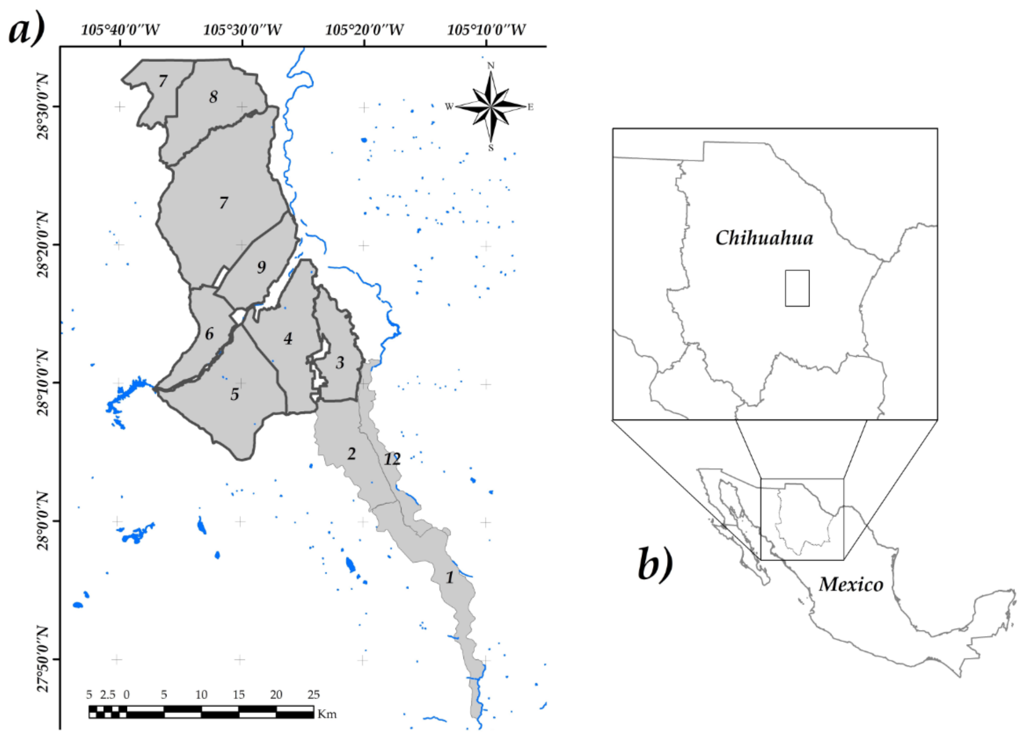

Figure 1. Geographical location of the study area: a) Irrigation District 005 location (■), boundary of the studied area (—), water bodies (—), bold numbers denote the module number, b) Irrigation District 005 location in Chihuahua, Mexico.

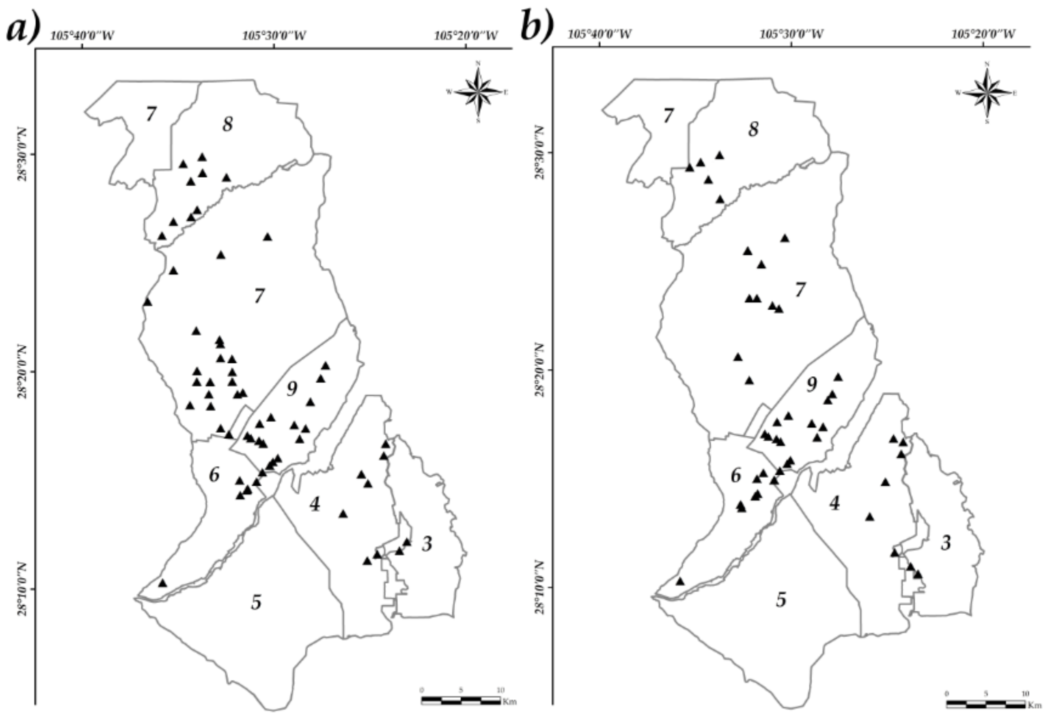

Figure 2. Sampling maps: a) Summer (Sampling 1 - left) and b) Fall (Sampling 2 - right). Studied modules (—), sampling points (▲), bold numbers denote the module number.

analysis. In the first period, samples were collected from 65 wells; while in the second period from 54 wells.

Multivariate Statistical Methods

For As determination, the samples were filtered with 0.2 mm ash paper Whatman No. 41. Subsequently, before the analysis, filtered with 0.45 μm Millipore filters. The As quantification was perform on an Atomic Absorption Spectrometry, Perkin Elmer AAnalyst 700 to which the FIAS 100 Hydride Generator was coupled, by following the standard NMX-AA-051-SCFI-2001 (SCFI, 2001). The detection limit of the equipment was 3.12 μg/L. The samples were analyzed in triplicate using the standard Trace Metals-Sand 1 Number CRM048 Sigma Aldrich with a recovery percentage of 99 %.

Moreover, different physicochemical parameters (PhP) were analyzed: Temperature, Electrical Conductivity (EC), Oxide Reduction Potential (ORP), Hardness, pH, Total Dissolved Solids (TDS) and Turbidity. These are listed in Table 1 along with their respective analytical method. All parameters were analyzed in triplicate. Temperature, pH and ORP were determined in situ, the rest in the laboratory.

Spatial Variability of the Physicochemical Parameters (PhP)

Prior to the multivariate analysis, a Pearson correlation analysis was performed to understand the relationships among the PhP. To know the magnitude of the relationship between the parameters, the Pearson value is classified in 33 % percentiles. The values of Pearson's linear correlation coefficient were classified as: Poor (0.0-0.3), Moderate (0.4-0.6) and Strong

Table 1. Physicochemical parameters on water, units, and analytical method.

Parameter	Unit	Analytical method
As	mg/L	AAS Perkin Elmer Aanalist 700, coupled HG FIAS 100
Temp	^oC	Potentiometer Hanna portable (in situ)
EC	µS/cm	Electrical conductivity meter CYBERSCAN
ORP	mV	Potentiometer Hanna portátil (in situ)
Hardness	mg/L	Titration net NET (indicator)
pH	Adimensional	Potentiometer Hanna portable (in situ)
TDS	mg/L	Electrical conductivity meter CYBERSCAN
Turb	NTU	Electrical conductivity meter CYBERSCAN

As = Arsenic.
EC = Electrical Conductivity.
Turb = Turbidity.
ORP = Oxide Reduction Potential.
Temp = Temperature.

(0.7-1.0, Simeonov *et al*., 2003; Jothivenkatachalam *et al*., 2010). Such analysis was performed in the SAS© 9.1.3 software (SAS, 2006). The multivariate analysis on the data from the ID005 was carried out by the PCA and CA methods (Wunderlin *et al*., 2001; Jolliffe, 2002). The PCA is a method for pattern recognition that attempts to explain the variance of a large set of correlated variables (PhP); transforming the data set into a smaller set of independent (uncorrelated) principal components (PC). SAS© 9.1.3 software was used to describe these patterns. The PCA is a dimensionality reduction technique that helps to simplify the data and make it easier to visualize by looking for a PC set (Shrestha and Kazama, 2007). The PCs are orthogonal variables calculated by multiplying the original correlated variables with a list of coefficients that can be described as shown in Equation 1:

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + . . . + a_{im}x_{mj} \quad (1)$$

where: z = the component´s coefficient, a = component weight, x = measured value of the variable, i = component number, j = sample number, and m = number of variables.

