Botanical Sciences 103 (2): 381-393. 2025
DOI: 10.17129/botsci.3598

Received: September 20, 2024, Accepted: January 31, 2025
On line first: March 15, 2025

Ecology / Ecología

A proposal of an updated potential vegetation map of Mexico

Enrique Ortiz, José Luis Villaseñor*

Departamento de Botánica, Instituto de Biología, Universidad Nacional Autónoma de México, Ciudad de México, Mexico.
*Corresponding author: vrios@ib.unam.mx

Abstract

Background: Jerzy Rzedowski's 1978 publication of Mexico's potential vegetation map has been widely used. However, Rzedowski acknowledged his map's weaknesses regarding its resolution and precision.

Question: Can the resolution and accuracy of Rzedowski's potential vegetation map be improved?

Data description: The vegetation types of Mexico.

Study site: Mexico.

Methods: Using Series VII of INEGI's Land Use and Vegetation map of Mexico as a base, the interpolation technique known as 'Thiessen polygons' was applied to assign one of the nine Rzedowski's vegetation types to each polygon. The new map thus created was validated against vegetation data recorded for 10,000 Asteraceae collection sites, by calculating Cohen's Kappa and other performance metrics, such as sensitivity, specificity, errors of omission, and commission. The same validation was used to evaluate Rzedowski's original potential vegetation map.

Results: The potential vegetation map produced using Thiessen polygons showed statistically better agreement with the Asteraceae validation points than Rzedowski's map.

Conclusions: The interpolation method was useful for generating a potential vegetation map whose resolution and precision are suitable for spatial analysis. The proposal presented here is a heuristic exercise that can serve as a hypothesis to support or falsify future vegetation analyses at different scales throughout the country.

Keywords: Cohen's Kappa, digital mapping, precision, Rzedowski, spatial resolution, Thiessen polygons.

Resumen

Antecedentes: El mapa de vegetación potencial de México, publicado en 1978 por Jerzy Rzedowski, ha sido ampliamente utilizado. Sin embargo, el mismo autor señala algunas debilidades en su mapa acerca de su resolución y precisión.

Pregunta: ¿Es posible mejorar la resolución y precisión del mapa de vegetación potencial de Rzedowski?

Descripción de datos: Tipos de vegetación de México.

Sitio de estudio: México.

Métodos: Empleando como base el mapa de Uso de Suelo y Vegetación, serie VII de INEGI, se aplicó la técnica de interpolación conocida como 'polígonos de Thiessen' para asignar a sus polígonos alguno de los nueve tipos de vegetación de Rzedowski. El nuevo mapa se validó con los datos de vegetación reportados para 10,000 sitios de recolecta de Asteraceae, calculando la Kappa de Cohen y otras métricas de desempeño, como sensibilidad, especificidad, errores de omisión y comisión. La misma validación se hizo para evaluar el mapa original de vegetación potencial de Rzedowski.

Resultados: El mapa de vegetación potencial producido con los polígonos de Thiessen estadísticamente mostró mejor correspondencia con respecto a los puntos de validación de Asteraceae, en comparación con el mapa de Rzedowski.

Conclusiones: El método de interpolación usado fue útil para generar un mapa de vegetación potencial, cuya resolución y precisión son adecuadas para realizar análisis espaciales. La propuesta aquí presentada constituye un ejercicio heurístico que puede servir como hipótesis para apoyar o refutar futuros análisis de vegetación a distintas escalas en el país.

Palabras clave: Cartografía digital, Kappa de Cohen, polígonos de Thiessen, precisión, resolución espacial, Rzedowski.

Vegetation maps are visual inventories of the types, geographic distribution, and extent of plant communities in a given area (Küchler 1984, Küchler & Zonneveld 1988). Vegetation maps provide crucial information for studies in a wide variety of disciplines, including ecology, taxonomy, biodiversity conservation, vegetation management, and ecological restoration, and they are an important tool for strategic decision-making at a national level (Küchler 1953, Paz-Pellat *et al.* 2020). They are also scientific tools that allow us to analyze the relationships between plants and the environment in which they occur, thus providing a valuable reference for observing and measuring temporary changes in vegetation (Küchler & Zonneveld 1988). However, the proper quantification of vegetation cover loss requires maps of potential (natural) vegetation, defined as maps that describe the geographic distribution of vegetation in the hypothetical absence of human intervention (Zerbe 1998, Chiarucci *et al.* 2010). Potential vegetation maps are just as valuable as maps of current vegetation, as they make it possible to estimate where or how much the vegetation cover has been transformed due to changes in land use or other events such as environmental catastrophes, in addition to being a useful tool in landscape management or ecological restoration (Carranza *et al.* 2003, Loidi *et al.* 2010).

There have been several efforts to represent Mexico's vegetation. González-Medrano (2004) summarized them; later, Meave *et al.* (2016) reviewed the processes and challenges of classifying the country's vegetation. We refer to those works rather than repeating here the information therein. In the 20 years since the publication of the treatment of González-Medrano (2004) on this topic, however, little work has been done aimed at understanding vegetation and documenting the results graphically. Examples of this work are the studies on the vegetation of Durango State (González-Elizondo *et al.* 2007), the southeastern tropical region of Mexico (Vaca *et al.* 2011), the Sierra Madre Occidental (González-Elizondo *et al.* 2012), and of Aguascalientes State (Siqueiros-Delgado *et al.* 2016). The scant research focused on vegetation mapping is reflected in the dearth of studies on the subject reported in a global analysis of vegetation studies (Ibarra-Manríquez *et al.* 2022), which reported that among 15,000 articles on vegetation analyzed, fewer than 5 % focused on vegetation mapping.

Rzedowski's potential vegetation map of Mexico was initially published in 1978 in his book "Vegetación de México" and is an important source of information on the distribution of vegetation in the country (Greller *et al.* 2000, Olson *et al.* 2001, Velázquez *et al.* 2016). It is widely used, since, for each vegetation type, the book describes, among other pieces of information, the species composition, aspects of its structure, and the environmental factors that define its distribution. The map lacks a spatial scale, although based on the included graphic scale, it is probably about 1:7,700,000. The map was later published as part of the National Atlas of Mexico (Rzedowski 1990) at a 1:4,000,000 scale, and its shapefile was published by CONABIO (National Commission for the Knowledge and Use of Biodiversity) in 2001 and is available on its Geoinformation Portal (http://www.conabio.gob.mx/informacion/gis/). In the shapefile, Rzedowski differentiated nine vegetation types and included two additional polygons as water bodies—the La Amistad and Falcón dams, on the border with the United States of America (Figure 1A). Based on the digital map, Table 1 shows the surface area occupied by each vegetation type.

