# Complex networks, an innovative methodology for functional zoning in forest management

Elizabeth Serrano-Ramírez [1],
José René Valdez-Lazalde [1],
Roman Anselmo Mora-Gutiérrez [2],
Héctor Manuel de los Santos-Posadas [1],
Gregorio Ángeles-Pérez [1]

Forest management planning requires a permanent collection of data on the distribution, composition, and structure of the stands that conform a woodland. These data serve as the basis for suggesting the most appropriate management scheme according to the natural resource conditions and management objectives. It is common for the collected databases' structure and dimension to hinder their analysis using traditional descriptive techniques. Therefore, alternative methodologies are required to facilitate both the exploration of data properties and their collective behavior. We used complex networks analysis to identify distribution patterns of topographic, biological, and productive conditions of a managed forest, suggesting its functional zoning. The forest was considered a graph consisting of nodes and edges; the stands served as nodes and interactions between them as edges. Degree, clustering coefficient, triangles, and modularity were used as segregation and connectivity metrics to evaluate forest properties and allocate stands to five predefined potential forest uses (zones). The clustering coefficient metric provided the better graph partition, allowing to obtain the best alternatives for zoning the forest in conservation areas, areas with potential for timber production, and carbon storage. Proposing forest functional zoning through complex network theory is a powerful methodological option to represent the spatial and non-spatial interactions among the relevant attributes defining a forest ecosystem condition.

Keywords: Forest Planning, Spatial Interactions, Segregation And Connectivity Metrics, Graph Theory

## Introduction

Current forest management process is a complex task that requires frequent updating on the distribution, composition, structure, and modification of the forest due to the various silvicultural practices applied, particularly harvesting (Zou et al. 2019).

Despite the technological advancement in computer systems (Zanin et al. 2016) and equipment for evaluating and monitoring forest resources (Matese 2020), there is

still the need to incorporate new approaches that efficiently generate supporting information for the planning of forest management programs. New methods should facilitate and strengthen decision-making processes associated with the sustainable use of forest resources.

Data that describe forest resources are constantly changing and show multidimensional spatial characteristics that make difficult their analysis using traditional descriptive techniques. Several methods have been used to characterize and group the multiple attributes that define a forest, commonly using geographic clustering techniques (Mahfuz et al. 2019, Zou et al. 2019). Most of these methods have been enhanced to cater to big data sets and relate forest variables to geographical patterns. Some popular are the Iterative Self Organizing Data Analysis (ISODATA), the Density-Based Spatial Clustering of Applications with Noise (DBSCAN), the Spatial "K"luster Analysis by Tree Edge Removal (SKATER), or the regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP – Kupfer et al. 2012, Liu et al. 2019).

However, techniques using algorithms such as the REDCAP or SKATEX directly incorporates a spatial contiguity constraint into a traditional hierarchical clustering which could be helpful if the intention is to

create homogeneous areas or when there is no need to have prior knowledge of the number of regions (Barbierato et al. 2020, Lawal 2020). Instead, suppose the purpose is to characterize the multiple interactions of spatial and non-spatial forest variables per management unit, dividing the entire dataset into separate groups of stands but not creating homogeneous areas. In that case, it is more suitable to use a non-hierarchical method that could complement the forest analysis attributes (Karkra et al. 2020, Opach et al. 2020).

The Network theory is a novel approach that can tackle the above cases. It facilitates the representation of interactions between the elements that conform the forest system. This theory emphasizes the visualization of systems present in nature as complex networks, which allows extracting knowledge of both the individual characteristics that compose the system (nodes) and its collective properties through the presence of links or connections between nodes (Cestero & Caballero 2018, Velandia 2020).

Considering the forest as a complex system allows to explain the reaction of different system components to various stimuli, such as forest harvest, and identify areas with different potentials for use (functional zones) to diversify the forest management scheme. Functional zoning is a process by

✉ (1) Forest Sciences Department, COLPOS, Mexico C.P. 56230 (Mexico); (2) Systems Department, UAM - Iztapalapa, Mexico City C.P. 02210 (Mexico)

@ José René Valdez-Lazalde
(valdez@colpos.mx)

which management units (stands) are classified into homogeneous strata according to their environmental characteristics and productive potential. Each zone (stratum) is defined by the quality and quantity of goods and services potentially obtained in each of them (Liu et al. 2019).

This study proposes an alternative methodology for functional forest zoning under the complex networks approach. The forest was considered a graph, where the stands were the nodes and the connectivity between stands the graph links. Connectivity and segregation metrics were used to analyze the graph properties and to identify areas with similar topographical, biological, and productive conditions. Resulting functional zones with different productive potential were proposed to diversify the forest management scheme.

## Material and methods

### Relevant complex networks theory

Mathematically, a complex network is made up of nodes $N = \{n_1, …, n_N\}$ linked by a set of links $L = \{a_{11}, …, a_{ij}\}$ where $a_{ij} = \{n_i, n_j\}$ denotes the link that joins two nodes of the network. The nodes can represent any kind of objects or individuals described by variables, and the links the interactions between them (Bullmore & Sporns 2009).

The data collection of as many variables as are available about the objects and interactions that make up the network, are helpful to describe it through a mathematical summary commonly known as an adjacency matrix (A). This matrix will represent the structural connectivity and the collective behavior patterns of the nodes in the network. The number of rows and columns of the A matrix will correspond to the number of nodes in the network, and the values inside the matrix ($a_{ij}$) will indicate the existence/type of the link between nodes (Rubinov & Sporns 2010, Latora et al. 2017).

The existence of a link between two nodes ($n_i$ and $n_j$) can be determined by assessing their similarity or "distance" in one or several given variables (characteristics). A popular method to compute such similarity is the Mahalanobis distance ($Md_{ij}$), which has been commonly used to perform classification in multi-dimensional data (Galeano et al. 2014, Berrendero et al.