The CA is an unsupervised pattern recognition technique that describes the structure of a data set (Al-Bassam, 2006). The hierarchical grouping is the most common approach in which groups are formed sequentially, starting with the pair of most similar objects forming groups from the union of these objects. The euclidean distance usually gives the similarity between two objects or groups of objects (Shrestha and Kazama, 2007). The resulting groups of objects should exhibit high internal homogeneity (within a group) and high external heterogeneity (among groups), where grouping is typically illustrated with a dendrogram (McKenna, 2003). The dendrogram provides a visual summary of the clustering processes, presenting an image of the groups and their proximity with a dramatic reduction in the dimensionality of the original data (Ward, 1963). The CA was applied to classify the sampling sites by ascending cluster analysis with the Ward (Eder *et al*., 1994) criterion, using the R^2 as a measure of explanation of variation and pseudo T^2 served to confirm the number of groups (Neil, 2002). It is possible to plot the pseudo values versus the number of clusters. If the values present a sudden change, the value of the group n + 1 that caused the change is a candidate for the number of groups to choose (Arlsan and Turan, 2015). The CA was performed in the SAS© 9.1.3 software.

Physicochemical Parameters (PhP) Analysis

The information of the PhP was used as input data to carry out an interpolation. To examine the spatial distribution of the studied variables, the interpolation method used was Inverse Distance Weight (IDW), available in ArcMap© 10.3 software. The interpolation, through IDW has been widely used to map the spatial distribution of water elements (Varol and Davraz, 2015; Ishaku *et al.*, 2016; Moreno, 2008). The IDW method uses the existing values that are around the area to estimate the concentration of the non-sampled sites. The values of the closest observations will have a greater influence than those that are further away, i.e. the influence decreases with distance (CONAGUA, 2017). Equation 2 shows the algorithm for IDW.

$$Z(S_0) = \sum_{i=1}^{N} \lambda * Z(S_i) \tag{2}$$

where: Z(S_0) = value to be estimated in the place S_0, N = number of observations near to the place to estimate, λ= weight assigned to each observation to be used, decreases with distance, and $Z(S_i)$ = observed value of the place S_i.

RESULTS AND DISCUSSION

Analysis of Physicochemical Parameters

Table 2 shows the results obtained from the analysis of As and the PhP analyzed to water samples, the maximum permissible levels (MPL) were established according to Mexican regulations for each parameter and the percentage of samples exceeding said limit.

In the two seasons, TDS was the parameter with the highest percentages of samples exceeding the MPL of the Mexican regulation (35 % in S1, 39 % in S2). Turbidity exceeded the MPL in S1 (29 %) in more samples than S2 (12 %). As concentrations were above the MPL of water for agricultural irrigation in 9 % (Summer) and 13 % (Fall) of the wells.

Multivariate Analysis

The correlation analysis reported the existence of significant positive and negative correlations ($P>0.05$ and $P<0.0001$) among the values of PhP from the first sampling (Table 3). As was moderately correlated with Turbidity, Hardness and pH. EC was moderately correlated with TDS, Hardness and ORP. Regarding TDS, it was moderately correlated with Turbidity and Hardness. Finally, pH and ORP were correlated strongly and moderately, respectively, with Temperature. The poor correlation between the other pairs of PhP indicates the presence of other sources of variation.

In S2 (Table 3) it was observed that As was moderately correlated with Hardness while it was strongly correlated with pH. The EC was moderately correlated with Hardness and strongly correlated with TDS. Likewise, TDS was moderately correlated with Turbidity, Hardness and ORP.

Table 2. Range of concentrations and maximum permissible levels according to Mexican regulations of the PhP analyzed.

Parameter	Concentration Range S1	Concentration Range S2	MPL	Normative	Sampling 1 (%)	Sampling 2 (%)
As (mg/L)	ND – 0.338	ND – 0.576	0.100	CONAGUA, 2017	9	13
Temp (°C)	22.1 – 30.1	22.8 – 27.5	-	Without regulation	-	-
EC (µS/cm)	13.8 - 1981.6	553.6 - 2600	-	Without regulation	-	-
ORP (mV)	85.6 - 267.7	98.1 - 306.3	-	Without regulation	-	-
Hardness (mg/L)	13.3 - 814	0 - 611	500	SSA, 1994	9	5
pH	7.5 - 9.6	7.3 - 9.0	6.0 - 9.0	CONAGUA, 2017	1.5	0
TDS (mg/L)	0 - 990	0 - 932.3	500	CONAGUA, 2017	35	39
Turb (NTU)	0 -1000	0.2 - 519	10	CONAGUA, 2017	29	12

As = Arsenic.
EC = Electrical Conductivity.
Turb = Turbidity.
ORP = Oxide Reduction Potential.
Temp = Temperature.

Table 3. Pearson correlation among the PhP in the wells of the ID005.

Sampling 1								
	As	EC	TDS	Tur	Hardness	Ph	ORP	Temp
As	1.00							
EC	0.07	1.00						
TDS	-0.17	0.625**	1.00					
Turb	0.42	-0.01	-0.452**	1.00				
Hardness	-0.477**	0.493**	0.586**	-0.23	1.00			
pH	0.441*	0.08	-0.04	0.17	-0.348*	1.00		
ORP	-0.092*	-0.44	-0.17	-0.18	-0.09	-0.398*	1.00	
Temp	0.389*	0.327*	0.23	0.13	-0.17	0.827**	-0.462*	1.00

Sampling 2								
	As	EC	TDS	Turb	Hardness	pH	ORP	Temp
As	1							
EC	-0.02	1						
TDS	0.08	0.89**	1					
Turb	-0.14	-0.2	-0.55**	1				
Hardness	-0.46*	0.54**	0.45*	-0.15	1			
pH	0.77**	-0.14	-0.03	-0.07	-0.60**	1		
ORP	0.03	-0.33*	-0.41*	0.32*	-0.13	-0.04	1	
Temp	-0.25	-0.19	-0.06	-0.21	0.04	-0.29*	-0.04	1

As = Arsenic.
EC = Electrical Conductivity.
Turb = Turbidity.
ORP = Oxide Reduction Potential.
Temp = Temperature.
* =Significative P>0.05.
** = Highly significant P >0.0001.

Principal Components Analysis (PCA)

The assumption that the parameters are linearly related was verified, then the PCA was carried out to explore the relationships among the eight PhP. The first four PCs in S1 explained 87 % of the variance (Table 4). In S1, PC1 contributed with 34 % of the variance, PC2 with 30 %, while PC3 and PC4 contributed with 12 % and 9 %, respectively. The dominant PhP in PC1 were As, pH and Temperature. Considering Table 3, there is a significant correlation between As and pH ($r = 0.44$, $P<0.05$). In PC2, the coefficients that contributed the most were EC, TDS and Hardness. The parameters correlated were: EC and TDS ($r = 0.625$, $P<0.0001$), EC and Hardness ($r = 0.493$, $P<0.0001$) and TDS with Hardness ($r = 0.586$, $P<0.0001$). The PC3 was influenced by Turbidity and PC4 by As and ORP.

Regarding S2, 86 % of the variance was explained by considering four PCs (Table 4). The components contributed with 35 %, 24 %, 16 % and 9 % to PC1, PC2, PC3 and PC4, respectively. The PC1 was influenced by EC, TDS and Hardness with weak coefficients. These parameters strongly and moderately correlated as follows: EC and TDS ($r = 0.89$, $P<0.0001$), EC and Hardness ($r = 0.54$, $P<0.0001$), Hardness and TDS ($r = 0.45$, $P<0.05$). The PC2 was influenced by As and pH, with moderate and highly correlated coefficients ($r = 0.77$, $P<0.0001$). The PC3 explained the Turbidity and Temperature variability, with moderate to strong coefficients and PC4 was influenced by ORP. In Table 3 it was observed that there is a highly significant correlation of EC with TDS and Hardness, which indicates that these three components explain a large amount of variation in the study area.

Table 4. Eigenvectors and eigenvalues of the PhP.

PhP	Sampling 1				Sampling 2			
	PC1	PC2	PC3	PC4	PC1	PC2	PC3	PC4
As	***0.441***	-0.11	0.108	***0.661***	-0.22	***0.56***	0.00	0.24
EC	0.066	***0.533***	0.312	0.261	***0.49***	0.21	0.31	0.07
TDS	-0.13	***0.544***	-0.19	0.318	***0.50***	0.32	0.03	0.17
Turb	0.305	-0.19	***0.727***	0.011	-0.28	-0.27	***0.53***	-0.31
Hardness	-0.33	***0.424***	0.263	-0.03	***0.47***	-0.23	0.20	0.16
pH	***0.518***	0.1	-0.34	-0.16	-0.28	***0.55***	-0.02	-0.03
ORP	-0.29	-0.33	-0.26	***0.604***	-0.29	-0.22	0.26	***0.87***
Temp	***0.48***	0.271	-0.27	-0.04	0.06	-0.26	***-0.72***	0.17
Eigenvalue	2.708	2.473	1.031	0.757	2.8	2.15	1.24	0.75
Varibility	0.338	0.309	0.128	0.094	0.35	0.26	0.15	0.09
Cumulative	0.338	0.647	0.776	0.871	0.35	0.62	0.77	0.86

As = Arsenic.
EC = Electrical Conductivity.
Turb = Turbidity.
ORP = Oxide Reduction Potential.
Temp = Temperature.
PC = Principal Component.
Bold letters indicate the dominant coefficients.