From 1997 to 2021, the National Institute of Statistics and Geography (INEGI, https://www.inegi.org.mx/temas/usosuelo/) published seven series of vegetation and land use maps of Mexico; all can be accessed as shapefiles. These maps show the geographic distribution of natural and induced vegetation, areas without apparent vegetation, and agricultural or urban land use caused by anthropization. Series VII (INEGI 2021), for example, includes 86,714 polygons equivalent to some of Rzedowski's (1978) vegetation types, representing 69.7 % of the country's surface area; by contrast, 92,381 polygons, which cover 30.3 % of Mexico's geographical extent, appear to lack natural vegetation cover or cannot be directly assigned to any of Rzedowski's vegetation types.

Rzedowski (1978) acknowledged the limitations of his vegetation map, indicating that it was illustrated in a schematic, rather than precise way, given the apparent complexity of their cartographic representation and the lack of knowledge of the limits between plant communities. He also disclosed the possible incorrect assignment to a given vegetation type in some areas. Finally, he stated that the scale at which the map was published (1:7,700,000 according to our interpretation) prevented the mapping of many small patches of various plant communities, which led to inaccurate generalizations. Later, Vaca *et al.* (2011), in addition to discussing the weaknesses of Rzedowski's map,

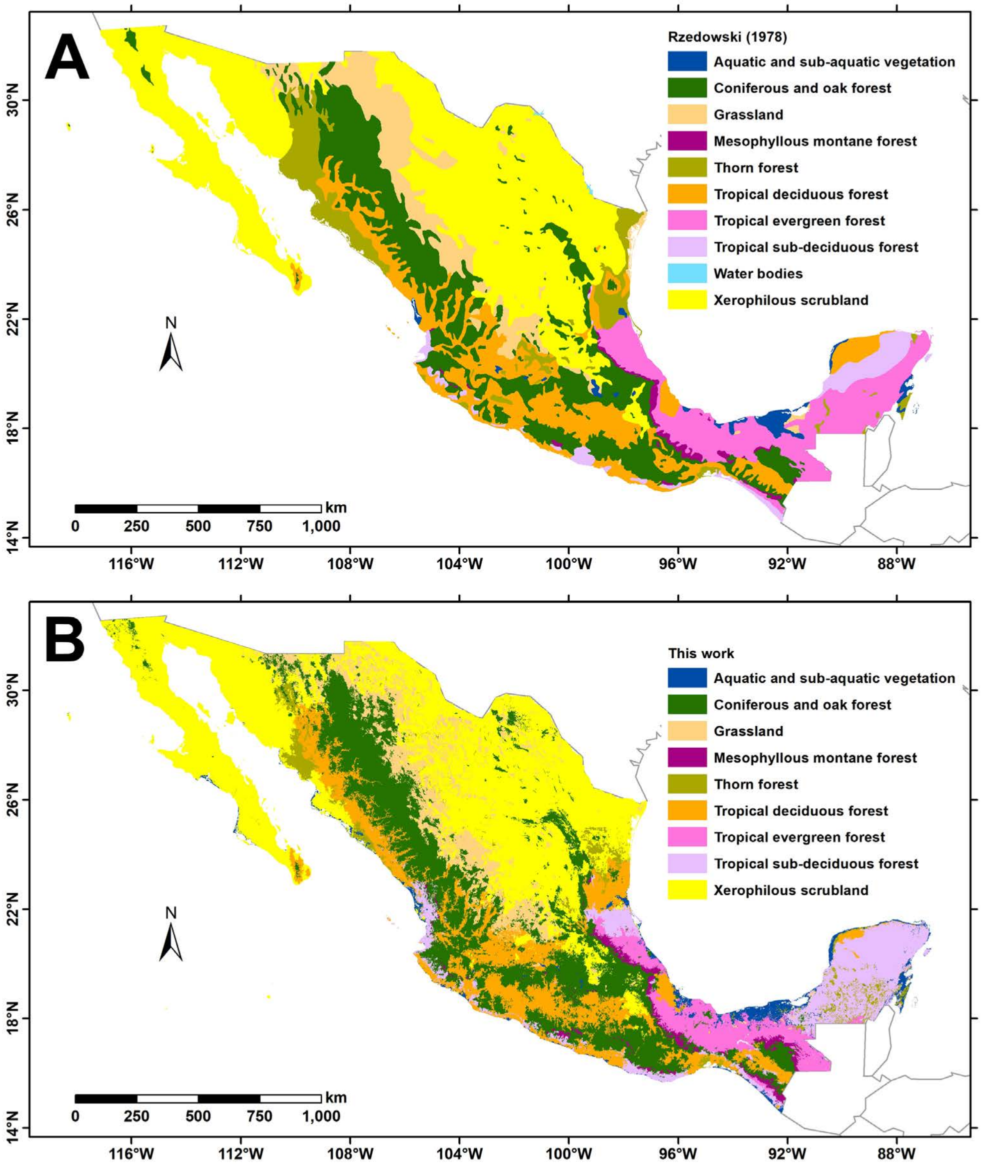

Figure 1. A. Rzedowski's (1978) potential vegetation map. B. Potential vegetation map by Ortiz and Villaseñor (this work).

characterized it as a communication tool that lacks the precision required of an analytical instrument and made a plea for an updated version.

In this study, we update Rzedowski's (1978) potential vegetation map and improve its resolution. Our objective was to make the map an instrument with the precision required for future studies focused on vegetation, ecology, biogeography, and systematics. Likewise, we expect that it will serve as a source of analysis and criticism that will assist in improving the resolution of future maps.

Materials and methods

First, the shapefile of Mexico's potential vegetation map (Rzedowski 1990) was downloaded from CONABIO's geoportal (http://www.conabio.gob.mx/informacion/gis/?vns=gis_root/usv/otras/vpr4mgw). The vegetation types and the surface area they occupy are shown in Table 1. The shapefile of the vegetation and land use map, series VII

(INEGI 2021), was downloaded from the same geoportal, at a scale of 1: 250,000. The 183 classes included in this latter map were assigned to the nine Rzedowski's vegetation types (Supplementary material 1). Classes that could not be assigned this way -for example, polygons without vegetation or with agricultural or urban use- were assigned to the "No vegetation" class. Polygons that were recognized as belonging to one of the nine Rzedowski's vegetation types were separated into a new layer. Subsequently, the centroid (coordinate calculated as the arithmetic mean of its minimum and maximum latitude and longitude coordinates) was calculated for each polygon and used to calculate Thiessen polygons (Figure 2A).

Table 1. Area (km^2) occupied by vegetation types according to Rzedowski's map (1990) and the map produced using Thiessen interpolation. The difference in area was calculated by subtracting the former from the latter. The number of sites with records of Asteraceae by vegetation type used in the conditional Kappa analyses is also shown.

