EGYPTIAN
JOURNAL
OF BOTANY
(EJBO)

Print ISSN: 0375-9237
Online ISSN: 2357-0350

Ecological niche and potential distribution of *Stenocereus queretaroensis* Weber in the Chapala subbasin, Mexico

Emmanuel Torres-Vázquez, Gustavo Cruz-Cárdenas, Salvador Ochoa-Estrada, José Teodoro Silva-García, Yareli Joselin Cadena-Rodríguez, Rebeca Flores-Magallón

PUBLISHED BY
THE EGYPTIAN
BOTANICAL SOCIETY

EGYPTIAN JOURNAL OF BOTANY (EJBO)

Ecological niche and potential distribution of *Stenocereus queretaroensis* Weber in the Chapala subbasin, Mexico

Emmanuel Torres-Vázquez[1], Gustavo Cruz-Cárdenas[1], Salvador Ochoa-Estrada[1], José Teodoro Silva-García[1], Yareli Joselin Cadena-Rodríguez[2], Rebeca Flores-Magallón[1]

[1]*Instituto Politécnico Nacional, Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional, Unidad Michoacán, COFAA, Justo Sierra No. 28, Jiquilpan 59510, Michoacán, Mexico*

[2]*Tecnológico Superior de Jalisco, Unidad Académica Tamazula, Carretera Tamazula-Santa Rosa 329, Tamazula de Gordiano, Jalisco*

Cacti provides ecosystem services and mitigates climate change. Some cacti are endemic to Mexico. One example is *Stenocereus queretaroensis* (family Cactaceae), which also has commercial value for the local community in the Chapala subbasin. The objectives of this work were to determine the ecological niche and potential distribution of *S. queretaroensis* and to identify the ideal areas for its conservation and use in the Chapala subbasin. Four types of predictors were used: climate, soil properties, vegetation indices, and relief attributes. Presence and absence data were obtained during the field work. Comparison of means and principal component analysis were applied to the environmental predictors to determine the ecological niche of this species. The potential distribution of this species was predicted to be using an ensemble of six algorithms and five suitability classes. The results of both the mean comparison test and principal component analysis revealed that eight variables define the ecological niche of *S. queretaroensis*. The highly suitable and moderately suitable classes are concentrated in the central part of the studied subbasin, specifically around the periphery of Lake Chapala, as well as in the northeastern and southeastern regions of the basin. Areas with Leptosols, which have shallow soil and good drainage, are the most suitable for *S. queretaroensis* conservation.

Keywords: Machine learning algorithms; diversity and richness of species; remote sensing; Leptosols

ARTICLE HISTORY
Submitted: October 24, 2024
Accepted: December 26, 2024

CORRESPONDENCE TO
Gustavo Cruz-Cárdenas,
Instituto Politécnico Nacional, Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional, Unidad Michoacán, COFAA, Justo Sierra No. 28, Jiquilpan 59510, Michoacán, Mexico
Email: guscruz@ipn.mx

DOI: 10.21608/ejbo.2024.330886.3055

EDITED BY: M. Abd El-Ghani

INTRODUCTION

Cacti are important in arid and semiarid areas. Cacti provide essential ecosystem services; for example, they have been used as soil stabilizers to prevent erosion and as forage sources. Cacti produce mucilage, gums, and pectins, and their stems and fruits are used for consumption and medicinal purposes (García-Cruz et al., 2022). Additionally, cacti serve as raw materials for construction and are used in social and religious events (Wolf & Rio, 2003; Flores & Macías, 2008). Additionally, cacti can contribute to mitigating climate change through their CAM physiology, which makes them efficient for carbon dioxide capture (Ranjan et al., 2016). Seventy-five percent of cactus species are endemic to Mexico (Goettsch et al., 2019). *Stenocereus queretaroensis* is among the important endemic cacti species because of the demand for its fresh fruit (Campos-Rojas et al., 2011; Ruán-Tejeda et al., 2014). Additionally, this species has a high antioxidant capacity (García-Cruz et al., 2022). However, studies on the habitat of this species, specifically its ecological niche and potential distribution, are lacking. Identifying environmentally suitable areas can guide the conservation and use of this species. Previous works, such as those by Huerta-Martínez et al. (1999), generally mention the values of environmental factors where some populations of this species are present.

According to Hutchinson (1957), an ecological niche is defined as the multidimensional space where species develop, establish, and reproduce, owing to the biotic and abiotic environment present in this space. In recent years, predictive algorithms based on the concept of ecological niches have been developed. Some of these algorithms use correlative techniques to determine the ecological niche (ENM) and potential distribution (PDM) of species, relating the environment (biotic and abiotic variables) to the presence of one or more species (Vicente-Silva et al., 2022). In general, to determine the ENM and PDM, data on presence and absence, or pseudoabsence (created from records of the species), are used. However, the results of the models can be improved through ensemble methods (different algorithms); that is, the best prediction of each algorithm can be obtained by using weights (Gašparovičová et al., 2022). Importantly, most ENM and PDM studies use only presence data due to the difficulty of finding records of species absence, as these records are not captured in databases (Illoldi-Rangel et al., 2012; Vicente-Silva et al., 2022; Cavalcante & Sampaio, 2022). Therefore, Brotons et al. (2004) and Václavík & Meentemeyer (2009) recommend the use of absence data whenever available to improve prediction accuracy, as including absence data makes the modeling of ecological niches and species distributions more reliable.