The grouping of the sites is shown in the displacement plane of the first two PCs (Figure 3). The PhP for S1 and S2 were organized into 4 groups. Group 1: As, pH and Temperature; Group 2: EC, TDS and Hardness; Group 3: Turbidity; and Group 4: As and ORP. In regards to S2, Group 1 was composed of: Hardness, EC, TDS; Group 2: As and pH; Group 3: Turbidity and ORP; and Group 4: Temperature. Gebreyohannes *et al*. (2015) determined in their area of study that TDS, Hardness and EC were positively associated and that these in turn were negatively associated with pH and Turbidity.

The comparison plots of PC1 vs PC2 in each sample indicate the displacement through components 1 and 2. These components, together explain more than 60 % of the total variation. On the one hand, As, pH and Temperature in S1 move to the right side, indicating its dominance in PC1. While in S2, only As and pH remain dominant in PC1. On the other hand, observing the displacement in PC2 in S1 the variable with the greatest influence is ORP and in S2 it is Hardness, Temperature, Turbidity and ORP. In the particular case of ORP, the change indicates that in S1 (spring) only its variation was high with respect to PC2, while in S2 (summer) it was a great source of variation for PC1 and PC2. The change in the station strongly influences the way in which the parameters are expressed in the well water, which explains the displacement of the parameters between stations.

Cluster Analysis (CA)

The definition of the number of groups was made considering the value of R^2 and the criterion of pseudo T^2. The value of R^2 indicated that, with four groups, up to 76 % of the variability in S1 was explained, while in S2 with four groups 85%

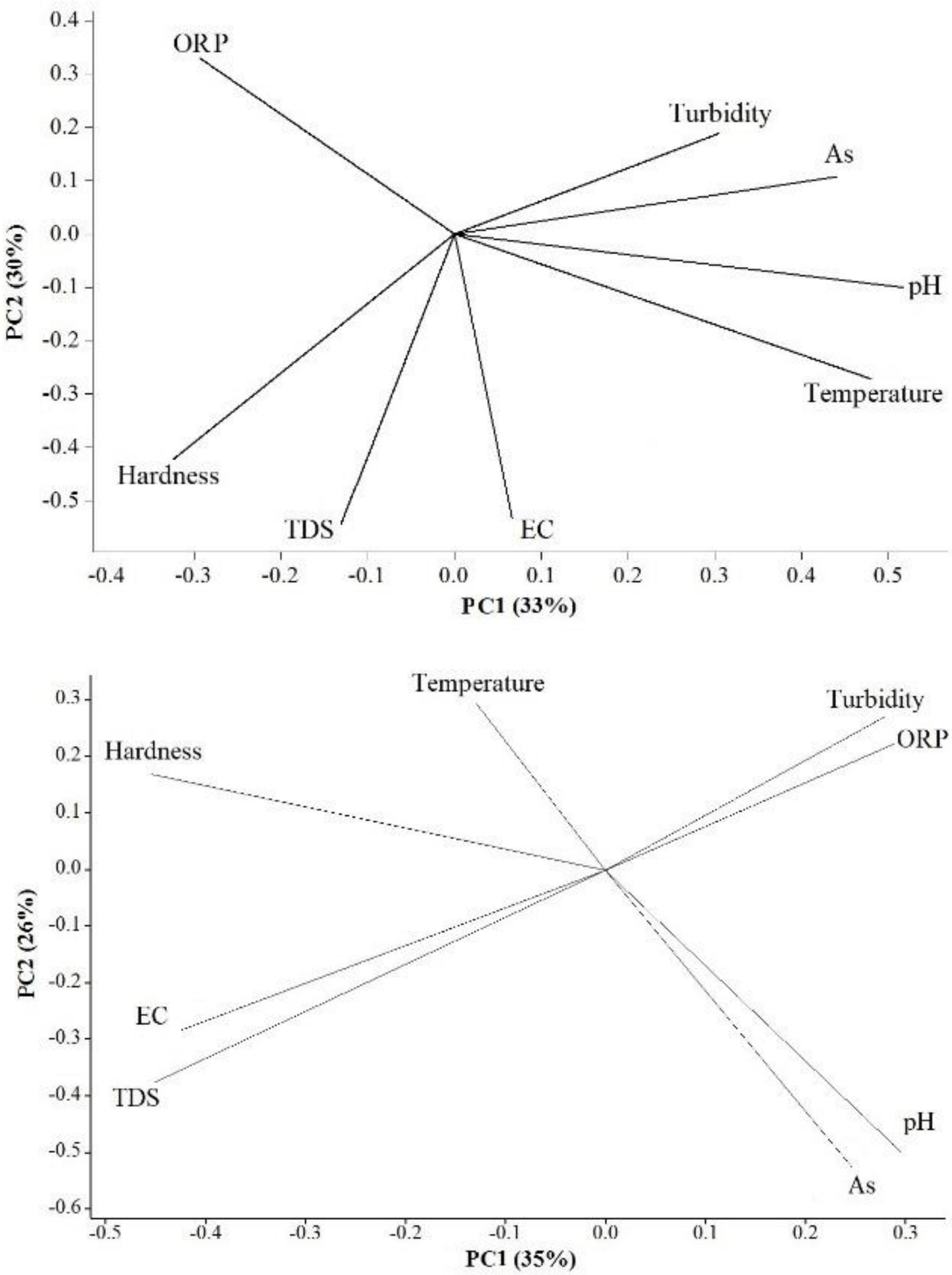

Figure 3. Plot S1 comparison PC1 vs PC2 (Top), Plot S2 comparison PC1 and PC2 (Down).

was explained. A line was drawn in both dendrograms to confirm the value of R^2 (Figure 4). This line is imaginary and is traced in the dendrogram to support the definition of the groups (Güler *et al*., 2002). The pseudo T^2 was useful to reaffirm the decision of the four groups, showing a value of 77.3 and 9.2, respectively (Neil, 2002). The groups were significantly different based on the MANOVA test (F = 25.65, λ of Wilk's = 0.002, P<0.0001).

In S1, Group 1 was made up of 9 wells; Group 2 was the largest with 38 wells; Group 3, the smallest with 7 wells and Group 4 included 9 wells. Each group was characterized with the average of the variables per group, presented in Table 5. In S1, Group 1 consisted of high values of As (0.098 mg/L), Turbidity (687.6 NTU), pH (8.2) and EC (1,117.7 µS/cm), and low values of TDS and Hardness (93.8 and 144.3 mg/L, respectively). Group 2 had moderate values with respect to almost all PhP, only with a low Turbidity value (3.9 NTU). Group 3 also presented moderate values in most of the PhP, with the exception of EC at low concentrations (15.3 µS/cm), and Turbidity at high concentrations (295.6 NTU). Lastly, Group 4 showed high values of TDS (883.9 mg/L), Hardness (497.4 mg/L), EC (1,773.4 µS/cm) and Temperature (25.1 °C), while the lowest values corresponded to Turbidity (38.0 NTU). The remaining PhP had moderate values.

In the S2, Group 1 was the largest with 18 wells; Group 2 was the smallest with 6 wells; Group 3 was comprised of 17 wells; and Group 4 of 9 wells. Group 1 was formed with high Turbidity values (196.1 NTU). Group 2 showed the highest values of Turbidity (164.5 NTU) and the lowest values of As (0.005 mg/L) and TDS (78.9 mg/L) among all groups. Group 3 showed high values of As (0.106 mg/L), EC (1.102.5 µS/cm), TDS (563.4 mg/L), Turbidity (157.6 NTU) and pH