Vegetation type	Rzedowski (%)	Thiessen (%)	Difference	Number of sites (%)
Coniferous and oak forest (BCO)	378,579 (19.3)	437,660 (22.4)	-59,081	3,116 (31.2)
Thorn forest (TF)	114,096 (5.8)	51,810 (2.7)	+62,286	257 (2.6)
Mesophyllous montane forest (MMF)	17,912 (0.9)	26,984 (1.4)	-9,072	904 (9.0)
Tropical deciduous forest (TDF)	276,965 (14.1)	258,744 (13.3)	+18,221	1,559 (15.6)
Tropical evergreen forest (TEF)	193,904 (9.9)	93,932 (4.8)	+99,972	295 (2.9)
Tropical sub-deciduous forest (TSF)	56,092 (2.9)	161,954 (8.3)	-105,862	563 (5.6)
Water bodies	771 (0.04)	0 (0)	+771	—
Xerophilous scrubland (XER)	736,236 (37.6)	721,946 (37.0)	+14,290	2,768 (27.7)
Grassland (PAZ)	159,669 (8.2)	157,872 (8.1)	+1,797	353 (3.5)
Aquatic and sub-aquatic vegetation (VA)	23,212 (1.2)	39,955 (2.0)	-16,743	185 (1.9)

Thiessen polygons, also known as Voronoi diagrams or Dirichlet tessellations, can be used as a spatial interpolation method that predicts the attributes of unsampled areas based on the known attributes of neighboring sampled areas (Brassel & Reif 1979, Yang 2009). Each trio of neighboring centroids was used to form a triangle (Delauney triangles). A perpendicular line was drawn in the middle of each side of the triangle, forming the edges of the Thiessen polygons and their intersections became the vertices of the polygons (Figure 2B). The technique has been used in various areas where it is necessary to know the spatial influence of a given phenomenon, including previous applications for vegetation mapping (Sanders *et al.* 2004, Xiao *et al.* 2005, Biró *et al.* 2006).

Once the Thiessen polygons were calculated, covering the entire territory of Mexico, they were intersected with the "No vegetation" polygons, to interpolate the corresponding vegetation type (Figure 2C). When a "No vegetation" polygon fell into two or more Thiessen polygons of different vegetation, the vegetation type with the largest area was assigned to form larger, more continuous vegetation patches. Finally, the original polygons with vegetation and the polygons with interpolated vegetation were integrated into a single layer (Figure 2D). Spatial operations were performed in ArcGIS 10.5 (ESRI 2013). The updated potential vegetation map was exported as a shapefile at a scale of 1:250,000 (Supplementary material 2).

The accuracy of a map describes how well the mapped features match the features that occur on the ground. The accuracy can be assessed quantitatively or qualitatively. Quantitative methods use independent observation points to estimate how often the mapped vegetation attribute matches that observed in the field (Story & Congalton 1986). To assess the reliability of the generated map, it was compared with the vegetation types reported for 10,000 georeferenced sites from a database that records herbarium collections of Asteraceae species from Mexico, managed by one of the authors (JLV). This database uses the nomenclature proposed by Rzedowski (1978) in the vegetation file (Table 1). The Asteraceae family was selected because it has a wide distribution in the country and is well-represented in all vegetation types (Villaseñor 2018).

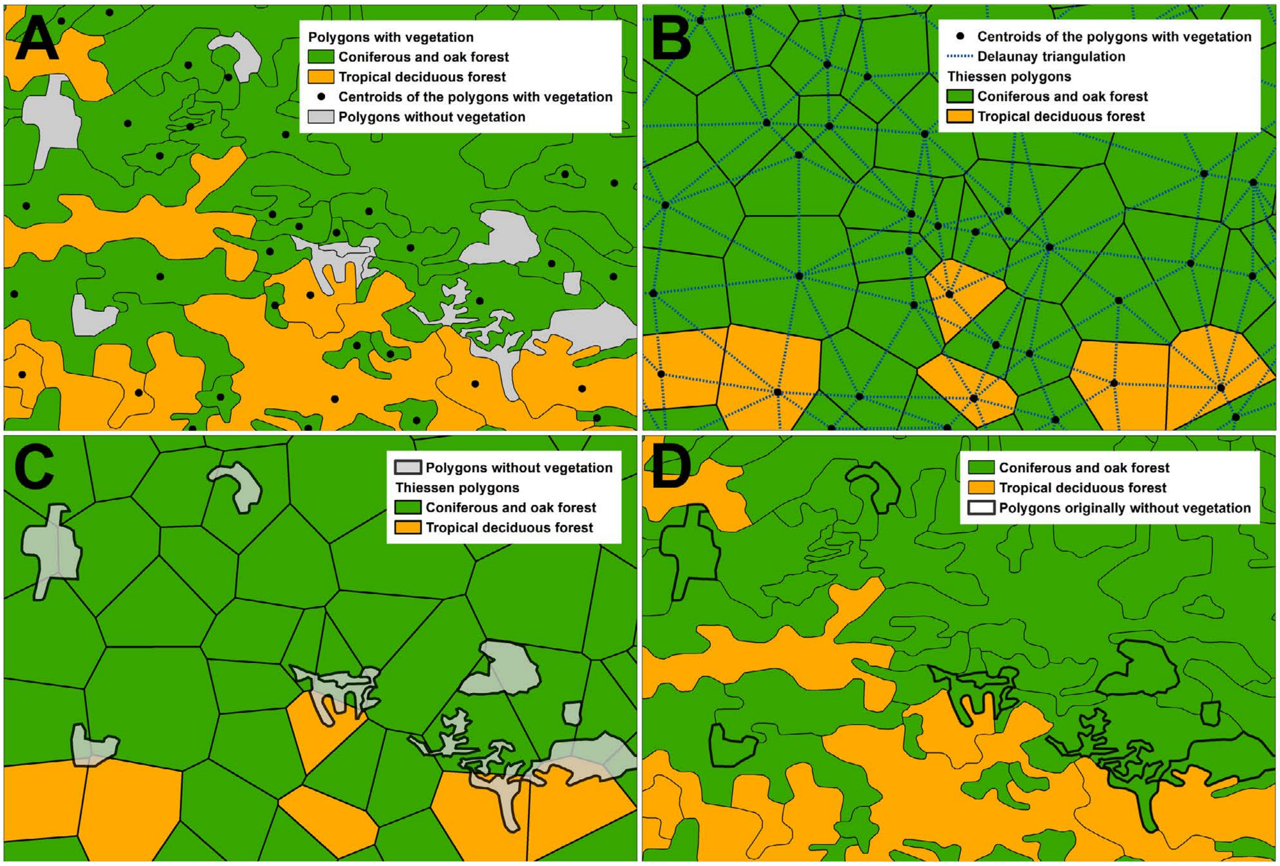

Figure 2. Example of the creation of Thiessen polygons and the updated vegetation map. A. The centroids of each polygon with an assigned vegetation type are shown as a point, and the polygons without recognized vegetation (in gray). B. The centroids were used to create Delaunay triangles and their sides to construct the bisector lines that constitute the margins of the Thiessen polygons. C. Thiessen polygons are used to interpolate the polygons without vegetation. D. The polygons with interpolated vegetation are merged to the polygons with vegetation in A.