Therefore, the objective of this work was to determine the ecological niche and potential distribution of *S. queretaroensis* using presence and absence records in an ensemble of algorithms. Once the probability of the potential distribution of the species was obtained, a classification was carried out to determine the ideal areas for its presence, future distribution, and conservation.

MATERIALS AND METHODS

Study area

The study area corresponds to the Chapala subbasin, which has a total area of 4715.89 km², and its geographical coordinates are 20°22’ North, 19°50’ South, 102°21’ East, and 103°34’ West (Figure 1). The Chapala subbasin is located within the states of Jalisco and Michoacán, and it includes a total of 38 municipalities, 13 in Jalisco and 15 in Michoacán. The dominant climates are semi-warm, sub-humid, and temperate (García & CONABIO, 2008). The land use in the subbasin is distributed as follows: 24% rainfed agriculture, 20% irrigated agriculture, 15% low deciduous forest, 9% grassland, and 5% oak forest (INEGI, 2021). The predominant soil types are Vertisols (49%) and Leptosols (13%), and the remaining 11% of soils are Luvisols (INEGI, 2014).

Environmental predictors

Four types of predictors were used: climate, soil properties, normalized vegetation indices, and relief attributes. In total, there were 25 variables (Table 1).

Climate: Information for the climatic predictors was obtained from 59 climatic stations of the National Meteorological Service (https://smn.conagua.gob.mx/es/), considering data from the period 1951-2010. Climatic stations were selected both inside and outside the study area to ensure more reliable interpolation. The climatic variables included minimum temperature, maximum temperature, precipitation, and evaporation.

Soil properties: Soil samples were collected during field trips, resulting in a total of 66 samples of 1 kg each that were obtained at a depth of 0–30 cm of these, 46 samples were collected from areas where S. *queretaroensis* was present, and 20 samples were collected from areas where the species was absent. From the samples obtained, the following soil properties, which influence plant adaptability, were determined (Mandal et al., 2020): pH, electrical conductivity, organic matter, texture, bulk density, total nitrogen, phosphorus, and potassium. The determinations were made by the official Mexican standard, which establishes specifications for fertility, salinity, and soil classification (SEMARNAT, 2002).

Vegetation indices: Data were downloaded from the remote Landsat 8 sensor via Earth Explorer (https://earthexplorer.usgs.gov/) via grids 28/46 and 27/46 from November 2021 to May 2022. The normalized difference vegetation index (NDVI) was generated for each of the scenes:

$$NDVI = \frac{NIR-RED}{NIR+RED} \quad (1)$$

where NIR = the near-infrared band, and RED = the red band.

Topographic attributes: The DEM (https://www.inegi.org.mx/app/geo2/elevacionesmex/) was used to generate the aspect variable.

Interpolation of climate and soil property predictors

For climatic and property predictors, it was necessary to apply geostatistics to generate rasters. This process consists of exploration data analysis, structural analysis, and spatial prediction. In the exploratory analysis of the data, the Shapiro–Wilk test of normality (Shapiro and Wilk, 1965) was applied to determine whether the data needed transformation if the p-value was less than 0.05.

In the structural analysis, seven theoretical models were evaluated (bessel, spherical, exponential, gaussian, linear, matern, and pentaspherical), and the one with the lowest error value was selected according to the equation of Kravchenko & Bullock (1999).

$$SSE = \sum_{i=1}^{m} (w_i[\hat{\gamma} - \gamma]^2 \quad (2)$$

where m is the number of lags (two locations separated by a given distance), γ is the semi-variance value of each distance, $\hat{\gamma}$ is the semi-variance value of the allowable prediction model, and w_i is the semi-variance factor obtained from the equation of Kravchenko & Bullock (1999).

$$w_i = \frac{N}{\gamma^2} \quad (3)$$

where N is the number of pairs of points used to calculate γ for each distance.

After selecting the theoretical model that best fits the experimental semivariogram, spatial predictions were carried out. Three kriging methods (simple, ordinary, and universal) were evaluated using a 10-fold cross-validation technique (Kozak & Kozak, 2003).

Table 1. Environmental predictors used for modeling *S. queretaroensis*.

Type	Source of information
Climate Ev= Evaporation (mm) Pp= Precipitation (mm) Tmax= Maximum temperature (°C) Tmin= Minimum temperature (°C)	Climatological normal of 59 Meteorological stations of the SMN from 1951-2010
Soil properties EC= Electrical conductivity (dS/m) BD= Bulk density (g/cm^3) pH OM= Organic matter (%) Sand (%) Clay (%) TN= Total nitrogen (%) P (mg/kg) K (mg/kg)	Determination in laboratory according to the NOM-021-RECNAT-2000
Normalized Difference Vegetation Indices 10 NDVI	Landsat L8 OLI/TIRS C2 L1. Row/path 28, 46 y 27, 46.
Topographic attributes dem= Altitude (m) Aspect (°)	Set of elevations Mexican from INEGI.

The kriging method with the lowest precision error was selected to perform the interpolation based on the mean error (ME), which should be close to 0.0. Similarly, the root means square error (RMSE) must be less than the sample variance and the mean square deviation ratio (MSDR).

$$ME = \frac{1}{N}\sum_{i=1}^{N}[\widehat{z(x_i)} - z(x_i)]^2 \qquad (4)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[\widehat{z(x_i)} - z(x_i)]^2} \qquad (5)$$

$$MSDR = \frac{1}{N}\sum_{i=1}^{N}\frac{[\widehat{z(x_i)} - z(x_i)]^2}{\sigma(x_i)} \qquad (6)$$

where $\widehat{z(x_i)}$ is the estimated value, $z(x_i)$ is the known value, *N* is the sample size, and $\sigma(x_i)$ is the variance of the prediction at point x_i.