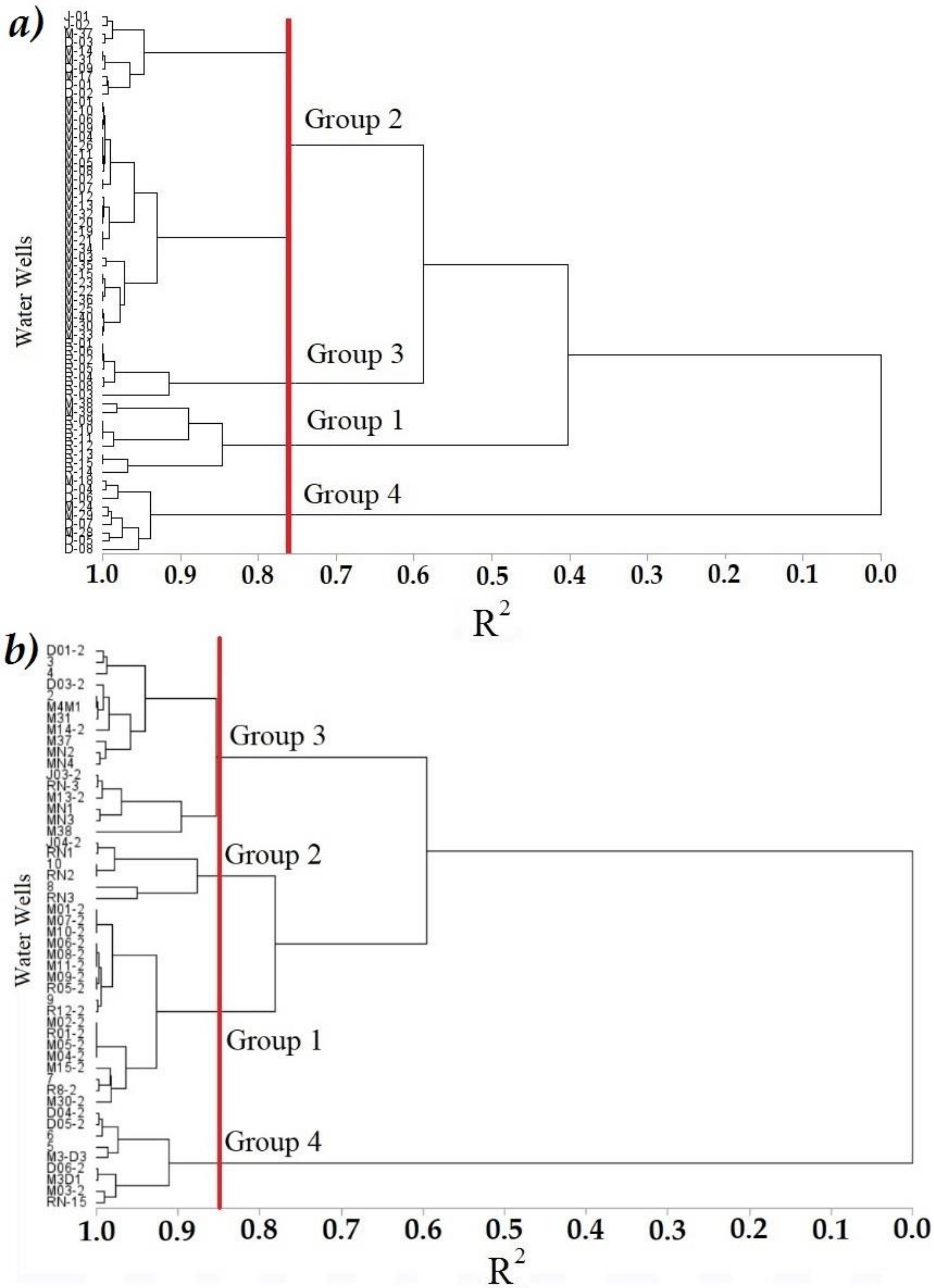

Figure 4. Dendogram. Grouping of sampling sites according to the PhP of ID005 for S1 (a) and S2 (b).

(8.02). Finally, Group 4 had the highest values of EC (1,779.1 μS/cm), TDS (885.3 mg/L), Hardness (384.7 mg/L) and the lowest Turbidity (5.2 NTU), pH (7.3) and ORP (195.2 mV) values when compared to the other groups (Table 5).

The spatial distribution of S1 and S2 was observed by linking the database derived from the CA with the vector file of wells, using ArcMap 10.3©. In S1, the distribution of Group 1 (high values of As, Turbidity and pH) was homogeneous in the northern part of the ID005, located in modules 7 and 8. Group 2, with the highest number of wells (moderate values of all PhP), included modules 4, 7, 8, 9 and 3. Group 3 (high Turbidity), was located in module 6, showing a homogeneous spatial grouping. Group 4 (high TDS, Hardness and EC) was defined in modules 4, 8, 7 and 3, being the group with the greatest geographical dispersion.

In S2, Group 1 (high Turbidity) was homogeneously distributed between module 9 and 6. Group 2 (high Turbidity) was placed in module 6, only with one observation in module 8. Group 3, with high magnitudes of As, EC, TDS, Turbidity and pH, was presented in module 7, with some observations in modules 8 and 4. Group 4, with high magnitudes of EC, TDS and Turbidity, was distributed in the boundary between module 3 and 4 in the southwest portion of the ID005 (Figure 5).

Spatial Variability of Physicochemical Parameters

The maps of the PhP are shown in Figures 6 and 7. The areas with low concentrations are colored in yellow, the blue colored areas represent moderate concentrations while the red colored areas represent high concentrations.

In S1, the PhP As, pH and Temperature showed a similar distribution where

Table 5. Average value of the PhP by groups.

	S1								S2							
G	As	EC	TDS	Turb	Hardness	pH	ORP	Temp	As	EC	TDS	Turb	Hardness	pH	ORP	Temp
1	0.098	1117.65	93.84	687.63	144.34	8.21	150.29	25.25	0.017	686.111	344.552	8.182	196.074	7.701	231.456	24.796
2	0.035	832.62	415.06	3.89	208.06	7.98	208.57	24.42	0.005	768.717	78.938	382.933	164.533	7.665	270.156	23.856
3	0.014	15.25	185.71	295.56	207.03	7.53	230.08	22.59	0.106	1102.486	563.371	5.753	157.647	8.024	210.418	24.170
4	0.008	1773.36	883.91	38.01	497.41	7.77	138.66	25.07	0.005	1779.137	885.278	5.228	384.678	7.525	195.219	24.203

As = Arsenic.
EC = Electrical Conductivity.
Turb = Turbidity, pH = pH, ORP = Oxide Reduction Potential.
Temp = Temperature.

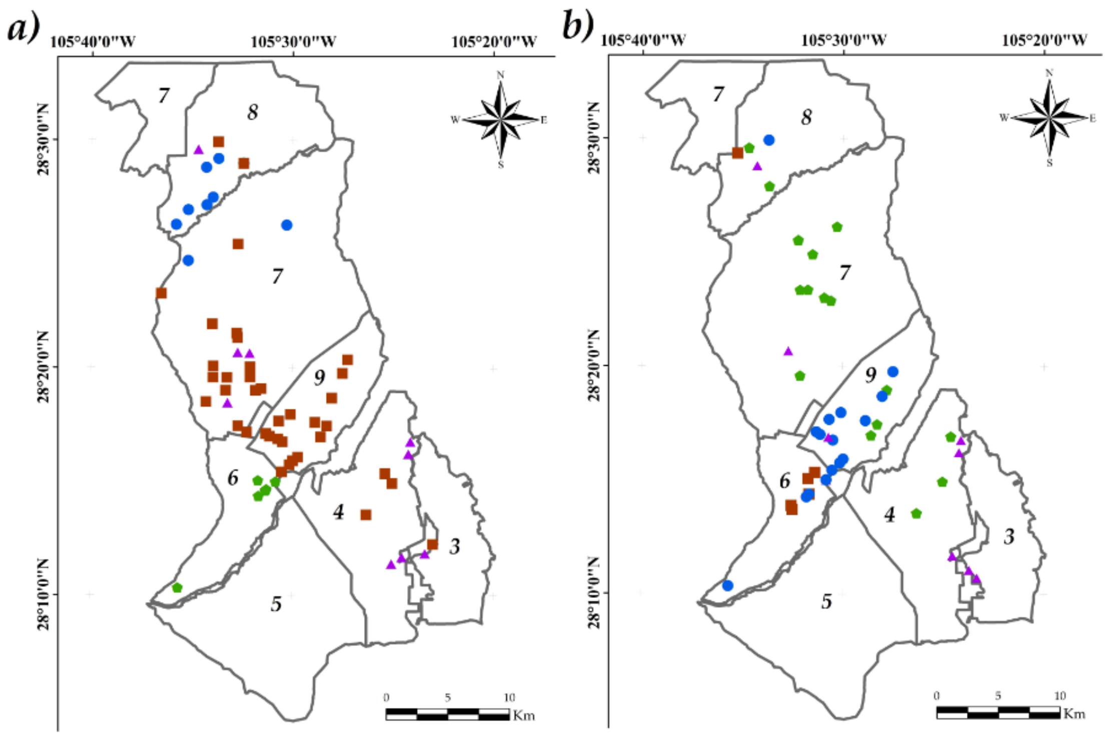

Figure 5. Spatial distribution of the groups in the IDDR005. S1 (left), S2 (right). Group 1 (●), Group 2 (■), Group 3 (⬟), Group 4 (▲), bold numbers denote the module number.