The correspondence or discrepancy between the validation points and the map was quantified and used to construct confusion matrices to summarize the number of true positive (TP), true negative (TN), false negative (FN), and false positive (FP) cases (Story & Congalton 1986, Sánchez-Muñoz 2016) for all vegetation types. Using the mesophyllous montane forest (*bosque mesófilo de montaña*) as an example, a true positive (TP) occurred when a validation point of this vegetation type coincided with a polygon of the same vegetation type. A true negative (TN) occurred when a point where the mesophyllous montane forest was not recorded coincided with a polygon not assigned to this vegetation type. A false negative (FN) occurred when a validation point of this forest type coincided with a polygon not associated with mesophyllous montane forest. A false positive (FP) occurred when a validation point not recorded as mesophyllous montane forest coincided with a polygon of this vegetation type. These quantifications were performed for each vegetation type. The confusion matrices allowed the calculation of three attributes of quantitative precision: the Kappa statistic, sensitivity, and specificity.

Cohen's Kappa statistic (Cohen 1960) compares the observed agreement between a pair of categorical data sets to that which would occur by chance alone. Kappa ranges from -1 to 1, with a value closer to -1 reflecting a greater degree of disagreement, 0 indicating agreement due to chance, and 1 indicating complete agreement between the data sets. Landis & Koch (1977) and Monserud & Leemans (1992) used the following qualitative descriptors to characterize the degree of agreement suggested by the Kappa statistic: very poor to poor (< 0.4), fair (0.4-0.55), good (0.55-0.7), very good (0.7-0.85), and excellent (> 0.85). If the number of agreements observed is smaller than expected by chance, Kappa will be negative, indicating disagreement. Sensitivity (TP / (TP + FN)) is the proportion of cases correctly classified to a vegetation type (TP) relative to the total number of cases classified to that vegetation type, either correctly (TP) or incorrectly (FN). In contrast, specificity (TN / (FP + TN)) is the proportion of cases correctly classified as not belonging to a vegetation type (TN) relative to the total number of cases classified as not belonging to a vegetation type, either correctly (TN) or incorrectly (FP). The Kappa statistic, specificity, and sensitivity were also calculated for Rzedoswki´s (1990) potential vegetation map. These analyses were implemented in R 4.3.1 (R Core Team 2024), using the "yardstick" package (Kuhn *et al.* 2024).

Determining which of the categories evaluated has more or fewer classification errors is also informative. Rosenfield & Fitzpatrick-Lins (1986) proposed the Conditional Coefficient of Agreement (conditional Kappa) to measure each category's precision. To perform this calculation, we used the same 10,000 Asteraceae sites described above (according to their geographic coordinates), with their vegetation type observed in the field and the vegetation type of the map interpolated with the Thiessen polygons. The conditional Kappa value was also calculated for Rzedowski's (1990) potential vegetation map. These calculations were performed in Excel, using the formula proposed by Wang & Hu (2017).

For both maps, the percentages of omission and commission errors were also calculated for each vegetation type. Returning to the example of the mesophyllous montane forest, an omission error occurs when a validation point of this vegetation type coincides with a polygon different from the focal vegetation type (FN). A commission error occurs when a validation point not coming from this forest type coincides with a mesophyllous montane forest polygon (FP). In other words, omission and commission errors measure the fraction of validation points incorrectly classified by the map. The desirable goal is a proportion of error as low as possible (Fielding & Bell 1997).

Results

Figure 1B shows the updated potential vegetation map, and Table 1 shows the area corresponding to each vegetation type. Xerophilous scrub (*matorral xerófilo*) is the most widely distributed vegetation type, covering 37 % of the country's surface area, whereas mesophyllous montane forest occupies the smallest area, only 1.4 % of the national territory.

Table 2 shows the confusion matrices of the 10,000 Asteraceae sites used as verification points for Rzedowski's map and the Thiessen polygons map produced here. The Cohen's Kappa values showed a substantial consensus increase in the map produced with the Thiessen polygons (0.927, $P < 0.001$) relative to Rzedowski's map (0.588, $P < 0.001$). The map obtained here had high sensitivity (0.912) and specificity (0.992) values, while Rzedowski's map had medium sensitivity (0.583) yet high specificity (0.956). The conditional Kappa value, which shows the precision values for each vegetation type (Table 3), was also higher for our map than for Rzedowski's. In the latter, the vegetation type with the best rating was xerophilous scrub (0.82), while the Thiessen polygons increased the precision for this vegetation type to 0.99. The vegetation type with the lowest accuracy in Rzedowski's map was evergreen tropical forest (*bosque tropical perennifolio*) (0.3), while the map produced here had a consensus of 0.85. Regarding the proportions of omission and commission errors, the Thiessen polygons map generally had lower values compared to Rzedowski's map (Table 3).

Discussion

Vegetation maps are an effective way to process and illustrate spatial information describing vegetation cover. They play a significant role in real-world applications, such as their relationship with climate, soil, or other ecological factors; they are also relevant for understanding biodiversity at the plant community level, and in the design and planning of protected areas, land use, and environmental assessment (Pedrotti 2013).

Vegetation maps can be constructed using very different philosophies and methods, subject to debate. One of these controversies is whether vegetation maps should be represented as vegetation mosaics with defined boundaries (like

Table 2. Confusion matrices of Rzedowski's map and that produced with the Thiessen polygons, based on 10,000 Asteraceae records as verification sites. The values on the diagonal (in bold) indicate the sites that were correctly assigned; the values below the diagonal indicate the omission errors or false negatives (records with the indicated vegetation type but located in a different type); the values above the diagonal indicate the commission errors or false positives (records that, not belonging to the vegetation type, are recorded in it).