2020). $Md_{ij}$ is defined as follows (eqn. 1):

$$Md_{ij} = \frac{1}{k}\left(Z_{ij}^{T} C^{-1} Z_{ij}\right) \qquad (1)$$
$$\forall \ j=1,2,…,n \ \ i=1,2,…,k$$

where $Z_{ij} = (X_{ij} - m_i)/s_i$ is the standardized observation of the $i$-th variable; $C^{-1}$ is the inverse of the correlation matrix; $X_{ij}$, $s_i$, and $m_i$ are the observations, the standard deviation and the mean of the $i$-th variable, respectively; and $k$ is the number of variables.

$Md_{ij}$ is calculated between each pair of nodes and the average of this measure for the data set (complex network) is obtained. Then, to build the adjacency matrix (A), a link will be set between two nodes if their computed distance is less than the average value of $Md_{ij}$ (Montes-Orozco et al. 2020). The main advantage of using the Mahalanobis distance, unlike the Euclidean distance, is that all the variables used to compare the nodes are considered equally important as it takes into account the average value, the ranges of acceptability (variance) between variables, and compensates for interactions (covariance) between them (Srinivasaraghavan & Allada 2006).

Depending on the links' nature, the adjacency matrix (A) can be directed or undirected, and it also can be binary or weighted. In an undirected matrix, the order of the indexes on the links does not alter the network's configuration; a link between node $n_i$ and node $n_j$ can be expressed indistinctively as $a_{ij}$ or $a_{ji}$. In a directed matrix, the order of the indices is relevant, the expression of the link between node $n_i$ and node $n_j$ can be $a_{ij}$ but not $a_{ji}$. A weighted matrix occurs when their links have a numerical value indicating the union's intensity (number of links or a cost value). In contrast, a binary matrix occurs if only the presence or absence of a link is considered (Zanin et al. 2016, Latora et al. 2017).

Various metrics can describe quantitatively the structure and connectivity of the network using the calculated adjacency matrix. These metrics can refer to the network elements (nodes or links) or the distribution of all of them (Fig. 1). Typically, individual element metrics reflect how a node is embedded in the network, while distribution metrics provide a more global network description (Rubinov & Sporns

2010, Jiménez 2017). The most common metrics for evaluating segregation and connectivity are degree, clustering coefficient, triangles, and modularity.

The degree ($K_i$) of a node is the number of connections ($a_{ij}$) linked to the rest of the network. The degree is the most fundamental network metric, and most other metrics are related to it (Hernández-González et al. 2020). $K_i$ is calculated as (eqn. 2):

$$K_i = \sum_{j=1}^{N} a_{ij} \qquad (2)$$

The clustering coefficient ($C_i$) calculates the clustering trend of the nodes. If a node is connected to its neighboring nodes and all of them are connected, then $C_i$ will have a maximum value of one. If its neighbors barely interact with each other, the value will be close to zero and, if there is no interaction, the value will be zero (Rubinov & Sporns 2010). This coefficient is given by the following (eqn. 3):

$$C_i = \frac{2T_i}{K_i(K_i-1)} \qquad (3)$$

$T_i$ is the triangles metric and represents the number of triangles that pass-through node $n_i$. In other words, the links of the nodes $n_i$ and $n_j$ from the neighborhood of $n_i$. $T_i$ is the sum of the cycles of length three, between a source node and its neighbors, for all the network nodes. The "triangle" can also be understood as a clique of size three. A clique is a graph (set of nodes) in which every pair of different nodes are adjacent, and there is an edge (link) that connects them (Jiménez 2017). The triangles metric is calculated as (eqn. 4):

$$T_i = \frac{t_e}{t_p} \qquad (4)$$

where $t_e$ and $t_p$ are the number of existing and possible triangles, respectively. Finally, modularity ($Q_i$) evaluates a community structure in the network, that is, subnets densely connected between them (Zanin et al. 2016). A high modularity value exists if the network has subnets densely connected internally but little connected externally. On the other hand, a low modularity value will be obtained if the network does not have dense subnets, but these are very well connected. Modularity calculates the minimum number of links that need to be removed to disconnect the network (eqn. 5):

$$Q_i = \frac{1}{2L}\sum_{ij}\left[A_{ij} - P_{ij}\right]B_{ij} \qquad (5)$$

where $P_{ij} = k_i k_j/2L$ and $B_{ij} = \delta(c_i, c_j)$. $L$ is the number of links in the network, $A_{ij}$ is the adjacency matrix, $P_{ij}$ is the probability that a link between two nodes is proportional to their degrees, and $B_{ij}$ is a binary measure that defines whether the nodes are from the same community. A community is a high concentration of links in some areas of the network (Clauset et al. 2004).
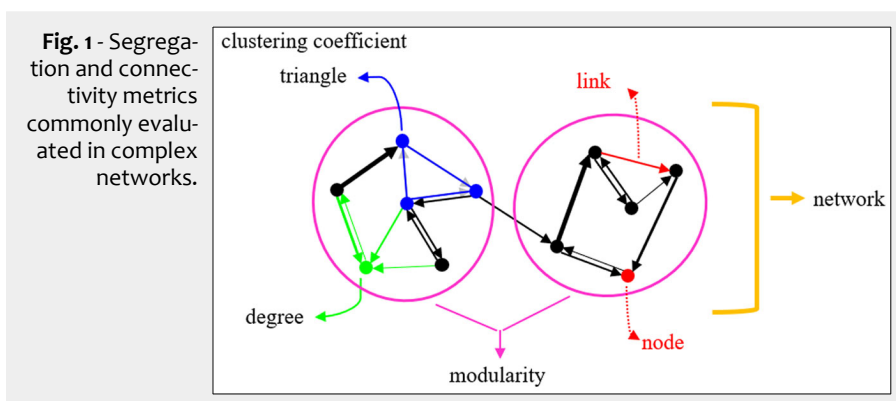
**Fig. 1** - Segregation and connectivity metrics commonly evaluated in complex networks.

*Data collection and processing*

The proposed methodology for defining a functional zoning scheme through a complex networks analysis (Fig. 2), at its first step, needs defining the system for analysis and its networks. Here we defined the forest as the system for complex network analysis, the stands as the nodes, and the existing connectivity (similarity) between stands as the links of the network.