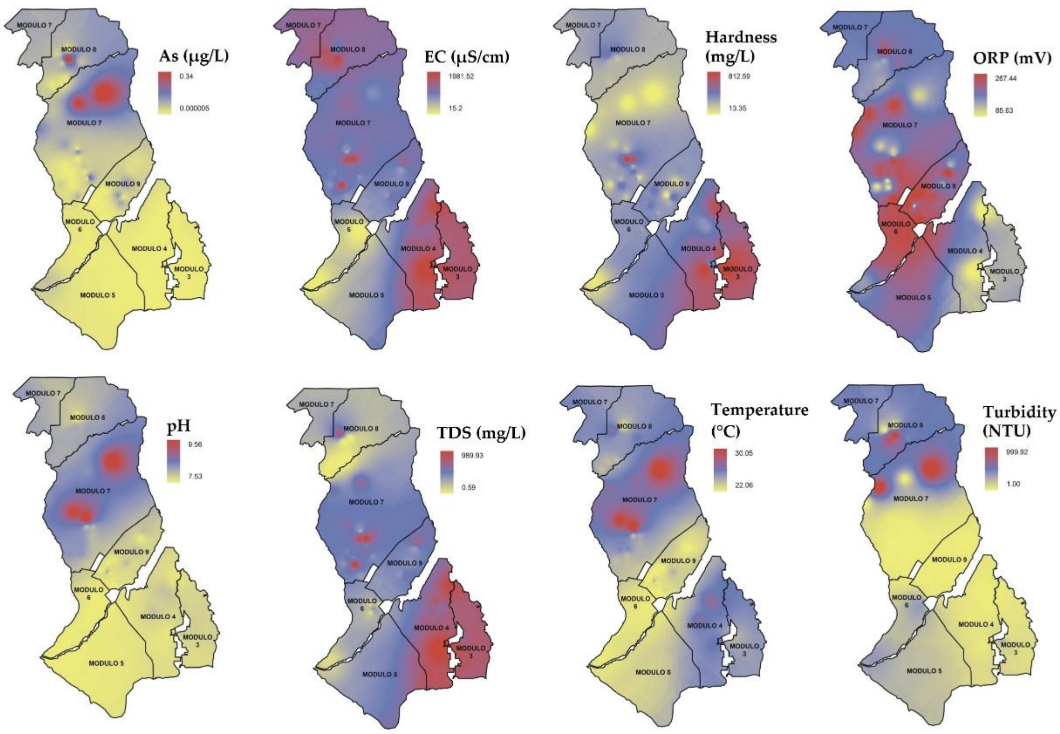

Figure 6. Spatial distribution of the PhP in the ID005 for S1. As = Arsenic, EC = Electrical Conductivity, pH = pH, ORP = Oxide Reduction Potential.

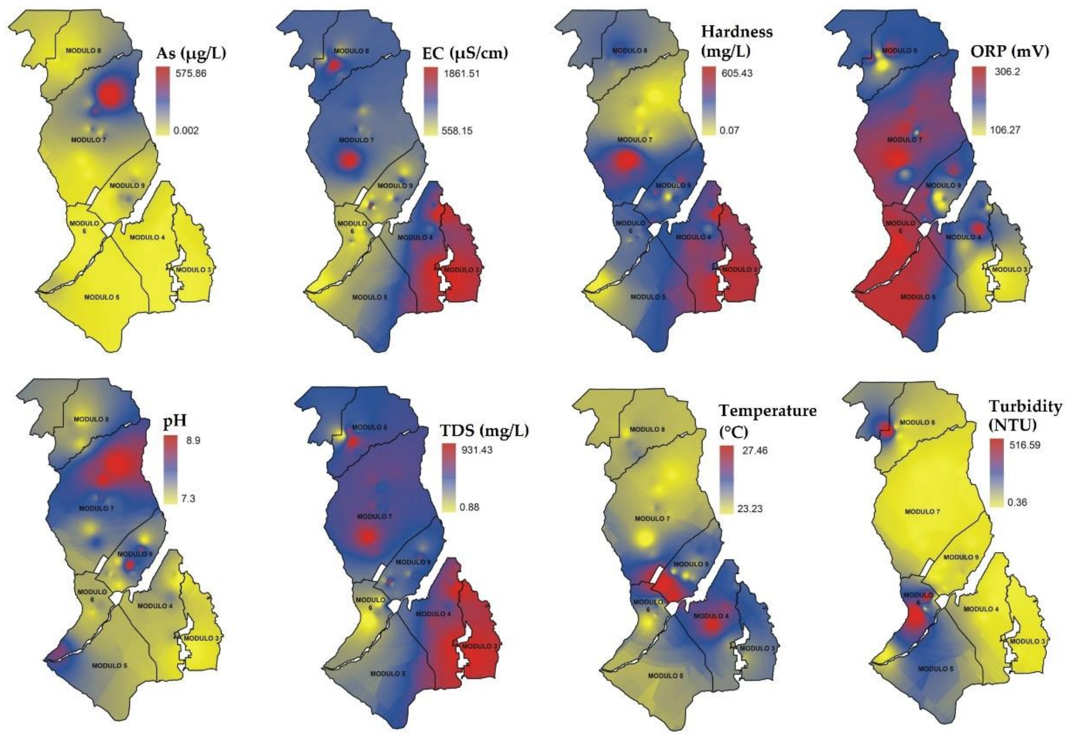

Figure 7. Spatial distribution of the PhP in the ID005 for S2, As = Arsenic, EC = Electrical Conductivity, pH = pH, ORP = Oxide Reduction Potential.

the highest and moderate concentrations are found in modules 6 and 7. The pattern of As (high concentration) may be due to a geological mineralization process (Bu *et al*., 2016), which seems to be present in these modules. Likewise, Hardness and TDS show a similar distribution. The highest concentrations are predominantly distributed in modules 3, 4 and 5. The EC shows a distribution pattern similar to Hardness and TDS but with some variations. The highest concentrations prevailed in modules 3, 4, 7 and 8.

In S2, As and pH showed a similar pattern with high concentrations in the northern part of module 7. The values of EC, TDS and Hardness showed a very similar spatial distribution in modules 8, 7, 4 and 3, at high concentrations. The Temperature and Turbidity PhP presented a similar pattern of high concentrations in module 6. Finally, ORP was the only variable that did not show a spatial behavior similar to the rest of the PhP.

The similarities in the spatial distribution among the PhP confirm the results of the multivariate analysis, where As-Turbidity, pH-Temperature, Hardness-TDS-EC and As-pH were grouped in S2, while TDS-EC-Turbidity were grouped in S1. In a study conducted by Li and Feng (2012), similarities in the spatial distribution of the elements were found. In both, the S1 and S2 samplings, the spatial behavior for As, EC, Hardness, ORP, pH and TDS was similar. The PhP that varied were temperature and Turbidity. Temperature showed a greater variability in S1 compared to S2, which may be associated with the variation of the rest of the PhP.

Likewise, the coefficients of each PC of the PhP were used together with the geographical coordinates of the wells to generate the interpolation of the PCs

(Figure 8). These interpolations spatially indicate the multivariate relationships among the PhP.

In S1, the interpolation of the PC1 coefficients (33 % of the variability explained) indicated that in the yellow areas (negative coefficients) low concentrations of As existed (0.000005 mg/L). These areas registered values of pH between 7.5-7.8 and Temperatures between 22-23 °C. Conversely, the areas in red are those with high concentrations of As (0.33 mg/L), pH (9.5) and Temperatures of 30 °C. PC2 (30 %), influenced by EC, TDS and Hardness, presented negative coefficients (areas in yellow), indicating the presence of low values of EC (13-16 μS/cm), TDS (20-100 mg/L) and Hardness (100-200 mg/L). PC3 (12 %), influenced by Turbidity, depicted areas with negative coefficients indicating the presence of low values for this parameter (1.0 NTU). Meanwhile, positive coefficients corresponded to areas with high concentrations (900 NTU). Although Turbidity has the highest coefficient in the matrix of eigenvectors, Temperature also shows a similar behavior in the database (not shown). In this database, the negative coefficients correspond to zones with temperatures of 30 °C and positive coefficients to areas with temperatures of 24 °C. Finally, PC4 (9 %) is influenced by As and ORP. The areas with negative coefficients correspond to low concentrations of As (0.000005 mg/L) and ORP (85-113 mV). The areas in red correspond to high concentrations of As (0.27-0.34 mg / L) and ORP (250 mV).

In the S2, PC1 (35%) represents the variability of EC, TDS and Hardness. In this component, the red areas represent EC values (1,863 μS/cm), TDS (932.33 mg/L) and Hardness (594 mg/L). The concentrations in yellow are for CE (1,078

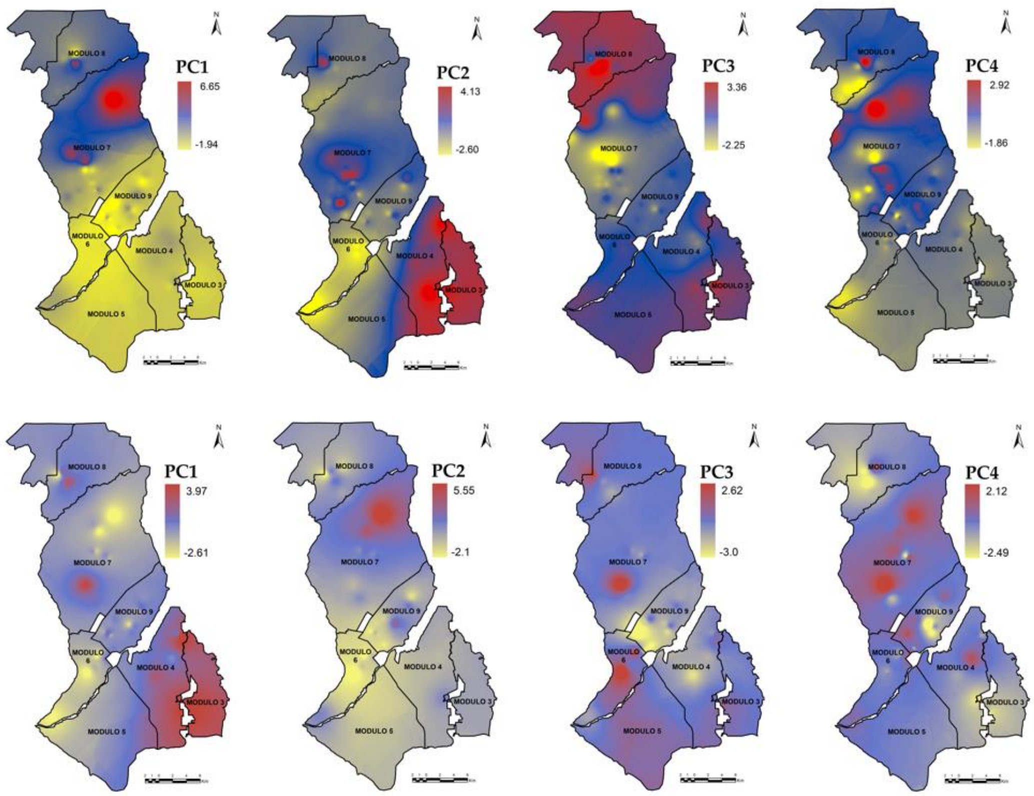

Figure 8. Spatial distribution of the PC scores of S1 (top) and S2 (bottom). PC = Principal component.