Rzedowski (1990)	**BCO**	**TF**	**MMF**	**TDF**	**TEF**	**TSF**	**XER**	**PAZ**	**VA**	**Total**
Coniferous and oak forest (BCO)	**2,336**	4	465	153	0	0	116	89	16	3,179
Thorn forest (TF)	11	**119**	0	105	1	16	65	0	10	327
Mesophyllous montane forest (MMF)	65	0	**304**	11	4	0	1	1	0	386
Tropical deciduous forest (TDF)	474	48	23	**1,189**	1	19	72	21	31	1,878
Tropical evergreen forest (TEF)	25	14	93	37	**278**	361	33	13	23	877
Tropical sub-deciduous forest (TSF)	25	3	4	22	3	**165**	9	1	6	238
Xerophilous scrubland (XER)	118	64	15	37	0	0	**2,084**	71	0	2,389
Grassland (PAZ)	48	3	0	1	1	2	344	**151**	3	553
Aquatic and sub-aquatic vegetation (VA)	14	2	0	4	7	0	44	6	**96**	173
Total	3,116	257	904	1,559	295	563	2,768	353	185	
This work										
Coniferous and oak forest (BCO)	**2,989**	0	56	64	0	0	21	9	9	3,148
Thorn forest (TF)	0	**227**	0	0	15	59	0	0	10	311
Mesophyllous montane forest (MMF)	58	0	**841**	0	0	0	2	5	0	906
Tropical deciduous forest (TDF)	40	12	3	**1,495**	0	0	17	3	5	1,575
Tropical evergreen forest (TEF)	14	0	3	0	**247**	0	14	10	0	288
Tropical sub-deciduous forest (TSF)	14	1	1	0	30	**504**	26	2	0	578
Xerophilous scrubland (XER)	0	16	0	0	0	0	**2,635**	0	0	2,651
Grassland (PAZ)	0	0	0	0	0	0	3	**324**	0	327
Aquatic and sub-aquatic vegetation (VA)	1	1	0	0	3	0	50	0	**161**	216
Total	3,116	257	904	1,559	295	563	2,768	353	185	

the maps analyzed in this work) or as vegetation gradients with fuzzy boundaries (Triepke 2017). The advantages or disadvantages of each of these approaches are beyond the scope of this study since only maps with sharp edges were analyzed here, although the interested reader is referred to Feilhauer *et al.* (2021).

More often than not, users must rely on existing maps, which may be based on data collected at different times, drawn at different scales, and with different accuracy levels (Küchler & Zonnelveld 1988, Demers 1991, Franklin 1995, Bredenkamp *et al.* 1998). Before using one map for a given analysis, users should evaluate the strengths and weaknesses of various maps, and then select one that meets their criteria for accuracy, precision, and currency that are compatible with their objectives. Selecting the most suitable map can ensure a certain degree of reliability in the results.

Although vegetation maps are useful in various applications, according to Demers (1991), they have been focused more on communication than analysis; in other words, the predominant purpose of vegetation maps has been to produce thematic visual patterns that correspond to the vegetation classification used. This communicative orientation underlying the production of some maps limits their analytical utility because attributes such as the map's accuracy, precision, and validity are left aside. These attributes are essential in the analytical paradigm of cartography, which focuses on the spatial operations that can be performed with a map, mainly with a geographic information system (Moellering 2000).

The process of creating vegetation maps of sufficient quality to be used for analytical operations requires abundant human and material resources. For example, Biró *et al.* (2006) worked for four years to create a vegetation map for a region in Hungary; their team consisted of 56 people, who used paper and digital cartography and satellite images processed in geographic information systems and field validation. Similarly, Su *et al.* (2020) updated China's vegetation map, involving 250 researchers over 20 years of work in the office and the field. In Mexico, INEGI regularly publishes (1997, 2001, 2005, 2009, 2013, 2016, 2021) updates to the country's land use and vegetation map, applying the institute's considerable human and material resources. Some alternatives consume fewer resources; for example, Eva *et al.* (2002) created a vegetation map for South America, using satellite images and an unsupervised automatic classification. Similarly, Vaca *et al.* (2011) produced a vegetation map for southeastern Mexico, using sites obtained from Rzedowski's map, to which they associated climatic variables, and with this, they trained a Random Forest model to predict the distribution of vegetation types. These applications can mean savings in the resources used to create vegetation maps.

Map accuracy describes the degree to which mapped attributes match attributes that occur on the ground. This can be assessed quantitatively or qualitatively. Quantitative methods use independent point observations to estimate how often the mapped vegetation state matches the observed state. Confusion matrices, which summarize the frequencies of matches and mismatches, can be used to calculate indices, such as the Kappa statistic, to estimate the agreement between mapped and observed attributes relative to that expected by chance (Legendre & Legendre 1998). To our knowledge, this study is the first to fully evaluate the accuracy of Rzedowski's (1990) digital vegetation map.

When comparing the predictive success using Cohen's Kappa coefficient, Rzedowski's potential vegetation map, with a Kappa value of 0.588, can be rated as good according to the criteria of Landis & Koch (1977) and Monserud & Leemans (1992), while the map based on Thiessen polygons (0.927) can be rated as excellent. Rzedowski's map also showed lower sensitivity (0.583) compared to the interpolated map with Thiessen polygons (0.912); this result indicates a lower capacity of the former to correctly classify the validation sites with the polygon, including the corresponding vegetation type. On the other hand, both maps have a high specificity, successfully identifying true negatives (TN).

When evaluating each vegetation type individually using conditional Kappa, Rzedowski's map shows comparatively lower agreement between the validation points and the vegetation polygons than the map obtained with the Thiessen polygons. This lower accuracy is also noted in the percentages of omission and commission errors, which are again higher in Rzedowski's map. A notable example of these differences is the tropical evergreen forest (TEF), with a conditional Kappa value of 0.3 (poor) in Rzedowski's map compared to the value of 0.85 (very good) in our map. This means that the former map performs poorly at correctly assigning the validation points of this vegetation type to the polygon corresponding to the TEF. Márcia-Barbosa *et al.* (2013) related omission and commission errors with the degree of underestimation or overestimation of a map. A high rate of omission errors is related to an underestimation of the map and a high rate of commission errors, with a map overestimation. The TEF polygon in Rzedowski's map has a commission error rate of 68 %; this high rate is mainly due to the high number of validation records from tropical sub-deciduous forests (TSF, 361 records, Table 2) that fall within it, which is substantially higher than the number of true positives (278 records). The high percentage of commission errors suggests that Rzedowski's TEF polygon is possibly overestimated. This overestimation is noticeable in the area of this vegetation type since it is more than twice as large as that of the polygon interpolated using Thiessen (Table 1), which has a lower percentage of commission errors (14 %). The confusion between these two types of vegetation may also stem from the fact that they are not readily distinguished, given their strong floristic and structural similarities. Such similarities often cause errors in identifying the local vegetation type by botanical collectors, who are not always sufficiently well-trained to discriminate against similar vegetation types. Another possible explanation for the difference between Rzedowski's map and ours is the change in vegetation due to successional or anthropogenic processes in the interval between both works. Similar changes have already been reported in other regions of the world (Pancer-Koteja *et al.* 2009, Kong *et al.* 2018, Zhang *et al.* 2022).

Another example of the contrasting accuracy of these maps is the mesophyllous montane forest (MMF); although Rzedowski's map has a very good conditional Kappa value (0.77), its percentage of omission errors is high

(66 %). The confusion matrix (Table 2) shows that the field records to validate the MMF classified by Rzedowski's map fall in TEF polygons (93 records) and coniferous and oak forests (65 records), which omit MMF polygons. The high percentage of omission errors suggests underestimating the polygon's area. The area of Rzedowski's MMF was only 66 % of that projected by our polygon (Table 1), which has better agreement with the validation points, as shown by an excellent conditional Kappa value (0.92) and low rates of omission (6 %) and commission errors (7 %). Vaca *et al*. (2011) and Cruz-Cárdenas *et al*. (2012) pointed out the underestimation of the surface occupied by the MMF in Mexico.