The forest owned by the community of San Pedro El Alto, Oaxaca, Mexico was used as study case. This community owns 29,250.58 ha of temperate forest, which it manages sustainably through its Community Forestry Company. The forest consists of 1827 minimum management units (stands), and the current Forest Management Plan (2016-2017 / 2025-2026) includes topographic, biological, and production data per stand.

The descriptive variables available for the stands allowed us to build five complex networks. Four networks (R1-R4) included variables that describe the stand by its geophysical, dasometric, habitat diversity, or productivity attributes (Tab. 1). The fifth network (R5) included all the information available per stand. Each one of the networks were made up of 1827 nodes (stands), and the existing links between them expressed their similarity.

After all data were standardized by computing the Mahalanobis distances between nodes, binary adjacency matrices were generated for networks R1-R4, setting the value of one if a connection (link) between a pair of stands existed, and zero otherwise. Differently, for the R5 network, the $a_{ij}$ positions took integer values from 0 to 4. These values resulted from the union of the matrices of the R1-R4 networks. The maximum possible value of the link is four (an existing link in all the R1-R4 networks), and the minimum possible value is zero if there are isolated stands (without connections in all the networks R1-R4). Due to isolated stands identified in some matrices, the matrices' size in networks R2, R3, and R4 were 1824, 1824, and 1715 nodes. These procedures were coded in MATLAB® software version R2020b (The MathWorks Inc. 2020).

Once the five adjacency matrices were obtained, the next step was to calculate the connectivity and segregation metrics on an a-per-stand basis, as described in the previous section. The degree ($K_i$) metric was used to evaluate the connectivity in networks' structure, while the triangles ($T_i$), clustering coefficient ($C_i$), and modularity ($Q_i$) were used to classify the stands into five classes (functional zones) previously defined by a panel of experts (Tab. 2). Fig. 2 synthesizes the flow of tasks performed to implement the described procedure. The triangles and clustering coefficient were calculated with the Latapy method (Latapy 2008), while modularity was calculated with the Louvain method (Blondel et al. 2008). The open-source software Gephi

**Tab. 1** - Data per stand included in each complex network (R1-R5) proposed.

| Network | Stand data (descriptor variables) |
|---------|-----------------------------------|
| R1 | Area (ha), perimeter (m), current land use, slope (%), and potential productive category. |
| R2 | Average total volume ($m^3ha^{-1}$), mean diameter at breast height (cm), average site index (SI), age (year), average basal area ($m^2ha^{-1}$), average number of trees per hectare, average estimated mean annual increment ($m^3ha^{-1}year^{-1}$), and habitat category (defined by tree genera dominance ). |
| R3 | Average volume per tree genera ($m^3ha^{-1}$). |
| R4 | Proportion per tree genera of the average total timber volume (0 - 1 values). |
| R5 | The conjunction of the data considered in the R1-R4 networks. |

version 0.9.2 (201709241107) was used for this purpose.

To minimize the variation in each class, we used the Natural Breaks (Jenks) algorithm incorporated in the ArcGIS® software version 10.3.1 (ESRI Inc. 2015), which has proven good performance in zoning processes (Liu et al. 2019). With this algorithm, similar values are better grouped, and differences between classes are maximized.

Each functional zone represents a combination of desirables criteria (productivity, slope, and habitat) suitable for a particular production scheme. For example, low-intensity harvesting (extensive management) is suggested in low-productivity stands with steep slopes to avoid harvesting over forest growth capacity and reduce soil erosion. Thereby, to determine the classes belonging to a specific functional zone, the representative characteristics of

the class (stand attributes) were matched with the desirable criteria of functional zones.

Because of the number of nodes and links existing in the complex networks (R1-R5), their visualization by drawing the network partition was useless as it looked like a fuzzy graph. For that reason, we choose the alternative of mapping the segregation metrics values per stand using ArcGIS® software for each complex network.

## Results

The results were analyzed in terms of the structure of the complex networks (degree values) and the performance of the segregation metrics triangles ($T_i$), clustering coefficient ($C_i$), and modularity ($Q_i$) to classify the stands in the five functional zones proposed. A summary of the descriptive statistics per metric evaluated in each complex
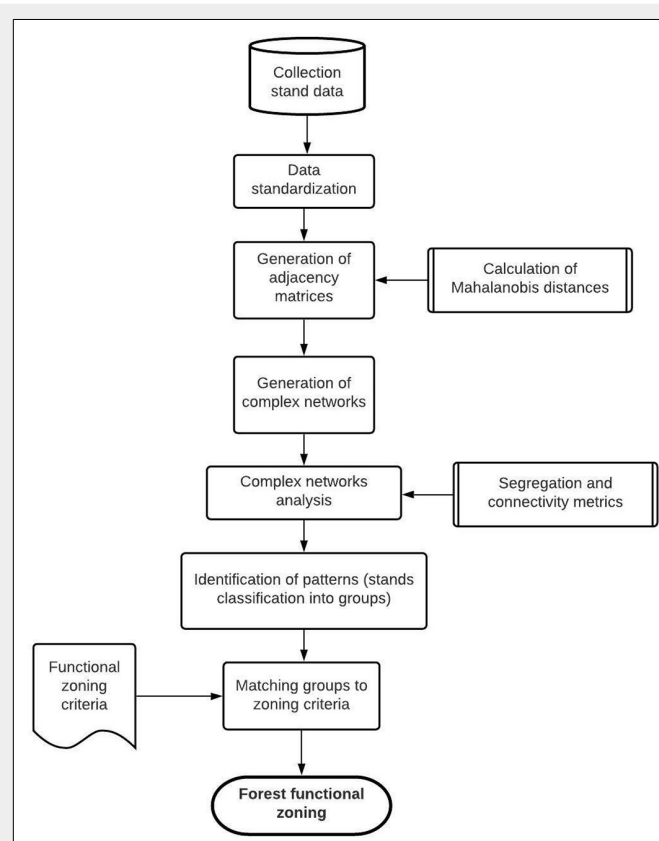