μS/cm), TDS (573 mg/L) and Hardness (0 mg/L), which are distributed in module 6 and the northern part of 7. PC2 (26 %) explains the variability of As and pH, where the high concentrations are distributed in module 7. In this module, the coefficients with positive value indicate As concentrations of 0.575 mg/L and pH of 8.9, while in modules 6 and 5, the concentrations are the lowest (As = 0 mg/L, pH = 7.4). The distribution of PC3 (15 %) explains the variation of Turbidity and Temperature. In these zones (red color), the PC explains Turbidity concentrations of 519 NTU and Temperature 23.8 °C. Turbidity values of 0.43 NTU and Temperature of 27.5 °C are reported in the yellow zones. PC4 (9 %) represents the variability of ORP. The zones in red tone indicate ORP concentrations of 306 mV and in yellow values of 106 mV.

The multivariate techniques and the interpolation were useful to interpret the relationship between the PhP. The PCA has been previously used to examine and interpret the behavior of groundwater quality parameters (Sánchez-Martos *et al*., 2001; Liu *et al*., 2003; Chapagain *et al*., 2010; Belkhiri *et al*., 2011). The relationship between the PhP provided significant information on the possible sources of these parameters. In this study, four components were needed to explain the original data set. The four components showed that the behavior of the PhP in the wells was governed by more than one process or phenomenon.

According to Yidana *et al*. (2010), in the analysis of dimensionality reduction of variables PC1 usually represents the most important mixture of processes in the study area. The PCA results suggest that most variations in water quality were found in S1 (summer) for As, pH and Temperature followed by EC, TDS and Hardness; and in S2 (Fall) for EC, TDS and Hardness, followed by

As and pH. According to Bonte *et al*. (2013), the increase in temperatures was associated with the As increase, which was shown in S1 where the main variables that explained the total variability were shown. The above was also demonstrated in the CA and the spatial interpolation of the individual parameters and the main components where high temperature zones show a spatial distribution similar to As.

The variables that were grouped in PC1 and PC2 of each sampling season had similar coefficients, which imply the existence of some similarities in the way they influence the groundwater concentration. It was observed that As, EC, TDS, Hardness and pH were shown in components 1 and 2 in both samples. This was consistent with the results of the interpolation, showing that the distribution of these parameters had a similar dispersion in the ID005. The interpolations of the PC coefficients are similar to the maps of the main components derived from the water sampling from the wells. Previous studies have shown similar results to improve the interpretation of PC (Lu *et al*., 2012; Mueller and Grunsky, 2016). These variables together accounted for more than 60 % of the variability of the original data set. Also, it was observed that there was an exchange of these variables in PC1 and PC2. This behavior may have been caused the result of the different sampling seasons.

These relationships agree with the natural dynamics of water PhP. The pH is the main factor that controls the concentrations of soluble metals (Acosta *et al.*, 2011). As well as, the arid climate leads to evaporation which can interfere in the concentration of As (Brahman *et al*., 2013) and cause seasonal variations. It was observed that As concentrations higher than the MPL (0.1 μg/L) of water for agricultural irrigation established in the Mexican regulation (Ayers and Westcot, 1994), were presented in the northern area of the territory in both S1 and S2. The EC showed a significant correlation with parameters such as Hardness, TDS (Patil *et al*., 2012), which can be related to water salinity (Brahman *et al*., 2013).

In wells that contain high amounts of As, the pH is also high. This was reaffirmed by the CA method and interpolation, where Group 1 showed the highest As and pH values for S1 and S2. The first two main components (PC1 and PC2) in both stations (S1 and S2) showed similar variations. This same behavior was observed in wells where high concentrations of EC, TDS and Hardness were obtained, which was also observed by the CA and spatial interpolation.

For this study, the wells near the city showed the highest concentrations of EC, TDS and Hardness. The EC and TDS measurements for the S1 and S2 samples showed that the salinity is classified according to the Food and Agriculture Organization of the United Nations (FAO) as moderate (EC 700 - 3000 μS/cm, TDS 450 - 2,000 mg/L), especially in the southern area of ID005 (Kolsi *et al*., 2013). EC and high TDS limit the absorption of water by crops because of the salt that stays in the roots. Due to the difficult access to water, the growth rate of plants is reduced, which limits agricultural production (Eder *et al*., 1994). The EC and the TDS in groundwater samples are significantly correlated with cations and anions (Ca^{2+}, K^{+}, Na^{+}, Cl^{+}, NO_{-3} and $SO_4{}^{2}$), which can be the result of ionic changes in the aquifer (Brahman *et al*., 2013).

In the case of Hardness, it was within the MPL established by the Mexican regulation (500 mg/L) (SSA, 1993). However, according to Gebreyohannes et al.

(2015), water with Hardness greater than 151 mg/L is classified as hard water. Considering this criteria, 75 and 77 % of the samples (S1 and S2, respectively) were classified as hard water. This classification may indicate that there are deposits with high Mg^{2+} and Ca^{2+} contents (Gebreyohannes *et al*., 2015). Likewise, it is considered that hard water is not suitable for industrial and agricultural purposes (Patil *et al*., 2012).

The interpretation of the spatial behavior of the main processes present in the study area on water quality was possible when the scores derived from the PC were mapped. Previous studies have used geostatistical methods to map the scores resulting from PCA and used the resultant maps to predict the factors that may be impacting groundwater quality (De Freitas *et al*., 2014; Ha *et al*.,2014). Based on the results of the PCA and the CA it was possible to understand the multivariate relationship of the set of parameters. In turn, with the application of the IDW interpolation technique on the scores of the PCA, it was possible to analyze the spatial variability. The combination of both methods is useful to examine patterns in common groups of parameters allowing to summarize the multiple relationship of variables on geographic regions for use in water quality analysis.

The PCA, the CA, the correlation coefficients and the interpolation were consistent with these interpretations. Although the results of the present study provided important conclusions regarding the origin of each PhP, more studies are needed to obtain a better understanding of the sources of the PhP and their concentrations.

CONCLUSIONS AND RECOMMENDATIONS

The As is an element present in the ID005, it is at levels that can cause a risk to agricultural production, mainly in the northern region. In addition, it is important to continue this investigation to determine the As traceability in the medium and to identify the risk of introducing this metalloid to the food chain by diet intake.

A slight issue was observed with indicators that affect the salinity of water. If such high levels persist, it can be detrimental to the optimal development of crops. Therefore, it will be necessary to look for alternatives to ameliorate this situation. Perhaps, starting with a continuous monitoring of wells in the ID005.

Multivariate statistical methods and spatial interpolation can be useful to identify locations of priority concern and potential sources of PhP, and to evaluate water quality in sludge from an agricultural area. The multivariate geographic information system (GIS) approach showed the spatial relationships between the PhP (As, pH, EC, TDS and Hardness), proving to be convenient for the confirmation and refinement of PhP interpretations through the statistical results.

REFERENCES

Acosta, J.A., A. Faz, S. Martínez-Martínez, R. Zornoza, D.M. Carmona, and S. Kabas. 2011. Multivariate statistical and GIS-based approach to evaluate heavy metals behavior in mine sites for future reclamation. J. Geochem. Explor. 109:8-17.

Aguirre-Grijalva, E. 2003. Estudio de factibilidad técnica y económica de 3 métodos de tecnificación del riego en el Modulo 07 del Distrito 005. Master dissertation, Master degree in Construction Administration, Chihuahua, Chih., Mexico.

Al-Bassam, A.M. 2006. Evaluation of ground water quality in Al-Qassim area, Saudi Arabia, using cluster and factor analyses. Kuwait Journal of Science and Engineering. 33:101.

AlSuhaimi A.O., K.M. AlMohaimidi, and K.A. Momani. 2017. Preliminary assessment for physicochemical quality parameters of groundwater in Oqdus Area, Saudi Arabia. Journal of the Saudi Society of Agricultural Sciences. In press.

Arslan, H., and N.A. Turan. 2015. Estimation of spatial distribution of heavy metals in groundwater using interpolation methods and multivariate statistical techniques; its suitability for drinking and irrigation purposes in the Middle Black Sea Region of Turkey. Environ. Monit. Assess. 187:516.