The vegetation map created by interpolation with Thiessen polygons generally shows better accuracy than the Rzedowski map. The tests indicate greater correspondence between the validation points relative to the polygons of the vegetation types in the interpolated map and lower accuracy in Rzedowski's map. Similar results were found by Vaca *et al.* (2011), who made a vegetation map of southeastern Mexico and compared it against that of Rzedowski using field sampling sites.

The map presented here is not error-free; Tables 2 and 3 list several of them. These flaws may arise from the mis-assignment of the vegetation types of the Land Use and Vegetation map (INEGI 2021) to one of the nine vegetation types of Rzedowski (1990) (Supplementary material 1). We may also have erroneously assigned "No vegetation" polygons that fell between two Thiessen polygons with interpolated vegetation. A further source of error could be the field validation records of Asteraceae, where the collector may not have recognized the vegetation type correctly or when the point fell in an ecotone (the fuzzy limits mentioned above); such is the case of tropical evergreen forest and tropical sub-deciduous forest discussed above.

Table 3. Conditional Kappa and percentages of omission and commission errors for the vegetation types of Rzedowski's map and the one produced with the Thiessen polygons. The values were obtained considering 10,000 verification sites of Asteraceae records.

Rzedowski (1990)	**Conditional Kappa**	**Omission (%)**	**Commission (%)**
Coniferous and oak forest (BCO)	0.61	25	26
Thorn forest (TF)	0.35	53	63
Mesophyllous montane forest (MMF)	0.77	66	21
Tropical deciduous forest (TDF)	0.57	23	36
Tropical evergreen forest (TEF)	0.3	5	68
Tropical sub-deciduous forest (TSF)	0.67	70	30
Xerophilous scrubland (XER)	0.82	24	12
Grassland (PAZ)	0.25	57	72
Aquatic and sub-aquatic vegetation (VA)	0.55	48	44
This work			
Coniferous and oak forest (BCO)	0.93	4	5
Thorn forest (TF)	0.72	11	27
Mesophyllous montane forest (MMF)	0.92	6	7
Tropical deciduous forest (TDF)	0.94	4	5
Tropical evergreen forest (TEF)	0.85	16	14
Tropical sub-deciduous forest (TSF)	0.86	10	12
Xerophilous scrubland (XER)	0.99	4	0
Grassland (PAZ)	0.99	8	0
Aquatic and sub-aquatic vegetation (VA)	0.74	12	25

A map created mainly with the aim of communication, such as Rzedowski's vegetation map, cannot be used under the analytical paradigm just because it can be incorporated into a geographic information system. Rzedowski's intention in 1978 was likely to provide a broad overview of the distribution of vegetation types in the country rather than providing a guide to deciding what vegetation type was present in a given location. It is important to recognize that maps mainly oriented toward communication may not prioritize accuracy and that the ability to upload the map into a GIS does not guarantee that it is appropriate for analytical applications; unfortunately, despite this reality, using such maps for purposes different from their original intention is becoming a common practice. In this paper, we calculated the same precision values for Rzedowski's map and the map interpolated with Thiessen polygons; these values are important because they will allow users to estimate the precision in future analyses that use them. This paper presents a method for creating a vegetation map, and the result (Supplementary material 2) is accessible as the product of this method. The methodological procedure can be adjusted to produce a vegetation map under any other classification scheme of vegetation types. This effort thus constitutes a heuristic exercise that can serve as a hypothesis to support or refute future vegetation analyses at different scales in the country. The updated map is available to users to evaluate whether it is appropriate for their analysis. To date, there is no other proposal for a vegetation map (potential or natural) for Mexico that would replace the great effort made by Rzedowski almost five decades ago. We are only aware of regional or state efforts (for example, González-Elizondo *et al.* 2007, Vaca *et al.* 2011, González-Elizondo *et al.* 2012, Siqueiros-Delgado *et al.* 2016). However, Rzedowski's map is the main source used to discuss vegetation types in the country, despite its coarse scale and the difficulty of specifying many of the types he used in his classification locally. Proof of this is the number of omission and commission errors found. A better alternative is undoubtedly the land use and vegetation maps generated by INEGI (2022), although these, besides natural vegetation, include areas that humans have transformed. Unfortunately, those maps come with their own difficulties for interpretation and use (too many categories, many of them slight variations of one or another, or poor knowledge of interpretation using a geographic information system), which limits its usefulness for many users interested in the subject.

We hope that the map presented here, and the accompanying shapefile format, will be useful in future studies that require interpretation of the vegetation types in Mexico. But above all, we hope it will be a basis for critique, review, and update as new information is obtained. We hope that the vegetation synthesis presented here, which updates Dr. J. Rzedowski's key contribution to the understanding of vegetation types in Mexico, will help to achieve, in a short time, a more generalized proposal that serves many other objectives. As Ibarra-Manríquez *et al.* (2022) argue, this would greatly impact ecology and many other areas of biology and related sciences.

Acknowledgments

We would like to thank Drs. Guillermo Ibarra Manriquez, Miguel Murguía Romero, and Eduardo A. Pérez García for their valuable comments, which helped improve the quality and content of this work. Lynna Kiere and David Gernandt revised the language, and their remarks substantially improved its content. Comments from two anonymous reviewers also improved the content of this manuscript.

Supplementary material

The supplementary material for this article can be consulted here: https://doi.org/10.17129/botsci.3598

Literature cited

Biró M, Révész A, Horváth F, Molnár Z. 2006. Point-based mapping of the actual vegetation of a large area in Hungary-description, usability, and limitation of the method. *Acta Botanica Hungarica* **48**: 247-269. DOI: https://doi.org/10.1556/abot.48.2006.3-4.2

Brassel KE, Reif D. 1979. A procedure to generate Thiessen Polygons. *Geographical Analysis* **11**: 289-303. DOI: https://doi.org/10.1111/j.1538-4632.1979.tb00695.x

Bredenkamp G, Chytrý M, Fischer HS, Neuhäuslová Z, van der Maarel E. 1998. Vegetation mapping: Theory, methods and case studies: Introduction. *Applied Vegetation Science* **1**: 162-164.