**Fig. 2** - Flow chart of the proposed methodology to forest functional zoning through complex networks analysis.

**Tab. 2** - Desirables stand attributes for the forest functional zones proposed. (BC): biodiversity conservation forest; (CS): carbon storage forest; (LpFE): low productivity forest suitable for extensive management; (HpFE): high productivity forest suitable for extensive management; (HpFI): high productivity forest suitable for intensive management.

| Zones | Desirable criteria |
|-------|--------------------|
| BC | Stands with superior tree diversity (most trees genera present) and distributed in slopes steeper than 42%) or difficult to access. |
| CS | Stands with a medium to high site index (SI), with slopes less than 42%, sites with dominance of *Pinus, Abies*, and *Quercus* genera and preferentially distributed on the periphery of the property. |
| LpFE | Stands with a low site index (IS), with slopes of 29-42%, and sites with dominance of *Pinus* genera. |
| HpFE | Stands with a medium to high site index (SI), with slopes of 15-28%, and sites with dominance of *Pinus* genera. |
| HpFI | Stands with a high site index (SI), with slopes of 0-14%, and sites with dominance of *Pinus* genera. |

network is available in the Tab. S1 (Supplementary material).

### Structural analysis of the complex networks

All the evaluated networks were of the non-directed type. That is, the origin-destination order of the link present between two stands (nodes) is indifferent. However, the R1-R4 networks differ from the R5 by their weight. Networks R1-R4 have terms $a_{ij}$ of binary type (zero if there is no link, one otherwise) in their adjacency matrices, so they are considered simple non-weighted networks. On the contrary, the R5 network has terms $a_{ij}$ with values greater than or equal to zero, which indicate the magnitude of the connectivity, that is why it is considered a weighted network.

Regarding the network structure, evaluated by the degree metric ($K_i$), a high probability of finding stands (nodes) with a degree much higher than the average is evident (Fig. 3). Although the dimensions of the networks are different, and therefore, the maximum and minimum degree values differ between networks, the behavior of the estimated density functions for each network are pretty similar. The degree value distribution in all networks presents a negative asymmetry value where the tail of the distribution is lengthened for values below the mean. However, in the R5 network, the data were highly concentrated around the mean due to its sharp distribution curve (Fig. 3).

The highest mean value of the triangles metric ($T_i$) was obtained by the network R5 with 707,570 cycles of length three. In comparison, the R2 network got the minimum mean value of 171,230 cycles, which is also a high number of triangles. Therefore, like the previous metric, these values suggest a high connectivity between the stands that compose the forest. However, negative kurtosis values in networks R1, R2, R3, and R5 indicate a higher proportion of stands with a high level of connectivity; meanwhile, the R4 network, which has a positive kurtosis value, shows that it includes more stands with a similar connection level (see Tab. S1 in Supplementary material).

In the clustering coefficient metric ($C_i$), the values obtained for R1-R4 networks were in the range of 0-1; and for the R5 network, the values were from 0.6 to 1. The highest mean value (0.9) was obtained in the R4 network and the minimum value (0.7) in R1 and R2 networks. This metric suggests a high clustering tendency between stands in all the networks due to the numerous interactions. Unlike the behavior shown by the triangles metric, most networks' clustering values are highly concentrated around the mean, except for R2, which presents a kurtosis value of -0.78 (see Tab. S1 in Supplementary material).

Finally, the modularity ($Q_i$) computed different values for R1-R5 networks. The range values for R1-R3 networks were 0-3, 0-4, and 0-2, respectively. In comparison, modularity values of R4 and R5 networks were higher, in a range of 0-5. But the average values calculated for the five networks were in a range of 0.5-1.4. Low modularity values mean that the network does not have densely connected communities (subnets densely isolated), but the subnets are well connected to each other (see Tab. S1 in Supplementary material).

### Forest functional zoning

In general, the classification of the stands in one of the proposed functional zones as a function of the segregation metrics evaluated per network (triangles, clustering coefficient, and modularity) showed consistency in the proportion of allocated area to each zone (class). However, there were zones in some networks in which no stand was allocated (Fig. 4).

The zone with the largest assigned area (number of stands) by all the metrics was HpFI (high productivity forest suitable for intensive management) with a range of 22-47% of the total forest area. Secondly, the HpFE zone (high productivity forest suit-
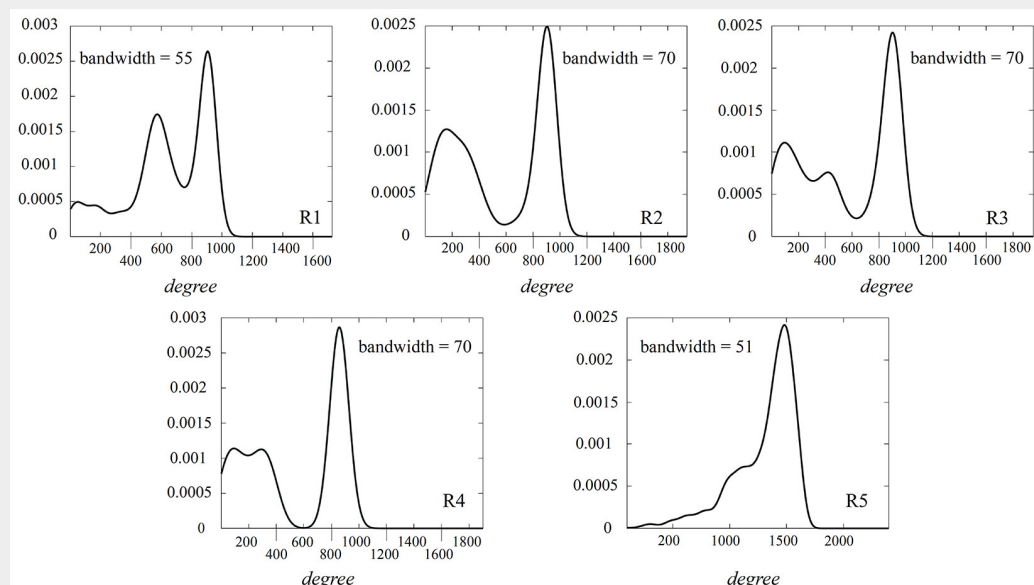


**Fig. 3** - Probability densities estimated with Gaussian kernel for degree values in the proposed complex networks (R1 - R5).

