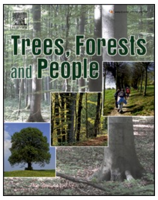
Trees, Forests and People 14 (2023) 100440

Contents lists available at ScienceDirect

Trees, Forests and People

journal homepage: www.sciencedirect.com/journal/trees-forests-and-people



Classifying the risk of forest loss in the Peruvian amazon rainforest: An alternative approach for sustainable forest management using artificial intelligence

Gianmarco Goycochea Casas [a,*], Juan Rodrigo Baselly-Villanueva [b], Mathaus Messias Coimbra Limeira [a], Carlos Moreira Miquelino Eleto Torres [a], Hélio Garcia Leite [a]

[a] *Department of Forest Engineering, Federal Univdaersity of Viçosa, Av. Purdue, s/n, Viçosa Campus, Zip Code 36570-900, Viçosa, MG, Brazil*
[b] *Estación Experimental Agraria San Roque, Dirección de Desarrollo Tecnológico Agrario, Instituto Nacional de Innovación Agraria (INIA), Calle San Roque 209, San Juan Bautista, Maynas, Loreto 16430, Peru*

ARTICLE INFO

Keywords:
Kohonen neural network
Forest conservation
Forest prevention
Risk classification

ABSTRACT

Peruvian Amazonian rainforests are constantly threatened by forest loss. Understanding changes in forest cover and assessing the level of risk is a permanent concern for numerous scientists and forest authorities. There are many conservation programs for Peruvian forests that involve collaborative efforts and employ diverse methodologies for forest monitoring. In this study, we propose an alternative approach to decision-making for forest preservation, aiming to classify the risk of forest loss in districts within the Peruvian Amazon rainforest. This classification enables sustainable forest management. To accomplish this, we utilized unsupervised learning artificial intelligence through Kohonen's neural network. The network was trained using a historical database spanning from 2001 to 2021, which includes variables such as forest cover and loss, climate, topography, hydrographic networks, and timber forest concessions. Through this approach, the network successfully established five clusters. Following preliminary analysis, we designated these clusters as: low, medium, high, very high, and extremely high risk of forest loss. Kohonen networks demonstrated their effectiveness in clustering forest loss and forest cover. The results indicate a shifting trend among the classes over time, with an increase in the categories exhibiting high and very high risk of forest cover loss. This study provides valuable information for decision-making in the prevention and conservation of Peruvian forests. We strongly recommend maintaining vigilance, particularly in districts classified as a very high or extremely high risk of losing forest cover.

1. Introduction

Peru has 56.51% of its surface area under forest cover, making it one of the ten countries with the highest forest density in the world and the second in South America (FRA, 2020). The Peruvian Amazon occupies the largest forest area (68,188,726 ha) followed by the Andean and dry forests (MINAM, 2016); therefore, it conserves a high biodiversity (Alvarez-Montalván et al., 2021), but ranks fifth worldwide in forest loss (Sierra Praeli, 2021). According to the GeoBosques platform of the Peruvian Ministry of Environment, 2,774,562 ha of land area was deforested between 2001 and 2021 alone (GeoBosques, 2021).

Primary forests in Peru are losing forest cover at an alarming rate each year. The highest rate of deforestation in the last 20 years occurred in 2020, when 203,272 ha of forest were lost. This was determined by satellite monitoring carried out by the Ministry of Environment through its National Forest Conservation Program for Climate Change Mitigation (PNCBMCC). The preservation of biodiversity, regulation of the hydrological cycle and carbon storage are one of the main sources of ecosystem services of the rainforest (Fearnside, 2008; Phillips and Brienen, 2017), services that are threatened due to an accelerated loss of forest cover (Brienen et al., 2015). The loss of forest cover in the Peruvian Amazon is caused by deforestation, which is mainly due to land use change, established by the search for lad for agriculture and cattle ranching by local people (Soares-Filho et al., 2006; Coomes et al., 2021; Lal, 2021).

The Peruvian government, in order to take strategic actions and

---

* Corresponding author.
E-mail address: gianmarco.casas@ufv.br (G.G. Casas).

https://doi.org/10.1016/j.tfp.2023.100440

implementation lines to reduce emissions from deforestation and forest degradation, approved the National Strategy on Forests and Climate Change (ENBCC) (Supreme Decree No. 007-2016-MINAM); which contains as a main action the monitoring of forests, which is being carried out progressively. In this context, governmental institutions have undertaken the task of classifying Peru's diverse forest regions. This classification process involves the utilization of advanced techniques such as remote sensing and geographic information systems. As a result of these efforts, six distinct ecozones have been identified: Coast, Highlands, High Accessible Forest, High Forest of Difficult Access, Low Forest, and Hydromorphic Zone. Remarkably, the Amazon rainforest contributes to four of these ecozones: High Accessible Forest, High Forest of Difficult Access, Low Forest, and Hydromorphic Zone. This stratification was guided by multifaceted criteria including floristic composition, physiographic characteristics, physiognomic attributes, carbon storage capacity, and accessibility factors (MINAM, 2014).

The Peruvian National Vegetation Cover Map was also used, conformed by special units defined and classified based on geographic, physiognomic, humidity condition and exceptionally floristic criteria (MINAM, 2015). The vegetation cover combines the criteria of vegetation formations with other physiographic, climatic, physiognomic, and anthropic criteria. It is stratified by six coverages: wetland forest, dryland forest, dryland scrub, wetland grassland, other vegetation formations and anthropic cover (SERFOR, 2016). The latter stands for the coverages in which there was human intervention, such as forest plantations, deforested areas, and agriculture.

A specific classification of the risk of losing forest cover in the Peruvian Amazon rainforest has not yet been studied. The classification of a forest is extremely complex due to the great biological interactions that exist between them, in addition to the physiographic and environmental conditions in which the areas are located (Moncrieff et al., 2016). The relationship between climate and plants occurs in multiple ways, influenced mainly by climate and geographic location (Zevallos and Lavado-Casimiro, 2022). To understand the numerous interactions in forest ecosystems, such as carbon deposition, climate buffering, and hydrological and erosional control, deeper investigations are essential (Ivanova et al., 2022). Simultaneously, the ability to collect extensive sets of forest-related data brings new perspectives, not only for the classification and monitoring of natural and altered habitats but also for their management (Gao et al., 2020; Xu et al., 2022; Zevallos and Lavado-Casimiro, 2022). However, the considerable amount of new data, characterized by their higher quality, structure, and analysis, poses a challenge (Ivanova et al., 2022).

Artificial neural networks (ANNs) can be used to classify forests in an alternative way. ANNs are a type of machine learning that can be used to solve a variety of problems, including regression, classification, and data compression (Ran and Hu, 2017). Among the various methods classified ANNs, an alternative classification approach is represented by self-organizing map (SOM) (Asan and Ercan, 2012).

The SOM was introduced by Kohonen (1982), Kohonen and Honkela (2007) in 1980s and it is known as Kohonen's neural networks or Kohonen map that can organize complex data into groups according to their relationships. The network is composed of an input layer and a Kohonen layer, in which the neurons of the input layer distribute the standard values for the Kohonen layer in a tabular form and keep the process updated during the whole training process (Kohonen, 1982; Dutra et al., 2021).

The SOM is an abstract mathematical model for mapping the topographical structure of visual sensors inspired by the structure of the cerebral cortex, which is organized in a two-dimensional grid of nodes, which are the individual neurons in the network. Each node represents a particular region of the input space, so that similar data points are mapped to neighboring nodes. This allows the SOM to represent the topological structure of the input data in a two-dimensional space (Yin, 2008). The SOM refers to an unsupervised neural network designed to condense input dimensions, thus portraying the distribution as a comprehensive map (Sacco et al., 2017), it has been used to develop new data compression and classification algorithms (Kohonen, 2001).