Data on the presence and absence of the species

Presence records were obtained during field trips throughout the study area, where populations of the species were observed, resulting in a total of 58 records. Absence records were also obtained during field trips, where no populations of the species were observed, resulting in a total of 21 absence records (Figure 1).

Ecological niche and potential distribution of the species

A mean comparison test (p-value 0.05) was applied to the environmental predictors to determine if there was a significant difference between the presence and absence records. In addition, principal component analysis (PCA) was used to project the environmental space (ecological niche) and select predictors for modeling the potential distribution of the species (Cruz-Cárdenas et al., 2014; Hu et al., 2016). For modeling, 58 presence records and 21 absence records were used. The presence data were partitioned, with 70% used for training and the remaining 30% used for algorithm validation. Bootstrap resampling with 10 repetitions was applied. The algorithms used included the domain, generalized linear model, Maxent, artificial neural network, support vector machine, and random forest algorithms. An ensemble was subsequently performed using the average weighting criterion of the area under the curve (AUC; Lastiri-Hernández et al., 2020).

Distribution suitability of the species

The ensemble species distribution model was classified into five groups (Acharya et al., 2019): highly suitable (0.8-1.0), suitable (0.6-0.8), marginally suitable (0.4-0.6), poorly suitable (0.2-0.4), and very poorly suitable (0.0-0.2).

Software

QGIS (QGIS.org, 2023) was used for spatial data configuration, and R (R Core Team, 2023) was used for statistical analyses and ecological niche and species distribution modeling. The libraries used in R were caret, dismo, do parallel, raster, rJava, RStoolbox, and sdm.

RESULTS

Interpolation of the climate and soil property predictors

Four of the thirteen variables had a normal distribution, and transformations were applied to the remaining variables to adjust for normality (Table 2). The theoretical models that best fit the empirical semi-variogram were as follows: bessel for evaporation (Ev) and precipitation; gaussian for maximum temperature, clay, and bulk density; spherical for phosphorus; exponential for minimum temperature and organic matter; linear for pH, electrical conductivity, and sand; and finally, pentaspherical for potassium and total nitrogen (Table 3).

Table 2. Results of the Shapiro-Wilk test of the variables used in the predictions.

Variable	*p-value*	Transformation
Ev (mm)	0.20	
Pp (mm)	0.01	
Tmax (°C)	0.09	$(x)^2$
Tmin (°C)	0.11	$(x)^2$
Clay (%)	0.34	
pH	0.64	1/x
EC (dS/m)	0.17	Log(x)
K (mg/kg)	0.10	Log(x)
Sand (%)	0.06	
BD (g/cm^3)	0.67	
TN (%)	0.49	√(x)
OM (%)	0.23	√(x)
P (mg/kg)	0.08	√(x)

Ev, evaporation; Pp, precipitation; Tmax, maximum temperature; Tmin, minimum temperature; EC, electric conductivity; BD, bulk density; TN, total nitrogen.

Table 3. Geostatistical results of theoretical models fitting.

Variable	Model	Partial sill	Nugget	Range
Ev (mm)	Bessel	37521	2166	10745
Pp (mm)	Bessel	10829	0	6135
Tmax (°C)	Gaussian	4683	4924	16115
Tmin (°C)	Exponential	1066	269	18887
Clay (%)	Gaussian	98.335	175	25778
pH	Linear	0.0002	0.0004	11717
EC (dS/m)	Linear	0.032	0.343	11308
K (mg/kg)	Pentaspherical	1.026	0.092	94310
Sand (%)	Linear	75.026	60.101	10715
BD (g/cm^3)	Gaussian	0	0.091	15556
TN (%)	Pentaspherical	0.001	0.029	127855
OM (%)	Exponential	0.623	0.005	4952
P (mg/kg)	Spherical	1.005	4.995	80361

Ev, evaporation; Pp, precipitation; Tmax, maximum temperature; Tmin, minimum temperature; EC, electric conductivity; BD, bulk density; TN, total nitrogen.

Universal kriging was selected for the spatial prediction of the variables of evaporation, precipitation, pH, potassium, bulk density, and phosphorus because it resulted in less error. On the other hand, ordinary kriging was used for the spatial prediction of the following variables: maximum temperature, minimum temperature, clay, electrical conductivity, sand, and total nitrogen (Table 4).

Ecological niche of Stenocereus queretaroensis

Fourteen variables with significant differences were identified: clay content, sand content, altitude, potassium content, organic matter content, vegetation indices 4 to 9, total nitrogen content, precipitation, and minimum temperature (Table 5). The sand, potassium, organic matter, total nitrogen, and minimum temperature variables were greater at the sites where the species was present. The clay content, altitude, NDVI 4-9, and precipitation were greater in places where the species was not present.

Principal component analysis revealed that the first four components explained more than 70% of the variance in the data (S1). In the first two components, the presence points are located toward the central part, whereas the absence points are associated with the ends of the graph (Figure 2). Additionally, there are points of absence within the clouds of presence points (61, 71, 75) because their values for altitude, NDVI (1, 7, 8, 10), phosphorus, pH, and minimum temperature are close to the means of the presence points. For components three and four, two absence points (67, 69) were within the clouds of presence points, and two presence points (4 and 14) were within the clouds of absence points.

Table 4. Summary of 10-fold cross-validation.