**Fig. 4** - Allocation of stands by functional zone, according to the metrics evaluated for complex networks (R1-R5) proposed. (BC): biodiversity conservation forest; (CS): carbon storage forest; (LpFE): low productivity forest suitable for extensive management; (HpFE): high productivity forest suitable for extensive management; (HpFI): high productivity forest suitable for intensive management.
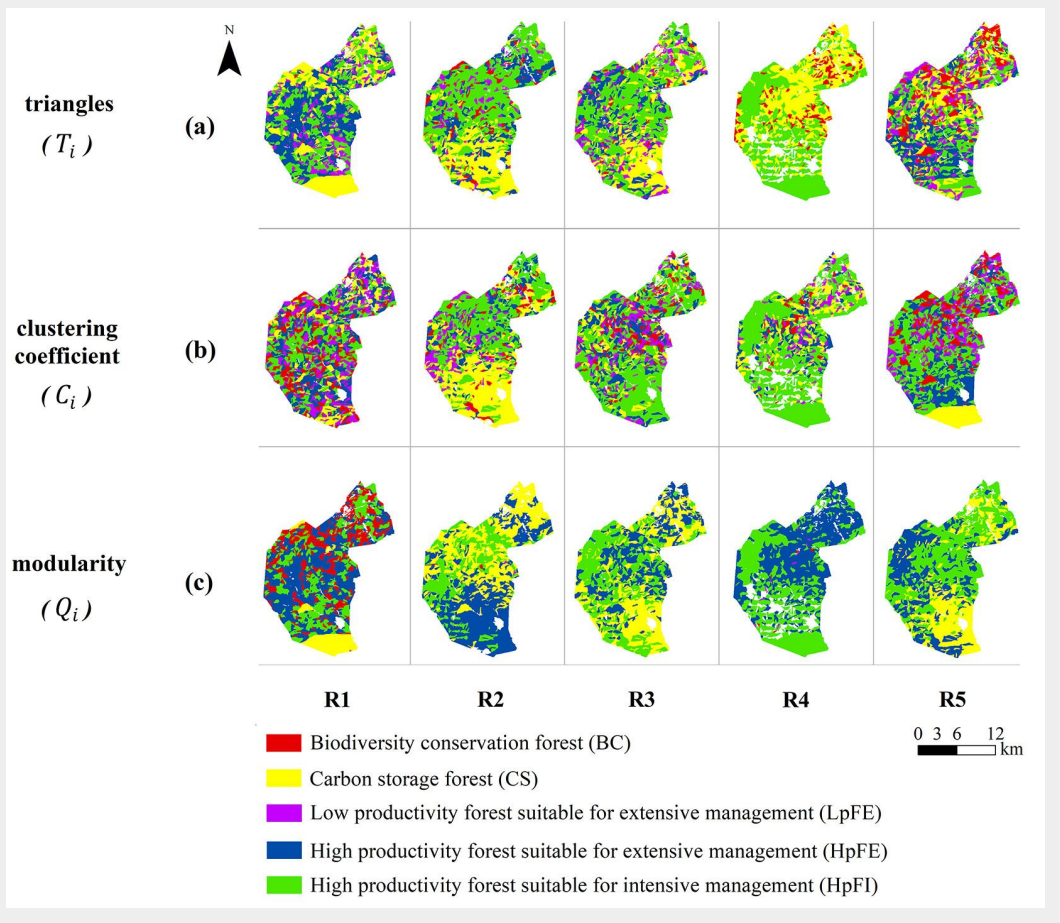


able for extensive management) with 0-46% of the total forest. The CS zone (carbon storage forest) occupied 0-39% of the area; the BC zone (biodiversity conservation forest) occupied 0-23%, and the zone with the smallest area was the LpFE (low productivity forest suitable for extensive management) with a range of 0-21% of the territory (Fig. 5).

If the spatial distribution of the stands that conform a zone is of interest, it can be seen that the triangles metric, calculated with all the stand's variables available, generated zones spatially dispersed. Differently the same metric, using only geophysical and dasometric data, generated carbon storage and timber production agglomerated zones (Fig. 4a).

Contrary, the clustering coefficient metric created a single spatial agglomeration for the carbon storage zone using both all the variables per stand and dasometric variables, whereas using only the geophysical data, this metric generated a scattered scenario for all the zones (Fig. 4b). Finally, the modularity metric agglomerated the stands suitable for carbon storage and timber production in all the networks. Except in the network R1 (geophysical data), the biodiversity conservation zone was excluded from the zoning scheme (Fig. 4c).

Considering the spatial patterns described above, an expert panel selected as the best match between the stands allocation and the desirable criteria for each functional zone, the scenario obtained in the R5 complex network evaluated by the clustering coefficient metric. This functional zoning scenario was compared with the current classification of stands in the study area and validated as a viable scheme of forest zoning (see Fig. S1 in Supplementary Material for the comparing scenarios).

The present forest zoning in the San Pe-

**Fig. 5** - Area (ha) allocated to each functional zone, determined by the metrics evaluated in the complex networks (R1-R5) proposed.
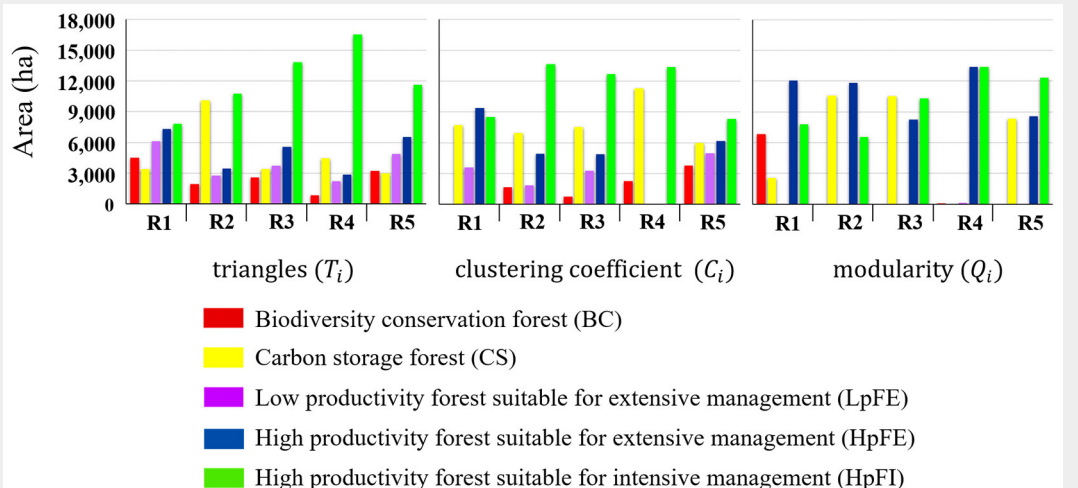
dro El Alto community is focused mainly on biodiversity conservation and intensive timber production. In contrast, the final functional zoning proposal showed the possibility to add productive potentials (different harvesting intensities and carbon sequestration areas) into their forest management scheme.