The learning algorithm employed by the SOM method can be classified into two distinct modalities: online learning and batch learning (Kikugawa et al., 2019). The application of batch learning in the evolution of the SOM has proven particularly relevant in practical applicability contexts (Kohonen, 2013). This learning process is based on the prior construction of a model based on historical data, which is subsequently used to evaluate the quality of recently obtained results (Ouidadi et al., 2023). Similarly, in operational scenarios where data is processed sequentially, the incremental training strategy emerges as the preferred approach (Mariño and Carvalho, 2022). The applicability of the online learning method involves continuous analysis of input data acquired during the monitoring process, promoting real-time updates of model parameters as new predictions are generated (Ouidadi et al., 2023).

The SOM has been applied in a broad sense that includes visualizations, feature map generation, pattern recognition and classification (Miljkovic, 2017). The SOM has helped to understand different behaviors of the biological sector through classification, such as classifying biological relevé and finding epigean species significantly associated with forest phytocenoses (Wolski and Kruk, 2020). The SOM also shows superior performance in environmental studies, such as the assessment of carbon stocks in forests, showing that it can reproduce the spatial pattern of carbon stocks (Stümer et al., 2010) and for the assessment of long-term changes in forest vegetation (Adamczyk et al., 2013). Likewise, the SOM is workable for the classification of land use from remote sensing data (Ji, 2000).

These neural networks could be used for classification problems where the output variable is difficult to obtain, such as classifying a forest area. The objective of this study was to classify the risk of loss of forest districts in the Peruvian Amazon rainforest using Kohonen neural networks, which allows and guides conservation and sustainable forest management actions.

2. Materials and methods

2.1. Localization and database

The study was conducted in the Amazon rainforest within the territory of Peru (Fig. 1). The Forest Cover Loss (Ha) and Forest Cover (Ha) database in tabular form used for classification were obtained from the Geobosques platform, which is managed by the Peruvian Ministry of Environment. This database comprised historical data from 2001 to 2021, encompassing 400 districts located in the Amazon rainforest and distributed across 15 departments. Additionally, climate, topography, hydrographic networks, and timber forest concessions data were incorporated to enhance the classification process, as these variables are closely associated with deforestation in the Peruvian Amazon (Bax and Francesconi, 2018; Cotrina Sánchez et al., 2021).

Climate data, including Corrected Precipitation (mm/day), Maximum Temperature at 2 M (°C), and Minimum Temperature at 2 M (°C), were acquired from the NASA Prediction of Worldwide Energy Resources, Topographic data, such as elevation (meters above sea level), and hydrographic networks information were sourced from the national chart generated by the Instituto Geográfico Nacional – IGN, while the timber forest concessions data was obtained from the Servicio Nacional Forestal y de Fauna Silvestre – SERFOR. Table 1 shows the source data used in the assessment of the risk of forest loss in the Peruvian Amazon rainforest. The variables used in the processing were rescaled to a spatial resolution of 250 m (Cotrina Sánchez et al., 2021).

2.2. Configuration of the Kohonen neural network

The data underwent processing within the computational environment of the R programming language R Core Team (2020), and this

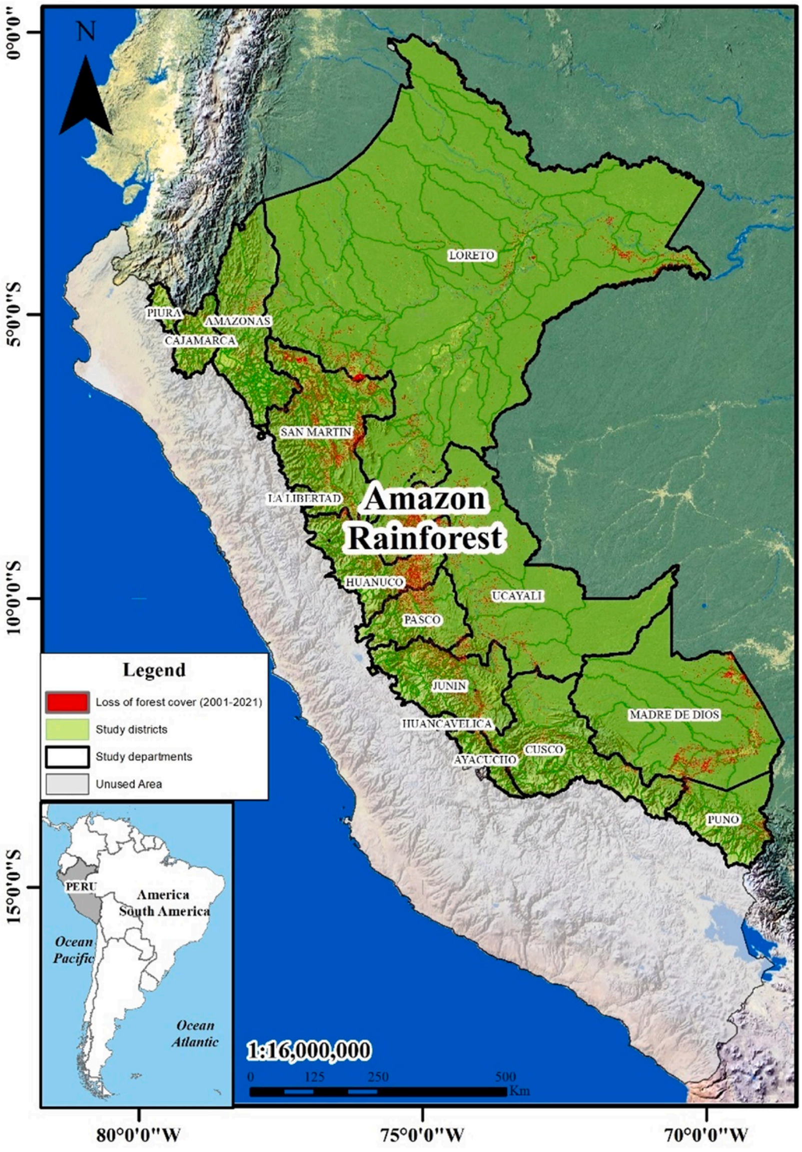


Fig. 1. Geographical distribution of districts in the Peruvian Amazon Rainforest biome.

procedure was facilitated by leveraging the functionalities provided by the aweSOM package (Boelaert et al., 2022). Specifically, the aweSOM package assisted in implementing a Self-Organizing Map (SOM) algorithm, a form of unsupervised machine learning widely utilized for pattern recognition and data visualization tasks.

As part of our methodology, we constructed a grid structure using the somgrid class, an object designed to systematically organize and record grid coordinates. To visualize and interpret the results of this grid arrangement, we employed the plot method (Venables and Ripley, 2002), which aided in the graphical representation of the grid’s layout and associated data points.

In the input layer, all possible combinations of the variables in our database were selected, establishing, and evaluating three training models. To facilitate effective learning and pattern extraction, we subjected the complete dataset to the SOM network on a total of 100 occasions. The learning rate, a crucial parameter influencing the rate of convergence during training, was dynamically configured ranging from an initial value of 0.05 and decreasing linearly to 0.01 in successive iterations.

In addition, the concept of neighborhood radius, representing the extent to which neighboring nodes influence each other’s updates, was operationalized as a vector. This vector spanned a spectrum from 2.65 to -2.65, thus modulating the extent of node interaction during training.

To ensure the compatibility of data structures and dimensions, the configuration of the array list was meticulously aligned with both the length of the data list and the total number of variables present within the dataset.