Variable	Kriging	ME	RMSE	MSDR
Ev (mm)	Universal	-3.216	135.749	1.044
Pp (mm)	Universal	7.177	120.260	1.598
Tmax (°C)	Universal	-1.205	98.829	1.133
Tmin (°C)	Ordinary	-1.841	32.471	1.165
Clay (%)	Ordinary	-0.231	15.654	1.184
pH	Universal	0.0002	0.023	0.788
EC (dS/m)	Ordinary	0.005	0.579	0.882
K (mg/kg)	Universal	0.0126	0.994	0.883
Sand (%)	Ordinary	-0.159	12.716	1.359
BD (g/cm^3)	Universal	0.002	0.272	0.766
TN (%)	Ordinary	0.0002	0.195	1.268
OM (%)	Ordinary	-0.009	0.818	1.381
P (mg/kg)	Universal	0.011	2.289	0.945

ME, mean error; RMSE, root mean square error; MSDR, mean square error ratio. Ev, evaporation; Pp, precipitation; Tmax, maximum temperature; Tmin, minimum temperature; EC, electric conductivity; BD, bulk density; TN, total nitrogen.

Table 5. Ranges of environmental predictors of presence and absence points.

Variable	Presence			Absence			p-value
	Minimum	Mean	Maximum	Minimum	Mean	Maximum	
Clay (%)	31.3	40.2	47.1	36.9	42.3	49.3	0.0287
Sand (%)	21.7	33.9	53.2	16.5	27	37.4	<0.001
Aspect (°)	5.12	156.47	344.72	14.55	169.3	351.75	0.609
EC (dS/m)	0.473	0.53	0.62	0.473	0.539	0.609	0.787
BD (g/cm^3)	1.24	1.36	1.52	1.26	1.400	1.52	0.063
Altitude (m)	1532	1621	1881	1505	1816	2181	0.006
Ev (mm)	1711	1922	2165	1480	1813	2181	0.065
K (mg/kg)	0.623	1.08	1.674	0.512	0.906	1.534	0.033
OM (%)	1.72	9.97	19.72	2.16	4.44	9.65	<0.001
ndvi_1	0.102	0.261	0.41	0.167	0.283	0.448	0.13
ndvi_2	0.089	0.211	0.357	0.144	0.221	0.314	0.327
ndvi_3	0.08	0.194	0.29	0.129	0.206	0.317	0.265
ndvi_4	0.059	0.156	0.326	0.131	0.183	0.251	0.005
ndvi_5	0.078	0.155	0.25	0.129	0.194	0.31	0.008
ndvi_6	0.076	0.153	0.237	0.125	0.194	0.338	0.005
ndvi_7	0.077	0.147	0.254	0.126	0.194	0.344	0.003
ndvi_8	0.077	0.155	0.279	0.139	0.188	0.312	0.009
ndvi_9	0.065	0.152	0.294	0.128	0.184	0.276	0.005
ndvi_10	0.088	0.168	0.262	0.128	0.185	0.293	0.133
TN (%)	0.215	0.235	0.252	0.212	0.229	0.244	0.019
P (mg/kg)	20.19	27.07	34.34	15.36	25.8	34.53	0.364
pH	5.75	6.33	6.95	5.76	6.23	6.85	0.255
Pp (mm)	730	788	890	734	821	958	0.043
Tmax (°C)	25.3	26.7	27.9	24.4	26.3	28.2	0.133
Tmin (°C)	8.6	11.1	13.5	7.4	10.5	13.6	0.036

Ev, evaporation; Pp, precipitation; Tmax, maximum temperature ; Tmin, minimum temperature; EC, electric conductivity; BD, bulk density; TN, total nitrogen.; ndvi 1-10, Normalized Difference Vegetation Indices.

The two presence points have values close to the mean of the presence points for the sand, organic matter, and minimum temperature predictors. The two absence points have values close to the mean of the presence points for the NDVI1 and NDVI10 predictors (S2).

Modeling the potential distribution of Stenocereus queretaroensis

The variables used for the distribution model of *S. queretaroensis* were clay, organic matter, evaporation, mean temperature, and NDVI (5, 6, and 8), according to the principal component analysis (Annex 1). The lowest AUC value was obtained with the domain algorithm, whereas the highest was obtained with the random forest algorithm (Table 6). Figure 3 presents the potential distributions predicted by the six algorithms. Green color indicates a greater probability of the species' presence, whereas lighter colors indicate a lower probability. In general, the six algorithms predict that the southern region has the lowest probability of species presence.

Distribution suitability of the species

The potential distribution shows that the highly suitable (HS) and suitable (S) classes are concentrated in the central part of the subbasin, specifically around the periphery of Lake Chapala, as well as in the northeastern and southeastern regions (Figure 4). In these classes, the soil has higher contents of sand and organic matter; evaporation and minimum and maximum temperatures are greater; and precipitation is lower. The relief is characterized by the presence of steep slopes and altitudes between 1532 and 1881 meters.

Table 6. Prediction accuracy of the potential distribution of *S. queretaroensis*.