## Discussion

### Structural analysis of the complex networks

The high values of degree ($K_i$), triangles ($T_i$), and clustering coefficient ($C_i$) metrics in all the complex networks indicate that stands are connected significantly at stand level. This "connectivity" refers to the linkage existing between two nodes of the network. Those stands highly connected can be identified as the ones which have a high probability of dominating the existing interactions in the forest (system). In case of severe harvesting or eliminating any of these stands, there will likely be a significant impact on the rest of the forest's ecological dynamics, *i.e.*, the distribution patterns of animal and plant species would be affected. Therefore, the activities carried out in any of these stands will indirectly affect the dynamics and condition of adjacent and neighboring stands.

Despite the modularity metric ($Q_i$) identifying stand communities into the forest, these are well connected both externally and internally. There is not an isolated community that can disappear by not having interaction in the system. A drastic removal or alteration of a large number of stands would be required to stop or prevent the flow of information and interactions (*i.e.*, dispersal of biological material) that exist between the subnets or communities identified in the forest. A community, in this case, is made up of stands that share common functional characteristics and relationships.

To increase the adaptiveness of forest management plans to future environmental changes and promote the ecosystem services prevalence, we need to develop methodologies that incorporate landscape structural metrics at the stand scale such as proximity, land cover classes, diversity, and size (Cosović et al. 2020, Rodrigues et al. 2021). In response to this need, several studies have developed different methodological frameworks to explore the relationships among the landscape functions and define a proper multifunctional forest zoning (Müller et al. 2020, Baskent et al. 2021, Marques et al. 2021, Rodrigues et al. 2021). However, none has used complex network theory to date.

Complex networks theory has been recently applied to diverse scopes such as medicine, logistics and transportation, the World Wide Web, and in ecology to model ecological niches and species food chains (Amaral & Ottino 2004, Amaral & Uzzi 2007). In the forestry field, we identified the work by Messier et al. (2019) in which complex networks theory was used to analyze the spatial connectivity of forest stands in terms of seed dispersal and tree establishment capacity as a mechanism to improve forest resilience to global change. Beside this work, we did not find any other evidence about the application of complex network theory in forest management planning, specifically in forest functional zoning.

### Forest functional zoning

Independent of the information available for each complex network, all the forest stands were successfully classified into a functional zone by all the connectivity metrics. R2 and R3 networks allocated similar areas to the defined functional zones as a result of close connectivity metrics. This result is logical considering that the R2 network includes both dasometric and productivity variables per stand. Meanwhile, the R3 network synthesizes the set of dasometric variables' information through the available timber volume per stand as a variable. On the contrary, the metrics calculated for the R3 and R4 networks evaluated the system's behavior (forest) differently, although both contained similar information, *i.e.*, the available timber volume by genus, and the proportion by genus present in the stand, respectively. Finally, network R5, which included more information and assigned a frequency value to existing links between stands (weighted network), generated considerably different values in the evaluated metrics and, consequently, a different definition of functional zones. In all cases, the validation of specialists and decision-makers is essential to discern which classification is the most appropriate.

Spatially aggregated zoning schemes could facilitate operational forest management, since less spatial aggregation presents a greater challenge to implement day-to-day conservation and silvicultural practices (Liu et al. 2019). In our study, the zoning patterns that showed the highest spatial aggregation were determined by the modularity metric. However, this metric did not include all functional zones. That is why the expert panel and the community selected the zoning scheme presented by the clustering coefficient metric in R5 (Fig. 4). It achieved a better partition of the graph (network) by proposing the best classifications of all the forest's functional zones while producing a not negligible aggregation between them.

The San Pedro El Alto forest presents a surplus of quality timber inventory, which denotes a highly productive forest (De los Santos-Posadas et al. 2015). Thus, the criteria for allocating functional zones favor the prevalence of areas that benefit from the forest productiveness, as are the timber production and carbon storage zones. In contrast, the allocation of stands to the biodiversity conservation zone is much smaller. However, the latter is not because it is unimportant but due to the assignation criteria defined for this particular zone.

As a spatial entity, the quality of the forest is heavily influenced by geographical factors. Therefore, tends to present specific zoning patterns, although these patterns may behave differently as functional characteristics of the ecosystem are integrated (Wang et al. 2021). The methodology proposed in our study successfully analyzed the interactions of multiple criteria related to forest productivity, intensity of forest activities, and the ecological importance of the forest to build a zoning scheme. Compared with the current classification of the stands in the community, the functional zoning scenario proposed shows the possibility of expanding the productive potential the forest.

Multiple optional spatial distribution schemes on the multifunctionality of forest ecosystems increases precision in landscape management (Myasnikov 2018, Peng et al. 2019). This justifies the current demand to develop and adopt innovative methodologies for defining functional zoning schemes (Gao et al. 2014). Therefore, our results are outlined as valuable information to support forest managers in a previous stage before developing management plans, increasing the efficiency and the effectiveness in forest management planning.

### Limitations and future research directions

In this study, only five functional zones were considered according to the ecological and economic benefits that these implied to the community (forest owners). Also, all data used to analyze the forest as a complex system came from the current management plan (no new data recompilation).