We harnessed two distinct learning algorithms: the batch (offline) and online. Notably, when employing the batch learning algorithm, the utilization of the learning rate was omitted, suggesting a different approach to weight updates and convergence dynamics compared to the online counterpart.

2.3. Training quality

The quality of training was evaluated according to the following criteria (Boelaert et al., 2022; Kaski and Lagus, 1996; Kohonen, 2001): (1) Quantization error (QE): It is defined as the mean squared distance between the data point and the map prototype to which it is assigned. The lower the distance, the better the quality of the model. (2) Percentage of explained variance (EV%): As with other clustering methods, the part of the total variance explained by clustering is one minus the

Table 1
Geospatial information related to the variables used in the assessment of the risk of forest loss in the Peruvian Amazon rainforest.

Data	Type	Source	Link	Spatial resolution
Cover Loss (Ha)	.xls	Geobosques	https://geobosques.minam.gob.pe/geobosque/view/descargas.php?122345gxxe345w34gg	-
Forest Cover (Ha)	.xls	Geobosques	https://geobosques.minam.gob.pe/geobosque/view/descargas.php?122345gxxe345w34gg	-
Corrected Precipitation (mm/day)	Raster	NASA	https://power.larc.nasa.gov/data-access-viewer/	10 kilometers
Maximum Temperature (°C)	Raster	NASA	https://power.larc.nasa.gov/data-access-viewer/	2 meters
Minimum Temperature (°C)	Raster	NASA	https://power.larc.nasa.gov/data-access-viewer/	2 meters
Topographic data (elevation)	Raster	Instituto Geográfico Nacional - IGN	https://www.geoidep.gob.pe/instituto-geografico-nacional	1:500 000
Hydrographic networks	Vetor	Instituto Geográfico Nacional - IGN	https://www.geoidep.gob.pe/instituto-geografico-nacional	1:500 000
Forest concessions	Vetor	Servicio Nacional Forestal y de Fauna Silvestre - SERFOR	https://geo.serfor.gob.pe/visor/	-
15 departments and 400 districts	Vetor	Ministerio del Ambiente - MINAM	https://geoservidor.minam.gob.pe/	-

Note: The data downloaded in .xls format from the Geobosques source are derived from a raster with a spatial resolution of 30 meters.

ratio of the quantification errors to the total variance. The higher the reference value, the better the quality of the model. (3) Topographic error (TE): Measures the degree of preservation of the topographic structure of the data on the map. It is calculated as the percentage of observations in which the best coincidence node is not adjacent to the second-best coincidence node of the map. It is taken as a reference that a low value presents better quality of the model. (4) Kaski-Lagus error (K-L): It is the sum of the average distance between the most compatible point and the best compatible prototype and the average geodesic distance between the most compatible point and the second best compatible prototype.

Similarly, the classification achieved by the Kohonen neural network was validated by comparing it with the deforested areas (observed data), which are reported in the Geobosques plantform in vector format. This validation process helped assess the accuracy and reliability of the network's classification results.

2.4. Establishment and assessment of the number of classes

To set up the number of classes, the scree graph (Cattell, 1966) was evaluated, which corresponds to the maximum value of the numerical solution of the approximation of the second derivative by finite differences, which can be defined as:

$$f''(i) \approx f(i+1) - 2f(i) + f(i-1)$$

Note here that the number of classes to be kept should be above the elbow of the graph, moreover, if the classes are important, the class should be steep while when the classes correspond to the error variance, the slope should be flat (Basto and Pereira, 2012).

2.5. Map of Kohonen

The Kohonen map was created using a cloud as a scatter plot within each cell. We set the same continuous variables shown in the model training and added district location as a categorical variable. The size of each cell was 400 pixels in that each cell is a visual representation of a specific unit in the self-organizing map. A color palette was used to stand for the classes as cell backgrounds and transparency was also set up to focus the variables. The graph was configured to show the name of the categorical variable used in this study, which were the districts belonging to the Amazon rainforest, from the observations in each cell and class. Subsequently, the Kohonen map classification results were exported to set up the district political map with their respective classifications.

3. Results

3.1. Training quality assessment

Table 2 provides a comprehensive overview of multiple models developed to classify the risk of forest cover loss in the Peruvian Amazon rainforest using the Kohonen neural network. Each model is characterized by the number of variables considered, the algorithms employed, and various performance metrics.

In the first model, which utilized three variables (loss of forest cover, forest cover, and climate), both the batch and online algorithms were applied. The models demonstrated quantization errors (QE) of 0.7135 and 0.7022, respectively. The explained variance (EV) percentages were 91.06% and 91.20%. The topographic errors (TE) were 0.125 and 0.113, while the Kaski-Lagus errors (K-L) were 2.8526 and 2.5488, respectively.

The second model expanded the analysis by including an additional variable, topography, alongside the variables used in the first model. The batch and online algorithms were again applied, resulting in quantization errors (QE) of 0.7504 and 0.7005, respectively. The explained variance (EV) percentages improved to 91.64% and 92.20%. The topographic errors (TE) were 0.125 and 0.158, while the Kaski-Lagus errors (K-L) showed values of 2.3575 and 2.7730.

The third model incorporated a comprehensive set of six variables, including loss of forest cover, forest cover, topography, climate, hydrographic networks, and timber forest concessions. The batch and online algorithms were once again utilized, resulting in quantization errors (QE) of 0.4540 and 0.4065, respectively. The explained variance (EV) percentages significantly increased to 94.32% and 94.92%, indicating the enhanced ability of the models to capture the underlying patterns in the data. The topographic error (TE) increased slightly to 0.261 and 0.291, while the Kaski-Lagus error (K-L) values were recorded as 1.9467 and 2.0938.

The inclusion of additional variables led to improved explained variance percentages, indicating a better understanding of the underlying factors contributing to classifying forest cover loss, as was the case with model three. It was also evident that the use of the online algorithm presented better training quality.

3.2. Analysis of the clusters

We have analyzed the clusters and cells of the best-trained model (Model 3 – Online Algorithm). According to the test and the scree plot, it was necessary to classify five clusters (Fig. 2A). Kohonen's map consisted of a 10 × 10 grid, which successfully classified the districts into

Table 2
Evaluation of training and clustering for each model and algorithm developed for forest cover loss risk assessment in the Peruvian Amazon rainforest.

Model Number	Variables	Algorithms	Quantization error (QE)	Explained variance % (EV)	Topographic error (TE)	Kaski-Lagus error (K-L)
1	Loss of forest cover, Forest cover, Climate	Batch	0.7135	91.06	0.125	2.8526
		Online	0.7022	91.20	0.113	2.5488
2	Loss of forest cover, Forest cover, Topography, Climate	Batch	0.7504	91.64	0.125	2.3575
		Online	0.7005	92.20	0.158	2.7730
3	Loss of forest cover, Forest cover, Topography, Climate, Hydrographic networks , Timber forest concessions	Batch	0.4540	94.32	0.261	1.9467
		Online	0.4065	94.92	0.291	2.0938

five distinct groups. Each cluster exhibited a varying number of cells, with 6, 11, 20, 31, and 32 cells, respectively (Fig. 2B).

The analysis involved calculating the average loss and forest cover for each year within each cluster. To facilitate comparison, the data needed to be normalized and analyzed separately. Based on the grouping of the data, a classification system was implemented, assigning ratings of low, medium, high, very high, and extremely high risk of forest cover loss to each cluster, agreeing with the classification used by the PNCBMCC (MINAM, 2019).

In this classification, a low rating indicated that the forest cover exceeded the loss, while an extremely high rating indicated that the loss of forest cover exceeded the available forested area (Fig. 3). By applying this rating system, the study was able to effectively assess the relative risk levels associated with forest cover loss within each cluster.