Algorithm	AUC
Random forest	1.00
Generalized linear models	0.93
Artificial neural network	0.96
Domain	0.89
Support vector machine	0.99
Maxent	0.99

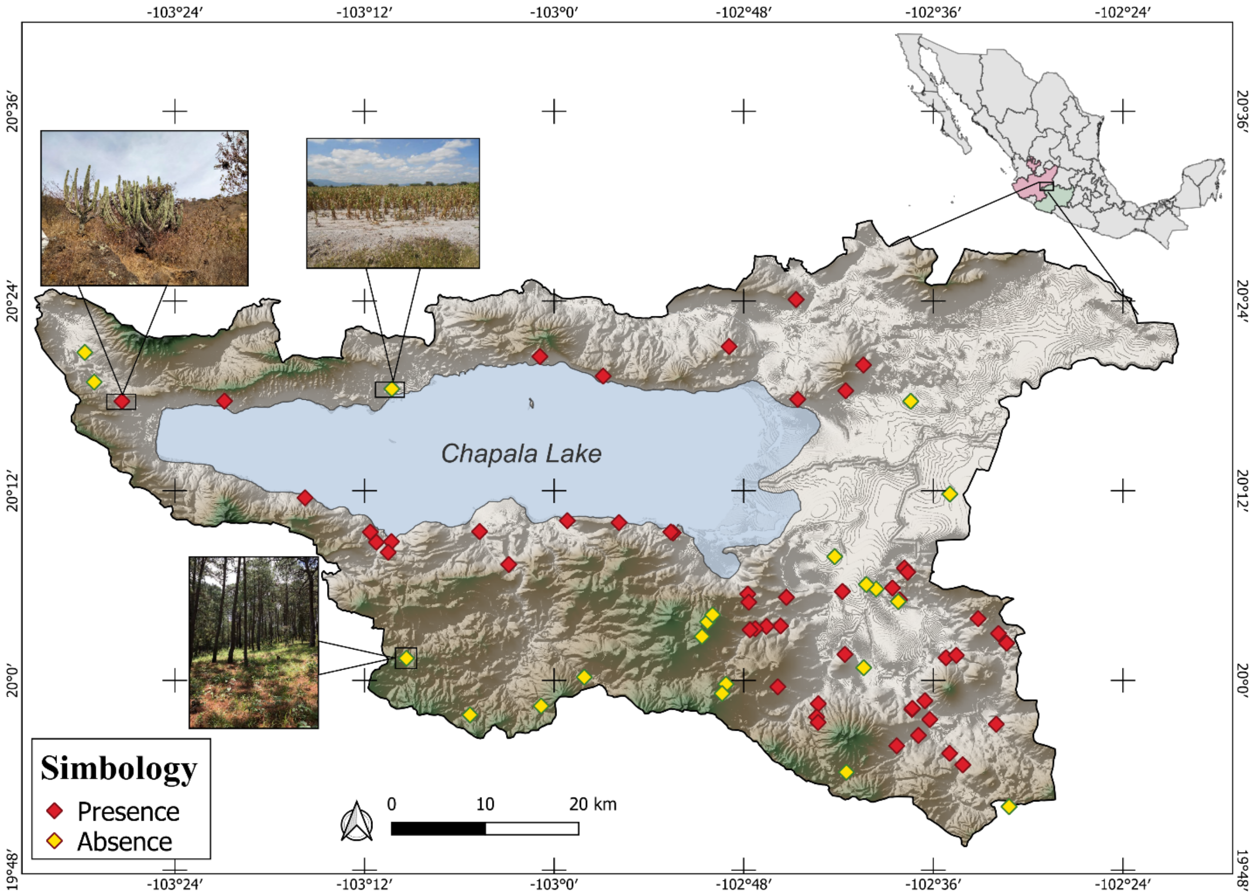

Figure 1. Study area and presence-absence records of *Stenocereus queretaroensis*.

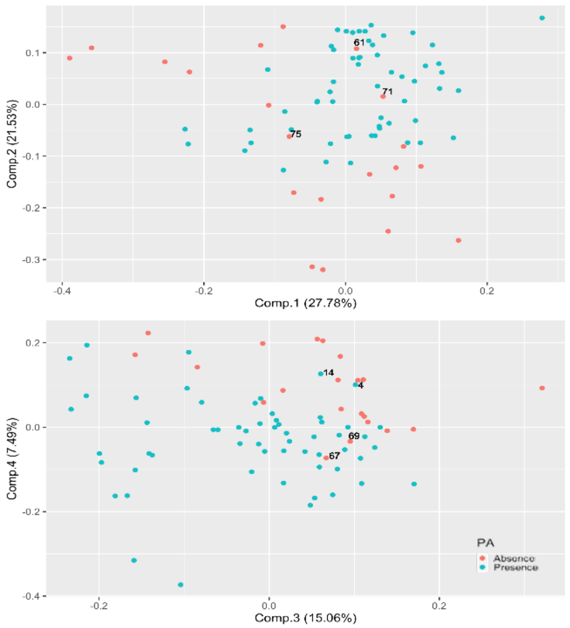

Figure 2. Principal component analysis (PCA) of the variables shows the differences in the environmental niche between the presences and absences of Stenocereus queretaroensis.

The poorly suitable (PS), and very poorly suitable (VPS) classes are found in the southern part of the subbasin (Figure 4). The prevailing climatic conditions include greater precipitation and low minimum and maximum temperatures. The soils are clayey and low in organic matter. The altitude ranges from 1505 to 1532 meters, corresponding to the cultivated lands of Irrigation District 024, Cienega of Chapala, which are flooded and have electrical conductivities greater than 4 dS m^{-1}. The other altitudinal range, between 1881 and 2181 meters, corresponds to the highest part of the subbasin, where oak and pine forests are commonly found.

DISCUSSION

The theoretical models of the fitted semivariograms of the variables EC, pH, and organic matter coincided with some previous works (Lastiri-Hernández et al., 2020; Estrada-Godoy et al., 2023). The ordinary and universal kriging results for the variables EC, evaporation, and precipitation were like those reported by Cadena et al. (2019) and Lastiri-Hernández et al. (2020).