Even though we arbitrarily selected the zoning criteria and available stand information, the general results show an easy way to incorporate complex network theory in forest management planning. The proposed methodology considers geographical attributes, productivity aspects, and biodiversity characteristics of the forest, besides the participation of the owners in the decision-making process of functional zoning. However, we emphasize that no environmental characteristics, ecosystem functions other than carbon storage, and socio-cultural priorities were considered in the analysis.

To efficiently achieve ecological and economic development objectives, and to reduce negative interactions in forest management areas, further exploration of zoning methods is needed. This challenge can be met by developing a conceptual framework for multifunctional ecosystem management (forestry) in which information systems (as the one presented in this study) are an important element to facilitate the analysis of alternative manage-

ment scenarios and, at the same time, support structured owners' participation on decision making (Zemp et al. 2017, Martínez-García et al. 2020, Marques et al. 2021).

## Conclusions

Complex network analysis is an innovative approach that allows taking advantage of all the information available on the conditions, characteristics, and functions identified in forest ecosystems for allocating stands to a predefined zoning scheme. It has two main benefits: (i) it facilitates the processing of large amounts of data composed of qualitative and quantitative variables; and (ii) it enables the deduction of forest properties that are not identifiable at first glance but results from the interactions between forest elements.

Although zoning is not new in forest management planning, the functional zoning strategy through the complex network approach shows a distinctive quality to encompass topographic, dasometric, production, productivity, diversity, and connectivity attributes associated with the ecosystem into the classification process. The methodology proposed in our study also showed a great capacity to operationalize a functional zoning scheme for the forest easily.

The key contribution of complexity theory to forest management is their practical analysis to achieve functional zoning and the support for decision-makers in the definition of management systems that explicitly recognize the multifunctional value of forests. In a subsequent stage, the proposed work could be integrated into optimization processes to assess the feasibility of different management scenarios, that deal with diverse political and organizational problems in forest management.

## List of abbreviations

SI: site index; BC: biodiversity conservation forest; CS: carbon storage forest; LpFE: low productivity forest suitable for extensive management; HpFE: high productivity forest suitable for extensive management; HpFI: high productivity forest suitable for intensive management.

## Acknowledgments

## References

Amaral LAN, Ottino JM (2004). Complex networks - Augmenting the framework for the study of complex systems. The European Physical Journal B 38: 147-162. - doi: 10.1140/epjb/e2004-00110-5

Amaral LAN, Uzzi B (2007). Complex systems - A new paradigm for the integrative study of management, physical, and technological systems.

Management Science 53: 1033-135. - doi: 10.1287/mnsc.1060.0696

Barbierato E, Bernetti I, Capecchi I, Saragosa C (2020). Integrating remote sensing and street view images to quantify urban forest ecosystem services. Remote sensing 12: 1-22. - doi: 10.3390/rs12020329

Baskent EZ, Borges JG, Vacik H, Reynolds KM, Rodriguez CE (2021). Management of multiple ecosystem services under climate change, bioeconomy and participation. Forests 3: 1-7. - doi: 10.3390/f12010104

Berrendero RJ, Bueno-Larraz B, Cuevas A (2020). On Mahalanobis distance in functional settings. Journal of Machine Learning Research 21: 1-33. [online] URL: http://www.jmlr.org/papers/volume21/18-156/18-156.pdf

Blondel VD, Guillaume J, Lambiotte R, Lefebvre E (2008). Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment 2008: 1-12. - doi: 10.1088/1742-5468/2008/10/P10008

Bullmore E, Sporns O (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. Nature Reviews Neuroscience 10: 312-312. - doi: 10.1038/nrn2618

Clauset A, Newmany MEJ, Moore C (2004). Finding community structure in very large networks. Physical Review E 70: 1-6. - doi: 10.1103/PhysRevE.70.066111

Cestero EV, Caballero MA (2018). Data science y redes complejas: métodos y aplicaciones [Data science and complex networks: methods and applications]. Centro de Estudios Ramon Areces S.A., Madrid, Spain, pp. 21-32. [in Spanish]

Cosović M, Bugalho MN, Thom D, Borges JG (2020). Stand structural characteristics are the most practical biodiversity indicators for forest management planning in Europe. Forests 3: 1-24. - doi: 10.3390/f11030343

De los Santos-Posadas HM, Valdez-Lazalde JR, Torres-Rojo JM (2015). San Pedro El Alto community forest, Oaxaca, Mexico. In: "Forest Plans of North America" (Siry JP, Bettinger P, Merry K, Grebner DL, Bostosn K, Cieszewski C eds). Elsevier, London, UK, pp. 199-208.

ESRI Inc. (2015). ArcGIS for Desktop version 10.3.1. ESRI Inc., Redlands, CA, USA.

Galeano P, Joseph E, Lillo RE (2014). The Mahalanobis distance for functional data withy applications to classification. Technometrics 57: 281-291. - doi: 10.1080/00401706.2014.902774

Gao Y, Feng Z, Wang Y, Liu JL, Li SC, Zhu YK (2014). Clustering urban multifunctional landscapes using the self-organizing feature map neural network model. Journal of Urban Planning and Development 140: 1-11. - doi: 10.1061/(ASCE)UP.1943-5444.0000170

Hernández-González S, Hernández-Torres JE, Hernández-Ripalda MD (2020). Análisis de las ventas empleando redes complejas: comunidades y centralidad de productos [Sales analysis using complex networks: communities and product centrality]. Ingeniería Industrial 2020: 177-191. [in Spanish] - doi: 10.26439/ing.ind2020.n038.4820

Jiménez CA (2017). Dinámica y efectividad de las fanpages de Facebook de candidatos a gobernador en los resultados electorales [Dynamics and effectiveness of the Facebook fanpages of gubernatorial candidates in the electoral re-

sults]. Innovaciones de Negocios 13: 221-238. [in Spanish]

Karkra R, Kaur S, Kaur M, Sharma R, Upadhyay RR (2020). Management zone delineation in precision agriculture using machine learning algorithms. Journal of Natural Remedies 21: 22-29.