3.3. *Assessment of forest cover loss in Peruvian districts*

Forest loss and forest cover data were exported and categorized in vector form on the map of political districts, allowing visualization of the risk classification of each district. The procedure was performed for 4 years (2006, 2011, 2016, and 2021) to evaluate the succession and variation of the risk of cover loss for each district. Additionally, the deforested areas (observed data) were used, showing the effectiveness of the neural network classification (Fig. 4).

The results reveal a notable trend over time, with a shift in the classification of risk levels for forest cover loss. Specifically, there is a tendency towards an increase in the number of districts classified as high or very high risk, mainly in districts located in central and southern Peru.

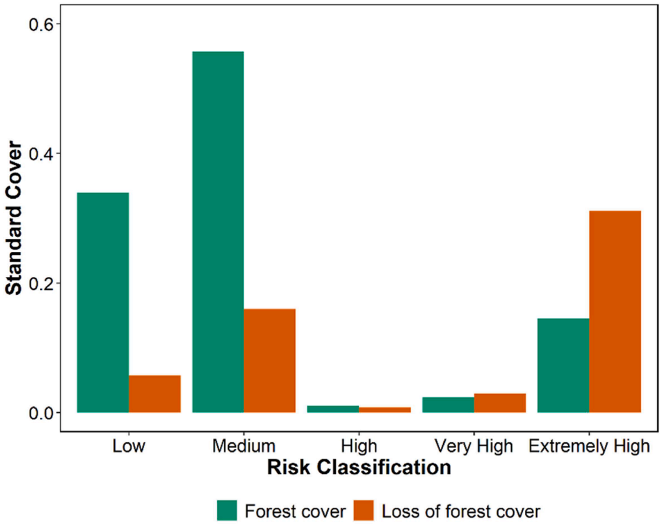


Fig. 3. Analysis of normalized average forest loss and forest cover data for the years 2001-2021 in the Peruvian Amazon rainforest.

In 2006, the districts of central Peru began to show an extremely high-risk classification of forest cover loss, as shown in Fig. 4A. However, it is important to highlight that during that year, the country still maintained considerable forest cover compared to the affected areas. The distribution of risk classifications for that year was: 53 districts classified as low risk, 9 districts as medium risk, 168 districts as high

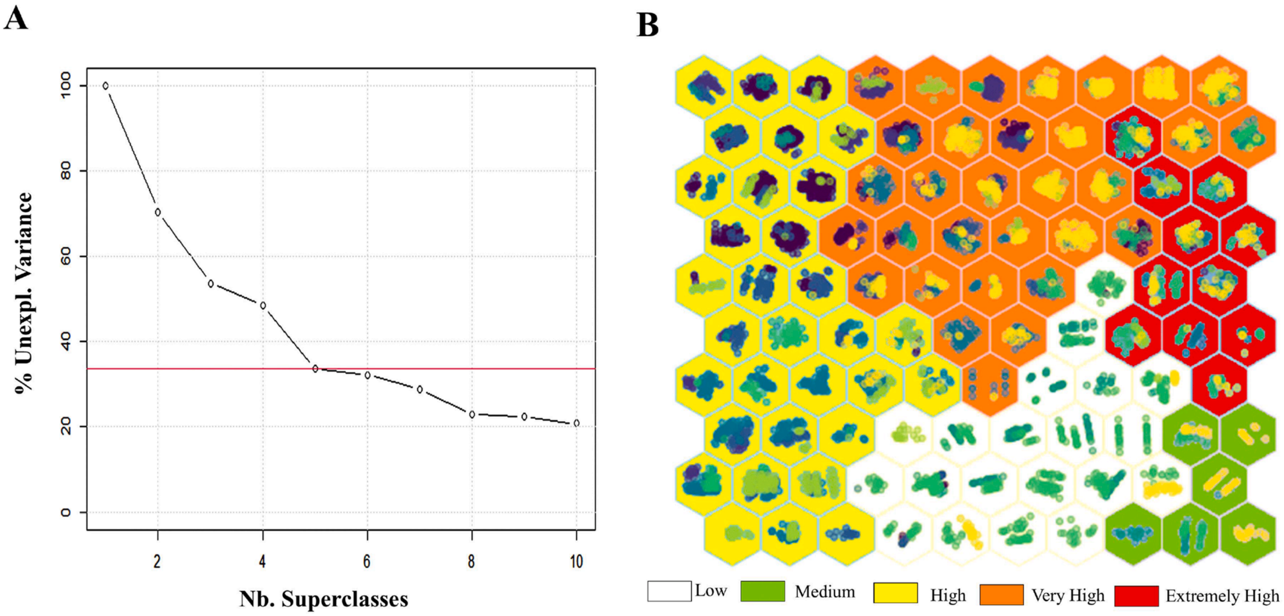


Fig. 2. (A) Determination of the number of classes using the scree plot, and (B) Kohonen map with their respective clusters. Note that the conglomerated points in each cell represent the districts as categorical variables.