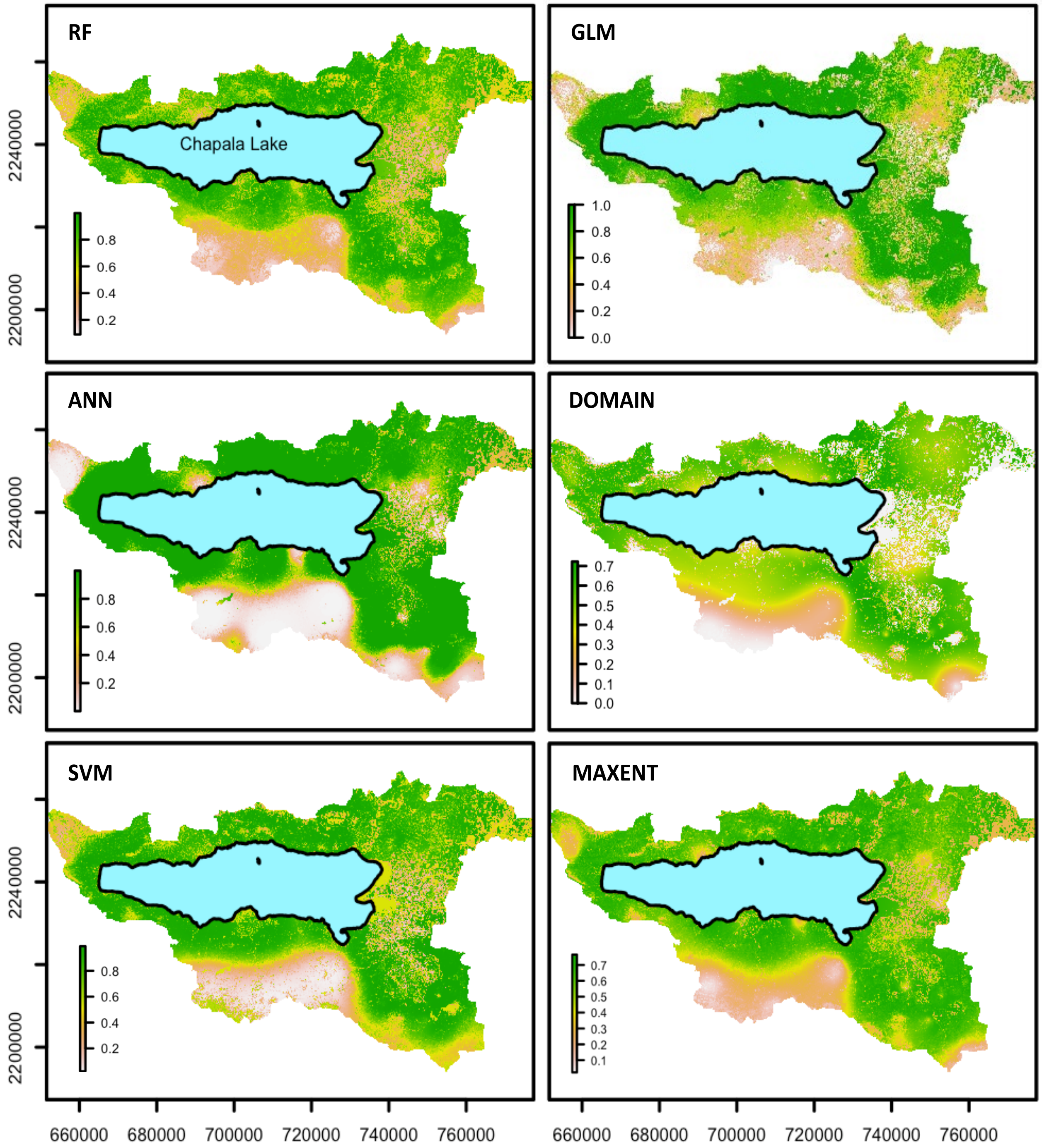

Figure 3. Potential distribution models of *Stenocereus queretaroensis*. RF= random forest; GLM= generalized linear models; ANN= Artificial neural network; SVM= Support vector machine.

The results of the mean comparison test and principal component analysis confirmed that eight variables were the most important for the ecological niche of *S. queretaroensis*. Three soil properties (clay, total nitrogen, and organic matter), two climate parameters (evaporation and minimum temperature), and altitude were measured. This means that 33% of the variables that define the ecological niche of this species are soil properties, as reported by Huerta-Martinez et al. (2020). At the regional level, temperature is the most important factor; however, at the local level, soil properties are predominant (Amaral et al., 2022).