Kupfer JA, Gao P, Guo D (2012). Regionalization of forest pattern metrics for the continental United States using contiguity constrained clustering and partitioning. Ecological Informatics 9: 11-18. - doi: 10.1016/j.ecoinf.2012.02.001

Latapy M (2008). Main-memory triangle computations for very large (sparse(power-law)) graphs. Theoretical Computer Science 407: 458-473. - doi: 10.1016/j.tcs.2008.07.017

Latora V, Nicosia V, Russo G (2017). Complex networks: principles, methods and applications. Cambridge University Press, Cambridge, UK, pp. 1-28. - doi: 10.1017/9781316216002

Lawal O (2020). Spatially constrained clustering of Nigerian states: perspective from social, economic and demographic attributes. International Journal of Environment and Geoinformatics 7: 68-79. - doi: 10.30897/ijegeo.588032

Liu Y, Zhao W, Wang S, Fu B (2019). Landscape functional zoning at a country level based on ecosystem services bundle: Methods comparison and management indication. Journal of Environmental Management 249: 1-11. - doi: 10.1016/j.jenvman.2019.109315

Mahfuz NM, Yusoff M, Ahmad Z (2019). Review of single clustering methods. IAES International Journal of Artificial Intelligence 8: 221.

Marques M, Reynolds KM, Marques S, Marto M, Paplanus S, Borges JG (2021). A participatory and spatial multicriteria decision approach to prioritize the allocation of ecosystem services to management units. Land 10: 1-22. - doi: 10.3390/land10070747

Martínez-García V, González O, Ortiz-Pulido R (2020). Hummingbird-plan network in a lowland dry forest in Yucatan, Mexico. Tropical Conservation Science 13: 1-12. - doi: 10.1177/1940082920973830

Matese A (2020). Editorial for the special issue: forestry applications of Unmanned Aerial Vehicles (UAVs). Forests 11: 406. - doi: 10.3390/f11040406

Messier C, Bauhus J, Doyon F, Maure F, Sousa-Silva R, Nolet P, Mina M, Aquilué N, Fortin MJ, Puettmann K (2019). The functional complex network approach to foster forest resilience to global changes. Forest Ecosystems 6: 1-16. - doi: 10.1186/s40663-019-0166-2

Montes-Orozco E, Mora-Gutiérrez RA, Obregón-Quintana B, De los Cobos-Silva SG, Rincón-García EA, Lara-Velázquez P, Gutiérrez-Andrade MA (2020). Inverse percolation to quantify robustness in multiplex networks. Complexity ID 8796360: 1-11. - doi: 10.1155/2020/8796360

Müller A, Olschewski R, Unterberger C, Knoke T (2020). The valuation of forest ecosystem services as a tool for management planning- A choice experiment. Journal of Environmental Management 271: 1-13. - doi: 10.1016/j.jenvman.2020.111008

Myasnikov AA (2018). Zonal forest communities and forest zoning of western Siberia (Russia). Iraqi Journal of Agricultural Sciences 49: 938-943. - doi: 10.36103/ijas.v49i6.125

Opach T, Scherzer S, Lujala P, Ketil RJ (2020). Seeking commonalities of community resilience to natural hazards: a cluster analysis approach. Norwegian Journal of Geography 74: 181-199. - doi: 10.1080/00291951.2020.1753236

Peng J, Hu X, Qiu S, Hu Y, Meersmans J, Liu Y (2019). Multifunctional landscapes identification and associated development zoning in mountainous area. Science of the Total Environment 660: 765-775. - doi: 10.1016/j.scitotenv.2019.01.023

Rodrigues AR, Marques S, Botequim B, Marto M, Borges JG (2021). Forest management for optimizing soil protection: a landscape-level approach. Forest Ecosystems 8: 1-13. - doi: 10.1186/s40663-021-00324-w

Rubinov M, Sporns O (2010). Complex network measures of brain connectivity: uses and interpretations. Neuroimage 52: 1059-1069. - doi: 10.1016/j.neuroimage.2009.10.003

Srinivasaraghavan J, Allada V (2006). Application of Mahalanobis distance as a lean assessment metric. International Journal of Advanced Manufacturing Technology 29: 1159-1168. - doi: 10.1007/s00170-005-0004-2

The MathWorks Inc. (2020). MATLAB for Desktop version R2020b. The MathWorks Inc., Natick, MA, USA.

Velandia LA (2020). Aportes de los sistemas y redes complejas para la transformación social [Contributions of complex systems and networks for social transformation]. Revista Logos, Ciencia and Tecnología 12: 204-216. [in Spanish] - doi: 10.22335/rlct.v12i1.1066

Wang L, Zhou Y, Li Q, Zuo Q, Gao H, Liu J, Tian Y (2021). Forest land quality evaluation and the protection zoning of subtropical humid evergreen broadleaf forest region based on the PSOTOPSIS model and the local indicator of spatial association: a case study of Hefeng County, Hubei Province, China. Forests 12: 1-25. - doi: 10.3390/f12030325

Zanin M, Papo D, Sousa PA, Menasalvas E, Nicchi A, Kubike E, Boccaletti S (2016). Combining complex networks and data mining: why and how. Physics Reports 635: 1-44. - doi: 10.1016/j.physrep.2016.04.005

Zemp DC, Schleussner CF, Barbosa HM, Hirota M, Montade V, Sampaio G, Staal A, Wang-Erlandsson L, Rammig A (2017). Self-amplified Amazon forest loss due to vegetation-atmosphere feedbacks. Nature Communications 8: 1-10. - doi: 10.1038/ncomms14681

Zou W, Jing W, Chen G, Lu Y, Song H (2019). A survey of big data analytics for smart forestry. IEEE Access 7: 46621-46636. - doi: 10.1109/ACCESS.2019.2907999

## Supplementary Material

**Tab. S1** - Descriptive statistics for the segregation and connectivity metrics values obtained in the proposed complex networks R1-R5.

**Fig. S1** - (a) Best functional zoning scheme obtained with the clustering coefficient segregation metric and R5 complex network; and (b) current classification of the stands in the San Pedro El Alto community.

**Link:** Serrano-Ramirez_3927@suppl001.pdf