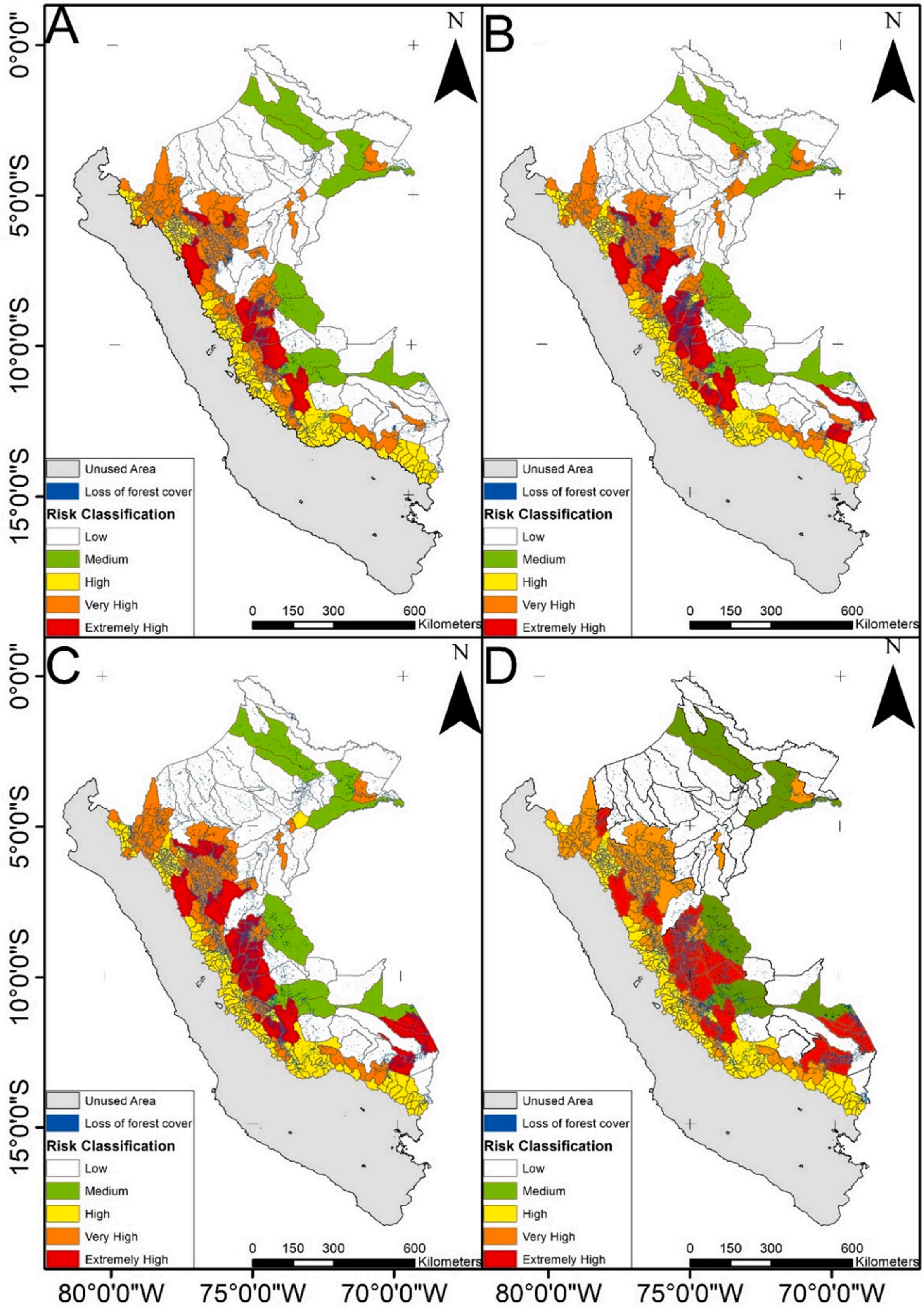


Fig. 4. District map of the risk of forest cover loss in the Peruvian Amazon rainforest in (A) 2006, (B) 2011, (C) 2016, and (D) 2021.
Note: The blue areas represent the observed data on forest cover loss for each year.

risk, 158 districts as very high risk, and 12 districts as extremely high risk.

In 2011, there was an expansion of districts categorized as extremely high-risk in central and southern Peru, as shown in Fig. 4B. The distribution of risk classifications for that year was: 46 low-risk districts, 9 medium-risk districts, 177 high-risk districts, 147 very high-risk districts, and 21 extremely high-risk districts.

In 2016, the focus remained on the central part of the country, with variations in risk classifications observed in the northern regions (Fig. 4C). The distribution of risk classifications for that year was: 47 low-risk districts, 9 medium-risk districts, 176 high-risk districts, 144 very high-risk districts, and 24 extremely high-risk districts.

In 2021, both the central and southern parts of the country experienced significant forest loss, resulting in higher risk classifications (Fig. 4D). The distribution of risk classifications for that year was: 45 districts classified as low risk, 9 districts as medium risk, 166 districts as high risk, 155 districts as very high risk, and 25 districts as extremely high risk.

Fig. 5A illustrates the dispersion of forest cover and forest cover loss in hectares for each risk classification in the years 2006, 2011, 2016, and 2021. Fig. 5B highlights the variation in risk classification, demonstrating an overall increase in the area classified as high or very high risk from 2001 to 2021. These graphs provide a visual representation of the changing landscape of risk levels over time, showing the increase in areas and districts with deforestation.

4. Discussion

The neural network of Kohonen's self-organizing map has been the attraction of several researchers in different areas due to its great classification potential (Ribeiro et al., 2014; Wolski and Kruk, 2020; Yu et al., 2019), including unsupervised learning, resulting in the ease of visualizing, and interpreting a classification with two or more variables with the linkage of categorical variables, which are sometimes not visible by conventional statistical methods (Moreira et al., 2019). Explaining how Kohonen networks work is complex, for reasons that are not self-explanatory (Stümer et al., 2010), in view of this, the evaluation of the quality of training is based on clustering methods. In this study, we evaluated different variables and their combination related to forest cover loss, keeping in mind that clustering is a complex task that

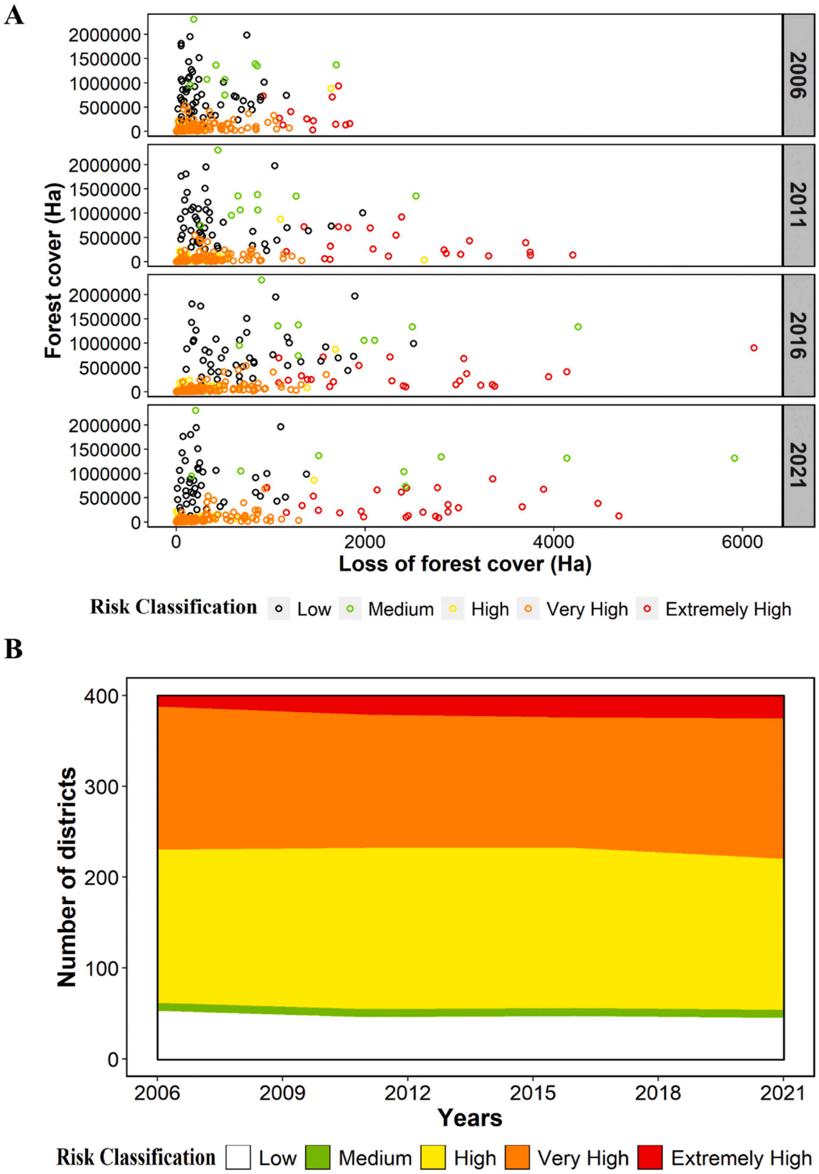


Fig. 5. (A) Scatter plot by risk classification between forest cover and forest cover loss per hectare and (B) Variation in risk classification; for the Peruvian Amazon rainforest districts.

requires careful consideration of many factors, in which we observed that the greater the number of variables related to each other, the better the quality of the clustering of the data (Table 2). Additionally, a high EV % was noted, indicating that the clustered data reflects the characteristics of the original dataset. This means that the technique was successfully preserving the essential information and structure of the data while discarding noise or redundant features (Boelaert et al., 2022).

The classification obtained in the present study agrees with the districts reported for forest cover loss by the Peruvian Forestry Authority (MINAM, 2019). However, in this study, we have evaluated different variables, understanding the relationship among all factors as the variables used, thus determining the risk of cover loss in specific districts. It is worth mentioning that this type of unsupervised learning helps to identify hidden patterns since the algorithm does not present guidance (Latif et al., 2019; Raza and Singh, 2021). Furthermore, unsupervised learning possesses the advantage of being unaffected by the presence of outliers (Lassoued and Abderrahim, 2013). This capacity stems from its integration of the missing value imputation (MVI) method, deeply rooted in machine learning methodologies (Hasan et al., 2021). Notably, the MVI method, specifically exemplified by the SOM, enhances the capacity to assign a singular data point to multiple clusters, a capability frequently harnessed within the realm of pattern recognition (Singh et al., 2015).