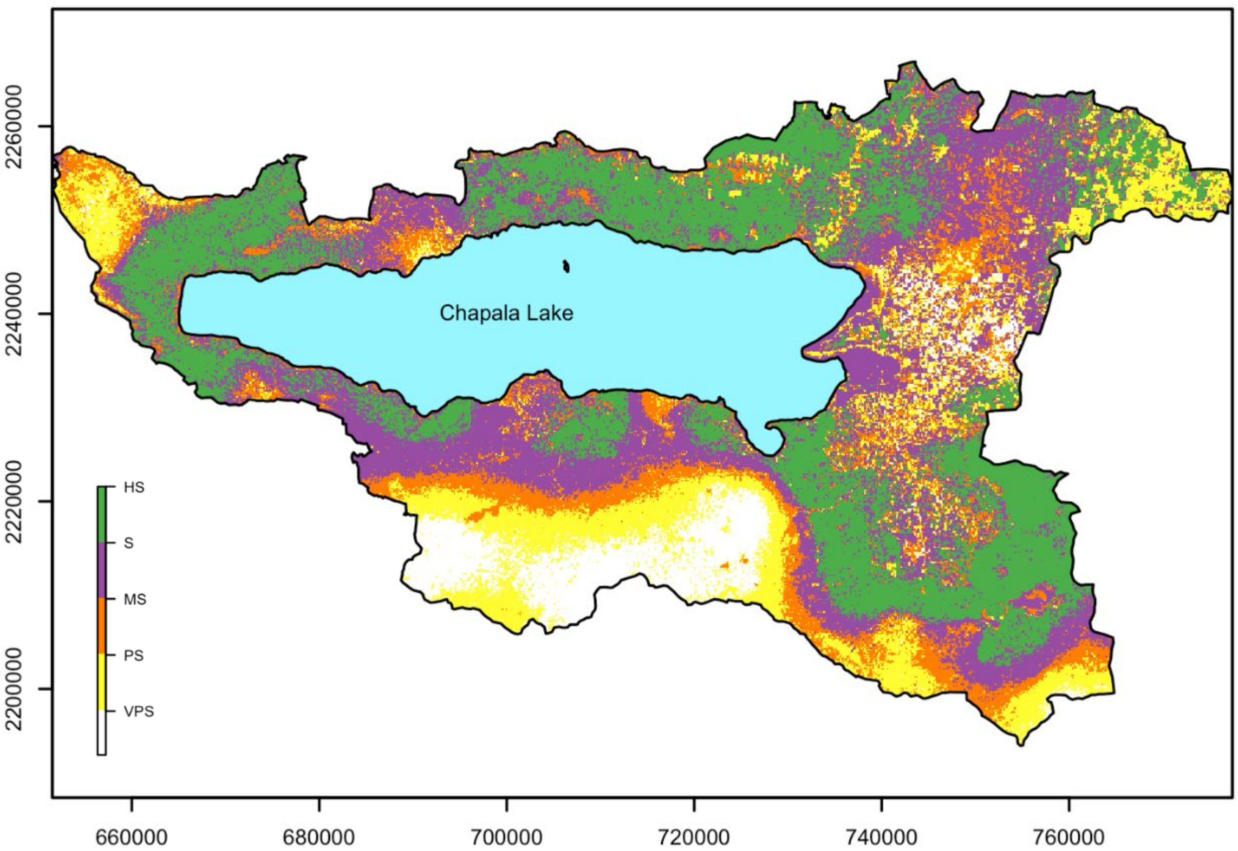

Figure 4. Potential distribution of *Stenocereus queretaroensis* assembled model. Highly suitable (HS, 0.8-1.0); Suitable (S, 0.6-0.8); Marginally suitable (MS, 0.4-0.6); Poorly suitable (PS, 0.2-0.4); Very poorly suitable (VPS, 0.0-0.2).

The suitable areas for the distribution of *S. queretaroensis* are those with Leptosols, which are shallow soils with good drainage. In general, these soils support the greatest distribution and diversity of cacti (Alanís-Rodríguez et al., 2015). Leptosols are stony and have clay contents greater than 35%, which favors an increase in the abundance of species in this family (Ferreira et al., 2016). The altitude is another factor that defines the distribution of *S. queretaroensis*. The higher the altitude, the greater the decrease in the presence of *S. queretaroensis*, which may be due to relatively high humidity, low temperatures, and the presence of tree species that inhibit plant growth (Diniz et al., 2021). The morphology of *S. queretaroensis* is affected at altitudes greater than 1881 meters, as individuals with shorter heights are observed. This contrasts with species of the genus *Neobuxbaumia*, in which fruit size and height are positively affected (Arroyo-Cosultchi et al., 2017).

It is important to define conservation areas for *S. queretaroensis* in its wild state. This will allow for the provision of ecosystem services, the mitigation of climate change, and the conservation of this species by targeting the protection of biodiversity, among other benefits. Additionally, Mexico is one of the countries with high endemism, phylogenetic endemism, and phylogenetic cactus diversity, which makes the designation of protection areas even more important (Amaral et al., 2022). Therefore, areas are proposed where Leptosols are located because these soils have the appropriate soil, topography, and climate properties to support the growth of *S. queretaroensis.* Additionally, areas for cultivation can be defined to provide income for local communities. For example, in the transition areas between Vertisols and Leptosols, where the climate and topography are suitable, soil conditions could be improved with agronomic management. Additionally, areas with Leptosols where rainfed maize is planted are suitable for the cultivation of *S. queretaroensis*.

CONCLUSIONS

Principal component analysis and mean comparison tests revealed that the clay content, total nitrogen content, organic matter content, evaporation, and minimum temperature are the most important variables that affect the ecological niche of *Stenocereus queretaroensis*. The distribution occurs at altitudes ranging from 1532 to 1881 meters. Areas with Leptosols are the most suitable for defining conservation areas for this species.

CREDIT AUTHOR STATEMENT

ETV: Conceptualization, Methodology, Investigation, Writing - Original Draft; GCC: Conceptualization, Methodology, Investigation, Writing - Original Draft; SOE: Conceptualization, Writing - Review & Editing; JTSG: Conceptualization, Writing - Review & Editing; YJCR: Methodology, Writing - Review & Editing; RFM: Methodology, Writing - Review & Editing.