Ayers, R.S., and D.W. Westcot. 1994. Water quality for agriculture. Food and Agriculture Organization of the United Nations. Rome, Italy.

Batayneh, A., and T. Zumlot. 2012. Multivariate statistical approach to geochemical methods in water quality factor identification; application to the shallow aquifer system of the Yarmouk basin of north Jordan. Research Journal of Environmental and Earth Sciences. 4:756–768.

Behera, S.K., and A.K. Shukla. 2015. Spatial distribution of surface soil acidity, electrical conductivity, soil organic carbon content and exchangeable potassium, calcium and magnesium in some cropped acid Soils of India. Land. Degrad. Dev. 26:71–79.

Belkhiri, L., A. Boudoukha, and L. Mouni. 2011. A multivariate statistical analysis of groundwater chemistry data. Int. J. Environ. Res. 5:537–544.

Bhuiyan, M.A., M.A. Rakib, S.B. Dampare, S. Ganyaglo, and S. Suzuki. 2011. Surface water quality assessment in the central part of Bangladesh using multivariate analysis. KSCE. J. Civ. Eng. 15, 995-1003.

Bhunia, G.S., P.K. Shit, and R. Maiti. 2016. Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC). Journal of the Saudi Society of Agricultural Sciences.

Bonte, M., B.M. van Breukelen, and P.J. Stuyfzand. 2013. Temperature-induced impacts on groundwater quality and arsenic mobility in anoxic aquifer sediments used for both drinking water and shallow geothermal energy production. Water. Res. 47:5088-5100.

Boyacioglu, H. 2006. Surface water quality assessment using factor analysis. Water. S.A. 32:389–393, http://dx.doi.org/10.4314/wsa.v32i3.5264.

Brahman, K. D., T. G. Kazi, H. I. Afridi, S. Naseem, S. S. Arain, and N. Ullah. 2013. Evaluation of high levels of fluoride, arsenic species and other physicochemical parameters in underground water of two sub districts of Tharparkar, Pakistan: a multivariate study. Water. Res. 47:1005-1020.

Brogna, D., A. Michez, S. Jacobs, M. Dufrêne, C. Vincke, and N. Dendoncker. 2017. Linking forest cover to water quality: A multivariate analysis of large monitoring datasets. Water, 9:176.

Bu, J., Z. Sun, A. Zhou, Y. Xu, R. Ma, W. Wei, and M. Liu. 2016. Heavy Metals in Surface Soils in the Upper Reaches of the Heihe River, Northeastern Tibetan Plateau, China. Int. J. Env. Res. Pub. He. 13:247.

Chaoyang, W.E.I., W.A.N.G. Cheng, and Y.A.N.G. Linsheng. 2009. Characterizing spatial distribution and sources of heavy metals in the soils from mining-smelting activities in Shuikoushan, Hunan Province, China. J. Environ. Sci. 21:1230-1236.

Chapagain, S.K., V.P. Pandey, S. Shrestha, T. Nakamura, and F. Kazama. 2010. Assessment of deep groundwater quality in Kathmandu Valley using multivariate statistical techniques. Water. Air. Soil Pollut. 210:277–288.

Chiu, W.C. 2000. Modelación y simulación para el drenaje de tierras, en la planicie aluvial del estado de Tabasco, México por Wilder Camacho Chiu. Doctoral dissertation, Doctor degree in Agricultural Sciences. N.L. Mexico.

CONAGUA (Comisión Nacional del Agua) 2017. Ley federal de derechos Disposiciones aplicables en materia de aguas nacionales. Available online: https://www.gob.mx/cms/uploads/attachment/file/105138/Ley_Federal_de_Derechos.pdf (accessed on 21 June 2017).

de Freitas Alves, S. M., D.M. de Queiroz, G.R. de Alcântara, and E.F. dos Reis. 2014. Spatial Variability of Physical and Chemical Attributes of Soil Using Techniques of Principal Component Analysis. Biosci. J. 30:22-30.

Eder, B.K., J.M. Davis, and P. Bloomfield. 1994. An automated classification scheme designed to better elucidate the dependence of ozone on meteorology. J. Appl. Meteorol. 33:1182-1199.

Fu, S., and C.Y. Wei. 2013. Multivariate and spatial analysis of heavy metal sources and variations in a large old antimony mine, China. J. Soil. Sediment. 13:106-116.

Gebreyohannes, F., A. Gebrekidan, A. Hedera, and S. Estifanos. 2015. Investigations of physico-chemical parameters and its pollution implications of Elala River, Mekelle, Tigray, Ethiopia. Momona Ethiopian Journal of Science. 7:240-257.

Giri, S., and Z. Qiu. 2016. Understanding the relationship of land uses and water quality in Twenty First Century: A review. J. Environ. Manage. 173:41-48.

Güler, C., G.D. Thyne, J.E. McCray, and K.A. Turner. 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. Hydrogeol. J. 10;455-474.

Gupta, I., S. Dhage, and R. Kumar. 2009. Study of variations in water quality of Mumbai coast through multivariate analysis techniques. Indian. J. Mar. Sci. 38:170-177.

Ha, H., J.R. Olson, L. Bian, and P.A. 2014. Rogerson, Analysis of heavy metal sources in soil using kriging interpolation on principal components. Environ. Sci. Technol. 48:4999-5007.

Helena, B., R. Pardo, M. Vega, E. Barrado, J.M. Fernández, and L. Fernández. 2000. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga river, Spain) by principal component analysis. Water Res. 34:807–816.

Ishaku, J.M., B.A. Ankidawa, and A.J.D. Pwalas. 2016. Evaluation of Groundwater Quality Using Multivariate Statistical Techniques, in Dashen Area, North Eastern Nigeria. British Journal of Applied Science & Technology 14:1.

Jackson, B.B. 1983. Multivariate Data Analysis: An Introduction, 0th ed.; Prentice Hall: Irwin, Homewood, Illinois, USA. 154–196, ISBN 978-0256028485

Johnston, K., J.M. Ver Hoef, K. Krivoruchko, and N. Lucas. 2001. Using ArcGIS geostatistical analyst. Redlands ESRI: CA. USA.

Jolliffe, I.T. 2002. Principal Component Analysis, 2 nd.; John Wiley & Sons Inc.: Charlottesville, V.A., USA. 150-166. ISBN 978-0-387-22440-4.

Jothivenkatachalam, K., A. Nithya, and M. S. Chandra. 2010. Correlation analysis of drinking water quality in and around Perur block of Coimbatore District, Tamil Nadu, India. Rasayan Journal Chemistry, 3:649-654.

Kazi, T.G., M.B. Arain, M.K. Jamali, N. Jalbani, H.I. Afridi, R.A. Sarfraz, J.A. Baig, and A.Q. 2009. Shah. Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. Ecotox. Environ. Safe. 72:301-309.

Kengnal, P., M.N. Megeri, B.S. Giriyappanavar, and R.R. Patil. 2015. Multivariate Analysis for the Water Quality Assessment in Rural and Urban Vicinity of

Krishna River (India). Asian Journal of Water, Environment and Pollution. 12:73-80.

Kengnal, P., M.N. Megeri, B.S. Giriyappanavar, and R.R. Patil. 2015. Multivariate Analysis for the Water Quality Assessment in Rural and Urban Vicinity of Krishna River (India). Asian Journal of Water, Environment and Pollution. 12:73-80.

Khan, S., Q. Cao, Y.M. Zheng, Y.Z. Huang, and Y.G. Zhu. 2008. Health risks of heavy metals in contaminated soils and food crops irrigated with wastewater in Beijing, China. Environ. Pollut. 152:686-692.

Kolsi, S.H., S. Bouri, W. Hachicha, and H.B. Dhia. 2013. Implementation and evaluation of multivariate analysis for groundwater hydrochemistry assessment in arid environments: a case study of Hajeb Elyoun–Jelma, Central Tunisia. Environ. Earth Sci. 70:2215-2224.

Li, X., and L. Feng. 2012. Multivariate and geostatistical analyzes of metals in urban soil of Weinan industrial areas, Northwest of China. Atmos. Environ. 47:58-65.

Liu, C.-W., K.-H. Lin, and Y.-M. Kuo. 2003. Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. Sci. Total. Environ. 313:77–89.

Loska, K., and D. Wiechuła. 2003. Application of principal component analysis for the estimation of source of heavy metal contamination in surface sediments from the rybnik reservoir. Chemosphere. 51:723–733.

Lu, A., J. Wang, X. Qin, K. Wang, P. Han, and S. Zhang. 2012. Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China. Sci. Total. Environ. 425:66-74.

Maas, S., R. Scheifler, M. Benslama, N. Crini, E. Lucot, Z. Brahmia, S. Benyacoub, and P. Giraudoux. 2010. Spatial distribution of heavy metal concentrations in urban, suburban and agricultural soils in a Mediterranean city of Algeria. Environ. Pollut. 158:2294-2301.

Markus, J., and A.B. McBratney. 2001. A review of the contamination of soil with lead II. Spatial distribution and risk assessment of soil lead. Environ. Int. 27:399–411.