In the present study, the Kohonen Networks not only allowed us to classify, but also to visualize, analyze and understand the historical successions of the risk of loss of forest cover over the years (Fig. 5), even though it is difficult to process forest data due to its complexity and the existence of outliers.

According to Lal (2021), the deforestation rampant in the Peruvian Amazon can be traced back to evolving land use patterns, predominantly driven by agricultural expansion and livestock rearing, as well as the extensive engagement in gold mining operations (Caballero Espejo et al., 2018). The construction of the Interoceanic Highway, which cuts through the Madre de Dios region, is highly likely to have a detrimental impact on forest loss. This is because it facilitates the influx of people from various parts of Peru into the Amazonia (Asner and Tupayachi, 2016). This demographic impact gives rise to the fulfillment of subsistence requirements, including agricultural and mining pursuits, which, in turn, engenders an extensive array of adverse socio-environmental consequences (Alarcón et al., 2016; Moody et al., 2020; Swenson et al., 2011; Velásquez Ramírez et al., 2020).

By 2021, 25 districts were classified as extremely high-risk, predominantly those located in the central and southern zones, followed by those located in northern Peru. The causes of cover change or deforestation in the central zone are the easy transportation of timber to the city of Lima (Bax et al., 2016); likewise, in recent years, cover changes have been made for the installation of coca and oil palm plantations in the Departments of Ucayali and Huánuco (Bax and Francesconi, 2018; Vijay et al., 2018). The increase in the number of districts with an extremely high risk of cover loss in southern Peru is mainly due to illegal mining in the Department of Madre de Dios (Asner and Tupayachi, 2016; Cotrina Sánchez et al., 2021).

In general terms, the number of districts classified as being at extremely high risk of deforestation increased between 2001 and 2021. These districts were predominantly located in the center and south of Peru, followed by those in the north. These findings are consistent with the study reported by Vicencio and de Vivanco (2023), who found that the concentration of extremely high forest loss is mainly located in the center and south of Peru.

There are many forest conservation programs that have mutual efforts and use different methodologies for forest monitoring (Cappello et al., 2022). It is most important to focus on the districts classified in this study as having an extremely high risk of losing cover and also on those districts that are about to move from a very high to an extremely high-risk class, as shown in Fig. 5. While the high risk of losing forest cover may persist in the long term, our classification study can be an alternative to prevent forest loss. This study provides valuable insights for forest management in Peru, offering essential information to guide decision-making and conservation efforts in the region.

Considering that the removal of vegetation cover and the felling of forest resources that do not have the corresponding authorization is a very serious offense according to the Forestry Peruvian Law (Supreme Decree No. 007-2021-MIDAGRI). In addition to intensifying management and conservation practices starting today, the world can reduce one of the main causes of climate change, a phenomenon that has already been modifying some ecosystem services (Reygadas et al., 2023).

Forests are complex ecosystems that interact dynamically. Many factors can influence forest loss, but some of them were not included in the study because of the lack of monitored data since the time of the study. These factors include human footprint, roads, socioeconomic, and ecological factors. It is important to continuously monitor these factors. This is especially important for modeling, as accurate modeling requires accurate data. Data analysis allows us to better understand the complex dynamics of forest loss. This knowledge can be used to make informed decisions about sustainable forest management and conservation.

5. Concussions

The utilization of Kohonen networks in this study has demonstrated their effectiveness in clustering forest loss and cover, enabling a quantitative assessment, and facilitating the establishment of a risk classification for forest cover loss in the districts of the Peruvian Amazon rainforest. This methodology also allows for the inclusion of various other variables that can influence forest cover, such as ecological, biological, and geological information, as well as changes in cover and the interaction with other biomes. These aspects should be further investigated in future studies.

It has been observed that out of the 12 districts initially classified as extremely high risk, 25 districts have now been identified as extremely high risk of losing their forest cover in the year 2021, indicating an alarming trend. This study provides crucial information to support decision-making for forest preservation and conservation in Peru. It offers technical expertise that can be used to develop Supreme Decrees or Ministerial Resolutions that prioritize monitoring and surveillance efforts in districts with a high or extremely high risk of forest cover loss.

CRediT authorship contribution statement

Gianmarco Goycochea Casas: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Juan Rodrigo Baselly-Villanueva:** Conceptualization, Methodology, Data curation, Formal analysis, Writing – review & editing. **Mathaus Messias Coimbra Limeira:** Formal analysis, Data curation. **Carlos Moreira Miquelino Eleto Torres:** Methodology, Supervision. **Hélio Garcia Leite:** Methodology, Conceptualization, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data are enabled online and the sources are written in the manuscript.

Acknowledgments

The present work was carried out with support from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Financing Code 001.

References

Adamczyk, J.J., Kurzac, M., Park, Y.S., Kruk, A., 2013. Application of a Kohonen's self-organizing map for evaluation of long-term changes in forest vegetation. J. Veg. Sci. 24, 405–414. https://doi.org/10.1111/j.1654-1103.2012.01468.x.

Alarcón, G., Díaz, J., Vela, M., García, M., Gutiérrez, J., 2016. Deforestación en el sureste de la amazonia del Perú entre los años 1999 - 2013; caso Regional de Madre de Dios (Puerto Maldonado – Inambari). Rev. Investig. Altoandin. J. High Andean Res. 18 https://doi.org/10.18271/ria.2016.221.

Alvarez-Montalván, C.E., Manrique-León, S., Fonseca, V.D., Cardozo-Soarez, J., Callo-Ccorcca, J., Bravo-Camara, P., Castañeda-Tinco, I., Alvarez-Orellana, J., 2021. Floristic composition, structure and tree diversity of an amazon forest in Peru. Sci. Agropecu. 12, 73–82. https://doi.org/10.17268/sci.agropecu.2021.009.

Asan, U., Ercan, S., 2012. An introduction to self-organizing maps. In: Kahraman, C. (eds) Computational Intelligence Systems in Industrial Engineering. Atlantis Computational Intelligence Systems 6, 295–31. 10.2991/978-94-91216-77-0_14.

Asner, G.P., Tupayachi, R., 2016. Accelerated losses of protected forests from gold mining in the Peruvian Amazon. Environ. Res. Lett. 12, 094004 https://doi.org/10.1088/1748-9326/aa7dab.

Basto, M., Pereira, J.M., 2012. An SPSS R -menu for ordinal factor analysis. J. Stat. Softw. 46 https://doi.org/10.18637/jss.v046.i04.

Bax, V., Francesconi, W., 2018. Environmental predictors of forest change: an analysis of natural predisposition to deforestation in the tropical Andes region, Peru. Appl. Geogr. 91, 99–110. https://doi.org/10.1016/j.apgeog.2018.01.002.

Bax, V., Francesconi, W., Quintero, M., 2016. Spatial modeling of deforestation processes in the Central Peruvian Amazon. J. Nat. Conserv. 29, 79–88. https://doi.org/10.1016/j.jnc.2015.12.002.

Boelaert, J., Ollion, E., Sodoge, J., Megdoud, M., Naji, O., Kote, A.L., Renoud, T., Hym, S., 2022. Interactive self-organizing maps.