Carranza ML, Ricotta C, Fortini P, Blasi C. 2003. Quantifying landscape change with actual vs. potential natural vegetation maps. *Phytocoenologia* **33**: 591-602. DOI: https://doi.org/10.1127/0340-269X/2003/0033-0591

Chiarucci A, Araújo MB, Decocq G, Beierkuhnlein C, Fernández-Palacios JM. 2010. The concept of potential natural vegetation: an epitaph? *Journal of Vegetation Science* **21**: 1172-1178. DOI: https://doi.org/10.1111/j.1654-1103.2010.01218.x

Cohen J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* **20**: 37-46. DOI: https://doi.org/10.1177/001316446002000104

Cruz-Cárdenas G, Villaseñor JL, López-Mata L, Ortiz E. 2012. Potential distribution of humid mountain forest in Mexico. *Botanical Sciences* **90**: 331-340. DOI: https://doi.org/10.17129/botsci.394

Demers MN. 1991. Classification and purpose in automated vegetation maps. *Geographical Review* **81**: 267-280. DOI: https://doi.org/10.2307/215631

ESRI. 2013. *ArcGIS*. Environmental Systems Research Institute. USA: Redlands.

Eva HD, de Miranda EE, Di Bella CM, Gond V, Huber O, Sgrenzaroli M, Jones S, Coutinho A, Dorado A, Guimarães M, Elvidge C, Achard F, Belward AS, Bartholomé E, Baraldi A, De Grandi G, Vogt P, Fritz S, Hartley A. 2002. *A vegetation map of South America*. Luxemburg: Office for Official Publications of the European Communities. ISBN: 92-894-4449-5

Feilhauer H, Zlinszky A, Kania A, Foody GM, Doktor D, Lausch A, Schmidtlein S. 2021. Let your maps be fuzzy!—Class probabilities and floristic gradients as alternatives to crisp mapping for remote sensing of vegetation. *Remote Sensing in Ecology and Conservation* **7**: 292-305. DOI: https://doi.org/10.1002/rse2.188

Fielding AH, Bell JF. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* **24**: 38-49. DOI: https://doi.org/10.1017/S0376892997000088

Franklin J. 1995. Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography* **19**: 474-499. DOI: https://doi.org/10.1177/030913339501900403

González-Elizondo MS, González-Elizondo MG, Márquez-Linares MA. 2007. *Vegetación y ecorregiones de Durango*. Durango, México: Plaza y Valdés. ISBN: 978-970-722-698-2

González-Elizondo MS, González-Elizondo M, Tena-Flores JA, Ruacho-González L, López-Enríquez IL. 2012. Vegetación de la Sierra Madre Occidental, México: Una síntesis. *Acta Botanica Mexicana* **100**: 351-403. DOI: https://doi.org/10.21829/abm100.2012.40

González-Medrano F. 2004. *Las comunidades vegetales de México: propuesta para la unificación de la clasificación y nomenclatura de la vegetación de México*. México: Instituto Nacional de Ecología. ISBN: 968-817-611-7

Greller AM. 2000. Vegetation in the floristic regions of North and Central America. *In*: Lentz D, ed. *Imperfect Balance: Landscape Transformations in the Pre-Columbian Americas*. New York Chichester, West Sussex: Columbia University Press. pp. 39-88. ISBN: 978-0-231-11157-7.

Ibarra-Manríquez G, González-Espinosa M, Martínez-Ramos M, Meave JA. 2022. From vegetation ecology to vegetation science: current trends and perspectives. *Botanical Sciences* **100**: 137-174. DOI: https://doi.org/10.17129/botsci.3171

INEGI [Instituto Nacional de Estadística y Geografía] - INE [Instituto Nacional de Ecología]. 1997. Uso del Suelo y Vegetación, escala 1:250 000, Serie I. Continuo Nacional. Digitalización de las cartas de uso del suelo y vegetación elaboradas por INEGI entre los años 1980-1991 con base en fotografías aéreas de 1968-1986. México, D. F.: Instituto Nacional de Ecología - Dirección de Ordenamiento Ecológico General e Instituto Nacional de Estadística, Geografía e Informática.

INEGI. 2001. Uso del Suelo y Vegetación. Escala 1:250 000, Serie II. Continuo Nacional. Aguascalientes, México: Dirección General de Geografía. Instituto Nacional de Estadística y Geografía.

INEGI. 2005. Uso del Suelo y Vegetación. Escala 1:250 000, Serie III. Continuo Nacional. Aguascalientes, México: Dirección General de Geografía. Instituto Nacional de Estadística y Geografía.

INEGI. 2009. Uso del Suelo y Vegetación. Escala 1:250 000, Serie IV. Continuo Nacional. Aguascalientes, México: Dirección General de Geografía. Instituto Nacional de Estadística y Geografía.

INEGI. 2013. Conjunto de Datos Vectoriales de Uso de Suelo y Vegetación. Escala 1:250 000, Serie V. Capa unión. Edición: 2. Aguascalientes, México: Instituto Nacional de Estadística y Geografía.

INEGI. 2016. Conjunto de Datos Vectoriales de Uso de Suelo y Vegetación. Escala 1:250 000, Serie VI. Capa unión. Edición: 1. Aguascalientes, México: Instituto Nacional de Estadística y Geografía.

INEGI. 2021. Conjunto de Datos Vectoriales de Uso de Suelo y Vegetación. Escala 1:250 000, Serie VII. Conjunto Nacional. Edición: 1. Aguascalientes, México: Instituto Nacional de Estadística y Geografía.

INEGI. 2022. Guía para la interpretación de cartografía: uso del suelo y vegetación: escala1:250 000: serie VII. México: INEGI.

Kong D, Miao C, Borthwick AG, Lei X, Li H. 2018. Spatiotemporal variations in vegetation cover on the Loess Plateau, China, between 1982 and 2013: Possible causes and potential impacts. *Environmental Science and Pollution Research* **25**: 13633-13644. DOI: https://doi.org/10.1007/s11356-018-1480-x

Küchler AW. 1953. Some uses of vegetation maps. *Ecology* **34**: 629-636. DOI: https://doi.org/10.2307/1929736

Küchler AW. 1984. Ecological vegetation maps. *Vegetatio* **55**: 3-10. DOI: https://doi.org/10.1007/BF00039975

Küchler AW, Zonneveld IS. 1988. *Vegetation mapping*. Vol. 10. Dordrecht: Kluwer Academic Publisher. ISBN: 978-94-010-7885-6. DOI: https://doi.org/10.1007/978-94-009-3083-4

Kuhn M, Vaughan D, Hvitfeldt E. 2024. *yardstick: Tidy Characterizations of Model Performance*. R package version 1.3.1, https://yardstick.tidymodels.org https://github.com/tidymodels/yardstick. Accessed May 25, 2024.