Maroju, S. 2007. Evaluation of five GIS based interpolation techniques for estimating the radon concentration for unmeasured zip codes in the state of Ohio. Master dissertation, Master of Science Degree in Civil Engineering, Ohio, U.S.

McKenna, Jr. J.E. 2003. An enhanced cluster analysis program with bootstrap significance testing for ecological community analysis. Environ. Modell. Softw. 18. 205e220.

Medellín-Vázquez, J.J. 2003. Análisis de la vegetación en un gradiente altitudinal mediante técnicas multivariadas, en el Campo Santa María, Lampazos de Naranjo, Nuevo León y Candela Coahuila. Master dissertation, Master in Forestry Sciences, N.L. México,

Miranda, J., E. Andrade, A. López-Suárez, R. Ledesma, T.A. Cahill, and P.H. Wakabayashi. 1996. A receptor model for atmospheric aerosols from a southwestern site in Mexico city. Atmos. Environ., 30:3471–3479.

Moreno, J.A. 2008 Sistemas y Análisis de la Información Geográfica. Manual de autoaprendizaje con ArcGIS. 2ª ed, Ra-Ma. Madrid, España. 908. ISBN 978-84-7897-838-0.

Morris, B.L., A.R. Lawrence, P.J.C. Chilton, B. Adams, R.C. Calow, and B.A. Klinck. 2003. Groundwater and its susceptibility to degradation: a global assessment of the problem and options for management Vol. 3. United Nations Environment Programme: Nairobi, Kenya, 2003. 118. ISBN: 92-807-2297-2.

Muangthong, S., and S. Shrestha. 2015. Assessment of surface water quality using multivariate statistical techniques: case study of the Nampong River and Songkhram River, Thailand. Environ. Model. Assess. 187:548.

Mueller, U.A., and E.C. Grunsky. 2016. Multivariate spatial analysis of lake sediment geochemical data; Melville Peninsula, Nunavut, Canada. Appl. Geochem. 75:247-262.

Navarrete Álvarez, M. 2012. Modelos geoestadísticos del precio de la vivienda: aproximación al conocimiento intraurbano de la ciudad de Madrid. Doctoral dissertation. Madrid, España.

Neil, H. T. 2002 Applied multivariate analysis, 1st ed.; Springer: Board, NY, USA. 532, ISBN:0-387-95347-7.

Núñez López, D., E.J. Treviño Garza, V.M. Reyes Gómez, C.A. Muñoz Robles, O.A. Aguirre Calderón, and J. Jiménez Pérez. 2014. Uso de modelos de regresión para interpolar espacialmente la precipitación media mensual en la cuenca del río Conchos. Rev. Mex. Cien. Agrí. 5:201-213.

Oketola, A.A., S.M. Adekolurejo, and O. Osibanjo. 2013. Water quality assessment of River Ogun using multivariate statistical techniques. J. Environ. Prot. 4:466.

Ortega-Gaucin, D. 2012. Reglas de operación para el sistema de presas del Distrito de Riego 005 Delicias, Chihuahua, México. Ing. Agric. Biosist. 4:31-39.

Ortega-Gaucin, D., E. Mejía Sáenz, E. Palacios Vélez, L. Rendón Pimentel, and A. Exebio García. 2009. Modelo de optimización de recursos para un distrito de riego. Terra Latinoamericana. 27:219-226.

Otto, M. 1998. Multivariate methods. In Analytical Chemistry; Kellner, R., Mermet, J.M., Otto, M., Widmer, H.M. Eds. Wiley-VCH, Weinheim, Germany. 916. ISBN 3-527-28881-3, DM 98.

Patil, P.N., D.V. Sawant, and R.N. 2012. Deshmukh. Physico-chemical parameters for testing of water-A review. Int. J. Environ. Sci. 3:1194.

Qadir, A., R.N. Malik, and S.Z. Husain. 2008. Spatio-temporal variations in water quality of Nullah Aik-tributary of the river Chenab, Pakistan. Environ. Monit. Assess. 140:43-59.

Rawlins, B.G., R.M. Lark, R. Webster, and K.E. O’Donnell. 2006. The use of soil survey data to determine the magnitude and extent of historic metal deposition related to atmospheric smelter emissions across Humberside, UK. Environ. Pollut. 143:416–426.

Sánchez-Martos, F., R. Jiménez-Espinosa, and A. Pulido-Bosch. 2001. Mapping groundwater quality variables using PCA and geostatistics: A case study of Bajo Andarax, southeastern Spain. Hydrol. Sci. J. 46:227–242

SAS (Statistical Analysis Software) Institute 2006. SAS Software. Version 9.1.3 Cary: SAS Inc.

SCFI (Secretaría de Comercio y Fomento Industrial) 2001. NMX-AA-051-SCFI. Análisis de agua medición de metales por absorción atómica en aguas naturales, potables, residuales y residuales tratadas- método de prueba. Available online: http://www.economia-nmx.gob.mx/normas/nmx/2010/nmx-aa-051-scfi-2016.pdf (accessed on 15 may 2017).

Sharma, R.K., M. Agrawal, and F. Marshall. 2007, Heavy metal contamination of soil and vegetables in suburban areas of Varanasi, India. Ecotox. Environ. Safe. 66:258-266.

Shrestha, S., and F. Kazama, 2007. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. Environ. Modell. Softw. 22:464-475.

Simeonov, V., J.A. Stratis, C. Samara, G. Zachariadis, D. Voutsa, and A. Anthemidis. 2003. Assessment of the surface water quality in Northern Greece. Water. Res. 37:4119–4124.

Singh, K.P., A. Malik, and S. Sinha. 2005. Water quality assessment and apportionment of pollution sources of Gomti river (India) using multivariate statistical techniques: a case study. Analytica Chimica Acta. 538. 355e374, Singh, K.P., A. Malik, D. Mohan, and S. Sinha. 2004. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)—a case study. Water. Res. 38:3980-3992.

Skeppström, K., and B. Olofsson. 2006. A prediction method for radon in groundwater using GIS and multivariate statistics. Sci. Total. Environ. 367:666-680.

SSA (Secretaría de Salud) 1993. NOM-014-SSA1. Procedimientos sanitarios para el muestreo de agua para uso y consumo humano en sistemas de abastecimiento de agua públicos y privados. Available online: http://www.salud.gob.mx/unidades/cdi/nom/014ssa13.html (accessed on 15 may 2017)

Varol, S., and A. Davraz. 2015. Evaluation of the groundwater quality with WQI (Water Quality Index) and multivariate analysis: a case study of the Tefenni plain (Burdur/Turkey). Environ. Earth. Sci. 73:1725-1744.

Vega, M., R. Pardo, E. Barrado, and L. Deban. 1998. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. Water. Res. 32:3581–3592.

Villatoro, M., C. Henríquez, and F. Sancho. 2008. Comparación de los interpoladores IDW y Kriging en la variación espacial de pH, Ca, CICE y P del suelo. Agron. Costarric. 32.

Walter, A.M., S. Christensen, and S.E. Simmelsgaard. 2002. Spatial correlation between weed species densities and soil properties. Weed. Res. 42:26-38.

Wang, S., G.H. Huang, Q.G. Lin, Z. Li, H. Zhang, and Y.R. Fan. 2014. Comparison of interpolation methods for estimating spatial distribution of precipitation in Ontario, Canada. Int. J. Climatol. 34:3745-3751.

Wang, Y., P. Wang, Y. Bai, Z. Tian, J. Li, X. Shao, L.F. Mustavish, and B.L. Li. 2013. Assessment of surface water quality via multivariate statistical techniques: a case study of the Songhua River Harbin region, China. Journal of hydro-environment research. 7:30-40.

Ward Jr, J.H. 1963. Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. 58:236-244.

Wunderlin, D.A., M.P. Diaz, M.V. Ame, S.F. Pesce, A.C. Hued, and M.A. Bistoni. 2001. Pattern recognition techniques for the evolution of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba-Argentina). Water Res. 35:2881–2894.

Xie, Y., T.B. Chen, M. Lei, J. Yang, Q.J. Guo, B. Song, and X.Y. Zhou. 2011. Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: accuracy and uncertainty analysis. Chemosphere, 82:468-476.

Yan, W., Q. Mahmood, D. Peng, W. Fu, T. Chen, Y. Wang, S. Li, J. Chen, and D. Liu. 2015. The spatial distribution pattern of heavy metals and risk assessment of moso bamboo forest soil around lead–zinc mine in Southeastern China. Soil. Till. Res. 153:120-130.

Yan-Feng, Z.H.A.O., S.H.I. Xue-Zheng, B. Huang, Y.U. Dong-Sheng, W.A.N.G. Hong-Jie, S.U.N. Wei-Xia, I Öboern, and K. Blombäck. 2007. Spatial Distribution of Heavy Metals in Agricultural Soils of an Industry-Based Peri-Urban Area in Wuxi, China. Pedosphere 17:44-51.

Yidana, S.M., B. Banoeng-Yakubo, and T.M. Akabzaa. 2010. Analysis of groundwater quality using multivariate and spatial analyses in the Keta basin, Ghana. J. Afr. Earth. Sci. 58:220-234.