Brienen, R.J.W., Phillips, O.L., Feldpausch, T.R., Gloor, E., Baker, T.R., Lloyd, J., Lopez-Gonzalez, G., Monteagudo-Mendoza, A., Malhi, Y., Lewis, S.L., Vásquez Martinez, R., Alexiades, M., Álvarez Dávila, E., Alvarez-Loayza, P., Andrade, A., Aragão, L.E.O.C., Araujo-Murakami, A., Arets, E.J.M.M., Arroyo, L., Aymard, C.G.A., Bánki, O.S., Baraloto, C., Barroso, J., Bonal, D., Boot, R.G.A., Camargo, J.L.C., Castilho, C.V., Chama, V., Chao, K.J., Chave, J., Comiskey, J.A., Cornejo Valverde, F., da Costa, L., de Oliveira, E.A., Di Fiore, A., Erwin, T.L., Fauset, S., Forsthofer, M., Galbraith, D.R., Grahame, E.S., Groot, N., Hérault, B., Higuchi, N., Honorio Coronado, E.N., Keeling, H., Killeen, T.J., Laurance, W.F., Laurance, S., Licona, J., Magnussen, W.E., Marimon, B.S., Marimon-Junior, B.H., Mendoza, C., Neill, D.A., Nogueira, E.M., Núñez, P., Pallqui Camacho, N.C., Parada, A., Pardo-Molina, G., Peacock, J., Peña-Claros, M., Pickavance, G.C., Pitman, N.C.A., Poorter, L., Prieto, A., Quesada, C.A., Ramírez, F., Ramírez-Angulo, H., Restrepo, Z., Roopsind, A., Rudas, A., Salomão, R. P., Schwarz, M., Silva, N., Silva-Espejo, J.E., Silveira, M., Stropp, J., Talbot, J., ter Steege, H., Teran-Aguilar, J., Terborgh, J., Thomas-Caesar, R., Toledo, M., Torello-Raventos, M., Umetsu, R.K., van der Heijden, G.M.F., van der Hout, P., Guimarães Vieira, I.C., Vieira, S.A., Vilanova, E., Vos, V.A., Zagt, R.J., 2015. Long-term decline of the Amazon carbon sink. Nature 519, 344–348. https://doi.org/10.1038/nature14283.

Caballero Espejo, J., Messinger, M., Román-Dañobeytia, F., Ascorra, C., Fernandez, L., Silman, M., 2018. Deforestation and forest degradation due to gold mining in the peruvian amazon: a 34-year perspective. Remote Sens. 10, 1903. https://doi.org/10.3390/rs10121903 (Basel).

Cappello, C., Pratihast, A.K., Pérez Ojeda del Arco, A., Reiche, J., de Sy, V., Herold, M., Vivanco Vicencio, R.E., Castillo Soto, D., 2022. Alert-driven community-based forest monitoring: a case of the peruvian amazon. Remote Sens. 14, 4284. https://doi.org/10.3390/rs14174284 (Basel).

Cattell, R.B., 1966. The scree test for the number of factors. Multivar. Behav. Res. 1, 245–276. https://doi.org/10.1207/s15327906mbr0102_10.

Coomes, O.T., Cheng, Y., Takasaki, Y., Abizaid, C., 2021. What drives clearing of old-growth forest over secondary forests in tropical shifting cultivation systems? Evidence from the Peruvian Amazon. Ecol. Econ. 189, 107170 https://doi.org/10.1016/j.ecolecon.2021.107170.

Cotrina Sánchez, A., Bandopadhyay, S., Rojas Briceño, N.B., Banerjee, P., Torres Guzmán, C., Oliva, M., 2021. Peruvian Amazon disappearing: Transformation of protected areas during the last two decades (2001–2019) and potential future deforestation modelling using cloud computing and MaxEnt approach. J. Nat. Conserv. 64, 126081 https://doi.org/10.1016/j.jnc.2021.126081.

Dutra, R.A., Reis, E.L., Reis, C., Fidêncio, P.H., Reis, C.D.G., de Carvalho Damasceno, O. I., 2021. Spatial and temporal analysis of physical and chemical data of superficial waters, by self-organizing maps (SOM). Braz. J. Dev. 7, 57578–57594. https://doi.org/10.34117/bjdv7n6-251.

Fearnside, P.M., 2008. Amazon Forest maintenance as a source of environmental services. Acad. Bras. Cienc. 80, 101–114. https://doi.org/10.1590/S0001-37652008000100006.

FRA, 2020. Global forest resources assessment 2020: main report. https://fra-data.fao.org/PER/fra2020/home/ (accessed 11.13.22).

Gao, Y., Skutsch, M., Paneque-Gálvez, J., Ghilardi, A., 2020. Remote sensing of forest degradation: a review. Environ. Res. Lett. 15, 103001 https://doi.org/10.1088/1748-9326/abaad7.

GeoBosques, 2021. Cobertura y pérdida de bosque húmedo amazónico. https://geobosques.minam.gob.pe/geobosque/descargas_geobosque/perdida/documentos/Reporte_Cobertura_y_Perdida_de_Bosque_Humedo_Amazonico_2021.pdf?Tue%20Feb%2021%202023%2012:05:58%20GMT-0500%20(hora%20est%C3%A1ndar%20de%20Per%C3%BA) (accessed 6.11.23).

Hasan, M.K., Alam, M.A., Roy, S., Dutta, A., Jawad, M.T., Das, S., 2021. Missing value imputation affects the performance of machine learning: A review and analysis of the literature (2010–2021). Inform. Med. Unlocked 27, 100799. https://doi.org/10.1016/j.imu.2021.100799.

Ivanova, N., Fomin, V., Kusbach, A., 2022. Experience of forest ecological classification in assessment of vegetation dynamics. Sustainability 14, 3384. https://doi.org/10.3390/su14063384.

Ji, C.Y., 2000. Land-use classification of remotely sensed data using Kohonen self-organizing feature map neural networks. Photogramm. Eng. Remote Sens. 66, 1451–1460.

Kaski, S., Lagus, K., 1996. Comparing self-organizing maps. pp. 809–814. 10.1007/3-540-61510-5_136.

Kikugawa, G., Nishimura, Y., Shimoyama, K., Ohara, T., Okabe, T., Ohuchi, F.S., 2019. Data analysis of multi-dimensional thermophysical properties of liquid substances based on clustering approach of machine learning. Chem. Phys. Lett. 728, 109–114. https://doi.org/10.1016/j.cplett.2019.04.075.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern. 43, 59–69. https://doi.org/10.1007/BF00337288.

Kohonen, T., 2001. Self-Organizing Maps. Springer Berlin Heidelberg, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-56927-2.

Kohonen, T., 2013. Essentials of the self-organizing map. Neural Netw.

Kohonen, T., Honkela, T., 2007. Kohonen network. Scholarpedia 2, 1568. https://doi.org/10.4249/scholarpedia.1568.

Lal, R., 2021. Soil Organic Carbon and Feeding the Future, 1st ed. CRC Press, Boca Raton. https://doi.org/10.1201/9781003243090. ed.

Lassoued, F.A.Z., Abderrahim, S.B.K., 2013. A Kohonen neural network based method for PWARX identification. IFAC Proc. Vol. 46, 742–747. https://doi.org/10.3182/20130703-3-FR-4038.00088.

Latif, J., Xiao, C., Imran, A., Tu, S., 2019. Medical imaging using machine learning and deep learning algorithms: a review. In: Proceedings of the 2nd International Conference on Computing, Mathematics and Engineering Technologies (ICoMET). IEEE, pp. 1–5. https://doi.org/10.1109/ICOMET.2019.8673502.

Mariño, L.M., de Carvalho, F.D.A., 2022. Vector batch SOM algorithms for multi-view dissimilarity data. Knowl. Based Syst. 258, 109994.

Miljkovic, D., 2017. Brief review of self-organizing maps. In: Proceedings of the 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO). IEEE, pp. 1061–1066. https://doi.org/10.23919/MIPRO.2017.7973581.