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. *Biometrics* **33**: 159-174. DOI: https://doi.org/10.2307/2529310

Legendre P, Legendre L. 1998. *Numerical Ecology*. Developments in Environmental Modelling 20. Amsterdam: Elsevier. ISBN: 0-444-89249-4

Loidi J, Del Arco M, Pérez de Paz PL, Asensi A, Díez-Garretas B, Costa M, Díaz-González T, Fernández-González F, Izco J, Penas A, Rivas-Martínez S, Sánchez-Mata D. 2010. Understanding properly the potential natural vegetation concept. *Journal of Biogeography* **37**: 2209-2211. DOI: https://doi.org/10.1111/j.1365-2699.2010.02302.x

Márcia-Barbosa A, Real R, Muñoz AR, Brown JA. 2013. New measures for assessing model equilibrium and prediction mismatch in species distribution models. *Diversity and Distributions* **19**: 1333-1338. DOI: https://doi.org/10.1111/ddi.12100

Meave JA, Ibarra-Manríquez G, Larson-Guerra J. 2016. Vegetación: Panorama histórico, rasgos generales y patrones de pérdida. *In*: Moncada-Maya JO, López-López Á, eds. *Geografía de México: Una Reflexión Espacial Contemporánea*. México: Universidad Nacional Autónoma de México. pp. 216-234. ISBN: 978-607-02-8276-8

Moellering H. 2000. The scope and conceptual content of analytical cartography. *Cartography and Geographic Information Science* **27**: 205-224. DOI: https://doi.org/10.1559/152304000783547858

Monserud RA, Leemans R. 1992. Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling* **62**: 275-293. DOI: https://doi.org/10.1016/0304-3800(92)90003-W

Olson DM, Dinerstein E, Wikramanayake ED, Burgess ND, Powell GV, Underwood EC, D'amico JA, Itoua I, Strand HE, Morrison JC, Loucks CJ, Allnutt TF, Ricketts TH, Kura Y, Lamoreux JF, Wettengel WW, Hedao P, Kassem KR. 2001. Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* **51**: 933-938. DOI: https://doi.org/10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2

Pancer-Koteja E, Szwagrzyk J, Guzik M. 2009. Quantitative estimation of vegetation changes by comparing two vegetation maps. *Plant Ecology* **205**: 139-154. DOI: https://doi.org/10.1007/s11258-009-9604-5

Paz-Pellat F, Marín MI, Bolaños-González M, Etchevers-Barra JD, de Jong B, Herrera J. 2020. La urgente necesidad de México de contar con datos de actividad a escala apropiada para el sector agricultura, forestería y otros usos del suelo. *Elementos para Políticas Públicas* **4**: 61-77.

Pedrotti F. 2013. *Plant and vegetation mapping*. Heidelberg: Springer Science & Business Media. ISBN. 978-3-642-30234-3

R Core Team. 2024. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rosenfield GH, Fitzpatrick-Lins K. 1986. A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering and Remote Sensing* **52**: 223-227.

Rzedowski J. 1978. *Vegetación de México.* DF, México: Editorial Limusa.

Rzedowski J. 1990. *Vegetación Potencial*. IV.8.2. Atlas Nacional de México. Vol. II. Escala 1:4,000,000. México: Instituto de Geografía, Universidad Nacional Autónoma de México. http://www.conabio.gob.mx/informacion/gis/?vns=gis_root/usv/otras/vpr4mgw

Sánchez-Muñoz JM. 2016. Análisis de calidad cartográfica mediante el estudio de la Matriz de Confusión. *Pensamiento Matemático* **6**: 9-26.

Sanders ME, Dirkse GM, Slim PA. 2004. Objectifying thematic, spatial and temporal aspects of vegetation mapping for monitoring. *Community Ecology* **5**: 81-91. DOI: https://doi.org/10.1556/ComEc.5.2004.1.8

Siqueiros-Delgado ME, Rodríguez-Avalos JA, Martínez-Ramírez J, Sierra-Muñoz JC. 2016. Situación actual de la vegetación del estado de Aguascalientes, México. *Botanical Sciences* **94**: 455-470. DOI: https://doi.org/10.17129/botsci.466

Story M, Congalton RG. 1986. Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing* **52**: 397-399.

Su Y, Guo Q, Hu T, Guan H, Jin S, An S, Chen X, Guo K, Hao Z, Hu Y, Huang Y, Jiang M, Li J, Li Z, Li X, Li X, Liang C, Liu R, Liu Q, Ni H, Peng S, Shen Z, Tang Z, Tian X, Wang X, Wang R, Xie Z, Xie Y, Xu X, Yang X, Yang Y, Yu L, Yue M, Zhang F, Ma K. 2020. An updated vegetation map of China (1: 1,000,000). *Science Bulletin* **65**: 1125-1136. DOI: https://doi.org/10.1016/j.scib.2020.04.004

Triepke FJ. 2017. Fuzzy classification of vegetation for ecosystem mapping. *In:* Remmel TK, Perera AH, eds. *Mapping Forest Landscape Patterns*. New York: Springer. pp. 63-103. ISBN 978-1-4939-7329-3

Vaca RA, Golicher DJ, Cayuela L. 2011. Using climatically based random forests to downscale coarse-grained potential natural vegetation maps in tropical Mexico. *Applied Vegetation Science* **14**: 388-401. https://doi.org/10.1111/j.1654-109X.2011.01132.x

Velázquez A, Medina García C, Durán Medina E, Amador A, Gopar Merino LF. 2016. *Standardized Hierarchical Vegetation Classification. Mexican and Global Patterns.* Cham: Springer. ISBN: 978-3-319-41221-4

Villaseñor JL. 2018. Diversidad y distribución de la familia Asteraceae en México. *Botanical Sciences* **96**: 332-358. DOI: https://doi.org/10.17129/botsci.1872

Wang M, Hu C. 2017. Predicting *Sargassum* blooms in the Caribbean Sea from MODIS observations. *Geophysical Research Letters* **44**: 3265-3273. DOI: https://doi.org/10.1002/2017GL072932

Xiao X, Gertner G, Wang G, Anderson AB. 2005. Optimal sampling scheme for estimation landscape mapping of vegetation cover. *Landscape Ecology* **20**: 375-387. DOI: https://doi.org/10.1007/s10980-004-3161-z

Yang X. 2009. Spatial Interpolation. *In*: Karimi HA, ed. *Handbook of research on geoinformatics*. Hershey, Pennsylvania: Information Science Reference. pp. 129-136. ISBN: 978-1-59140-996-0

Zerbe S. 1998. Potential natural vegetation: validity and applicability in landscape planning and nature conservation. *Applied Vegetation Science* **1**: 165-172. DOI: https://doi.org/10.2307/1478945

Zhang YX, Liu Q, Wang YK, Huang JH. 2022. Assessing the impacts of climate change and anthropogenic activities on vegetation in southwest China. *Journal of Mountain Science* **19**: 2678-2692. DOI: https://doi.org/10.1007/s11629-021-6984-z

---

Associate editor: Eduardo Ruiz Sanchez

Author contributions: EO wrote the first draft of the manuscript and performed the analyses and digital mapping. JLV developed the original idea, provided the data for analyses, and contributed to the final manuscript.

Supporting Agencies: This research received no specific grant from any funding agency

Conflicts of interest: The authors declare that there is no conflict of interest, financial or personal, in the information, presentation of data, and results of this article.