MINAM, 2014. Estimación de los contenidos de carbono de la biomasa aérea en los bosques de Perú.

MINAM, 2015. Mapa nacional de cobertura vegetal: memoria descriptiva /Ministerio del Ambiente, Dirección General de Evaluación. Valoración y Financiamiento del Patrimonio Natural, Lima, Perú.

MINAM, 2016. Programa nacional de conservación de bosques para la mitigación del cambio climático (accessed 6.11.23).

MINAM, 2019. Apuntes del Bosque N^{o} 1: Cobertura y Deforestación en Los Bosques Húmedos Amazónicos 2018. Programa Nacional de Conservación de Bosque para la Mitigación del Cambio Climático. https://sinia.minam.gob.pe/documentos/apuntes-bosque-no-1-cobertura-deforestacion-bosques-humedos\.

Moncrieff, G.R., Bond, W.J., Higgins, S.I., 2016. Revising the biome concept for understanding and predicting global change impacts. J. Biogeogr. 43, 863–873. https://doi.org/10.1111/jbi.12701.

Moody, K.H., Hasan, K.M., Aljic, S., Blakeman, V.M., Hicks, L.P., Loving, D.C., Moore, M. E., Hammett, B.S., Silva-González, M., Seney, C.S., Kiefer, A.M., 2020. Mercury emissions from Peruvian gold shops: potential ramifications for Minamata compliance in artisanal and small-scale gold mining communities. Environ. Res. 182, 109042 https://doi.org/10.1016/j.envres.2019.109042.

Moreira, L.S., Chagas, B.C., Pacheco, C.S.V., Santos, H.M., de Menezes, L.H.S., Nascimento, M.M., Batista, M.A.S., de Jesus, R.M., Amorim, F.A.C., Santos, L.N., da Silva, E.G.P., 2019. Development of procedure for sample preparation of cashew nuts using mixture design and evaluation of nutrient profiles by Kohonen neural network. Food Chem. 273, 136–143. https://doi.org/10.1016/j.foodchem.2018.01.050.

Ouidadi, H., Guo, S., Zamiela, C., Bian, L., 2023. Real-time defect detection using online learning for laser metal deposition. J. Manuf. Process. 99, 898–910.

Phillips, O.L., Brienen, R.J.W., 2017. Carbon uptake by mature Amazon forests has mitigated Amazon nations' carbon emissions. Carbon Balanc. Manag. 12 (1) https://doi.org/10.1186/s13021-016-0069-2.

R Core Team, 2020. R: a language and environment for statistical computing (en línea).

Ran, Z., Hu, B., 2017. Parameter identifiability in statistical machine learning: a review. Neural Comput. 29, 1151–1203. https://doi.org/10.1162/NECO_a_00947.

Raza, K., Singh, N.K., 2021. A tour of unsupervised deep learning for medical image analysis. Curr. Med. Imaging Former. Curr. Med. Imaging Rev. 17, 1059–1077. https://doi.org/10.2174/1573405617666210127154257.

Reygadas, Y., Spera, S.A., Salisbury, D.S., 2023. Effects of deforestation and forest degradation on ecosystem service indicators across the Southwestern Amazon. Ecol. Indic. 147, 109996 https://doi.org/10.1016/j.ecolind.2023.109996.

Ribeiro, F.A.L., Rosário, F.F., Bezerra, M.C.M., Wagner, R.de C.C., Bastos, A.L.M., Melo, V.L.A., Poppi, R.J., 2014. Evaluation of chemical composition of waters associated with petroleum production using Kohonen neural networks. Fuel 117, 381–390. https://doi.org/10.1016/j.fuel.2013.08.086.

Sacco, D., Motta, G., You, L., Bertolazzo, N., Carini, F., Ma, T., 2017. Smart cities, urban sensing, and big data: mining geo-location in social networks. In Big Data and Smart Service Systems, 59–84). 10.1016/B978-0-12-812013-2.00005-8.

SERFOR, 2016. Memoria descriptiva del mapa de ecozonas. Inventario nacional forestal y de fauna silvestre (INFFS) - Perú.

Sierra Praeli, Y., 2021. Perú alcanza cifra de deforestación más alta en los últimos 20 años. https://es.mongabay.com/2021/10/peru-aumenta-deforestacion-cifras-bosques/ (accessed 10.13.22).

Singh, N., Javeed, A., Chhabra, S., Kumar, P., 2015. Missing Value Imputation with Unsupervised Kohonen Self Organizing Map, in: Emerging Research in Computing, Information, Communication and Applications. Springer India, New Delhi, pp. 61–76. https://doi.org/10.1007/978-81-322-2550-8_7.

Soares-Filho, B., Nepstad, D., Curran, L., Cerqueira, G.C., Garcia, R.A., Ramos, C.A., Voll, E., McDonald, A., Lefebvre, P., Schlesinger, P., 2006. Modelling conservation in the Amazon basin. Nature 440, 520–523. https://doi.org/10.1038/nature04389.

Stümer, W., Kenter, B., Köhl, M., 2010. Spatial interpolation of *in situ* data by self-organizing map algorithms (neural networks) for the assessment of carbon stocks in European forests. For. Ecol. Manag. 260, 287–293. https://doi.org/10.1016/j.foreco.2010.04.008.

Swenson, J.J., Carter, C.E., Domec, J.C., Delgado, C.I., 2011. Gold mining in the Peruvian amazon: global prices, deforestation, and mercury imports. PLoS One 6, e18875. https://doi.org/10.1371/journal.pone.0018875.

Velásquez Ramírez, M.G., Barrantes, J.A.G., Thomas, E., Gamarra Miranda, L.A., Pillaca, M., Tello Peramas, L.D., Bazán Tapia, L.R., 2020. Heavy metals in alluvial gold mine spoils in the peruvian amazon. Catena 189, 104454. https://doi.org/10.1016/j.catena.2020.104454 (Amst).

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S. Springer New York, New York, NY. https://doi.org/10.1007/978-0-387-21706-2.

Vicencio, R.V., de Vivanco, D.D.R., 2023. Análisis geográfico de la concentración de la pérdida de bosques húmedos amazónicos del Perú. High Tech-Eng. J. 3, 2–20.

Vijay, V., Reid, C.D., Finer, M., Jenkins, C.N., Pimm, S.L., 2018. Deforestation risks posed by oil palm expansion in the Peruvian Amazon. Environ. Res. Lett. 13, 114010 https://doi.org/10.1088/1748-9326/aae540.

Wolski, G.J., Kruk, A., 2020. Determination of plant communities based on bryophytes: the combined use of Kohonen artificial neural network and indicator species analysis. Ecol. Indic. 113, 106160 https://doi.org/10.1016/j.ecolind.2020.106160.

Xu, C., Zhang, X., Hernandez-Clemente, R., Lu, W., Manzanedo, R.D., 2022. Global forest types based on climatic and vegetation data. Sustainability 14, 634. https://doi.org/10.3390/su14020634.

Yin, H., 2008. The self-organizing maps: background, theories, extensions and applications. pp. 715–762. 10.1007/978-3-540-78293-3_17.

Yu, X., Xiao, F., Zhou, Y., Wang, Y., Wang, K., 2019. Application of hierarchical clustering, singularity mapping, and Kohonen neural network to identify Ag-Au-Pb-Zn polymetallic mineralization associated geochemical anomaly in Pangxidong district. J. Geochem. Explor. 203, 87–95. https://doi.org/10.1016/j.gexplo.2019.04.007.

Zevallos, J., Lavado-Casimiro, W., 2022. Climate change impact on Peruvian biomes. Forests 13, 238. https://doi.org/10.3390/f13020238.