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.

REFERENCES

Acharya, P., Bradar, C., Louhaichi, M., Ghosh, S., Hassan, S., Moyo, H., & Sarker, A. (2019) Finding a suitable niche for cultivating cactus pear (*Opuntia ficus indica*) as an integrated crop in resilient dryland agroecosystems of India. *Sustainability*, 11(21), 5897. https://doi.org/10.3390/su11215897

Alanís-Rodríguez, E., Mora-Olivo, A., Jiménez-Pérez, J., González-Tagle, M. A., Yerena Yamallel, J. I., Martínez-Ávalos, J. G., & González-Rodríguez, L. E. (2015) Composición y diversidad del matorral desértico rosetófilo en dos tipos de suelo en el noreste de México. *Acta Botánica Mexicana*, 110, 105-117. https://doi.org/10.21829/abm110.2015.187

Amaral, D. T., Bonatelli, I. A., Romeiro-Brito, M., Moraes, E. M., & Franco, F. F. (2022) Spatial patterns of evolutionary diversity in *Cactaceae* show low ecological representation within protected areas. *Biological Conservation*, 273, 109677. https://doi.org/10.1016/j.biocon.2022.109677

Arroyo-Cosultchi, G., Arias, S., López-Mata, L., & Terrazas, T. (2017) Morphological plasticity of an endemic widespread columnar cactus and its congener. *Brazilian Journal of Botany*, 40, 1029-1040. https://doi.org/10.1007/s40415-017-0399-7

Brotons, L., Thuiller, W., Araújo, M. B., Brotons, A. H. H., Jo, A., Hirzel, M. B., & Thuiller,/W. (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27, 4. https://doi.org/10.1111/j.0906-7590.2004.03764.x

Cadena R, Y. J. Vázquez-Sánchez, M., Cruz-Cárdenas, G., & Villaseñor, J. L. (2019) Use of ecological niche models of plant species to optimize placement of apiaries. *Journal of Apicultural Science*, 63(2), 243-265. https://doi.org/10.2478/jas-2019-0017

Campos-Rojas, E., Pinedo- Espinoza, J. M., Campos-Montiel, R. G., & Hernández-Fuentes, A. D. (2011) Evaluación de plantas de pitaya (S*tenocereus spp*) de poblaciones naturales de Monte Escobedo, Zacatecas. *Revista Chapingo Serie Horticultura*, 17(3), 173–182. https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1027-152X2011000300009

Cavalcante, A. de M. B., & Sampaio, A. C. P. (2022) Modeling the potential distribution of cacti under climate change scenarios in the largest tropical dry forest region in South America. *Journal of Arid Environments*, 200. https://doi.org/10.1016/j.jaridenv.2022.104725

Cruz-Cárdenas, G., López-Mata, L., Villaseñor, J. L., & Ortiz, E. (2014) Potential species distribution modeling and the use of principal component analysis as predictor variables. *Revista Mexicana de Biodiversidad*, 85(1), 189-199. https://doi.org/10.7550/rmb.36723

Diniz, F. C., Ramos, M. B., Almeida, H. A. D., Pinto, A. S., & Lopes, S. D. F. (2021) Effects of topographic factors on distribution of cacti along an elevation gradient in Brazilian Caatinga. *Rodriguésia*, 72, e02222018. https://doi.org/10.1590/2175-7860202172129

Estrada-Godoy, F., Cruz-Cárdenas, G., Ochoa-Estrada, S., & Silva, J. T. (2023) Cartografía digital de suelos con regresión-Kriging y datos de sensores remotos. *Terra Latinoamericana*, 41. https://doi.org/10.28940/terra.v41i0.1617

Ferreira, P. S. M., Lopes, S. D. F., & Trovão, D. M. D. B. M. (2016) Patterns of species richness and abundance among cactus communities receiving different rainfall levels in the semiarid region of Brazil. *Acta Botanica Brasilica*, 30, 569-576. https://doi.org/10.1590/0102-33062016abb0084

Flores, G. J. A., & Macías, C. G. V. (2008). Importancia de las cactáceas como recurso natural en el noreste de México. *Ciencia-UANL*, 11, 5-11. https://www.redalyc.org/pdf/402/40211102.pdf

García-Cruz, L., Valle-Guadarrama, S., Guerra-Ramírez, D., Martínez-Damián, M. T., & Zuleta-Prada, H. (2022) Cultivation, quality attributes, postharvest behavior, bioactive compounds, and uses of *Stenocereus*: A review. *Scientia Horticulturae*, 304. https://doi.org/10.1016/j.scienta.2022.111336

García, & Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), (2008). "*Climas (clasificación de Koppen, modificado por García). Escala 1:1000000*". México.

Gašparovičová, P., Ševčík, M., & David, S. (2022) The Prediction of Distribution of the Invasive *Fallopia* Taxa in Slovakia. *Plants*, 11(11). https://doi.org/10.3390/plants11111484

Goettsch, B., Durán, A. P., & Gaston, K. J. (2019) Global gap analysis of cactus species and priority sites for their conservation. *Conservation Biology*, 33(2), 369-376

Hu, J., Broennimann, O., Guisan, A., Wang, B., Huang, Y., & Jiang, J. (2016) Niche conservatism in Gynandropaa frogs on the southeastern Qinghai-Tibetan Plateau. *Scientific Reports*, 6(1), 32624. https://doi.org/10.1038/srep32624

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22: 415-427. https//doi.org/10.1101/SQB.1957.022.01.039

Huerta-Martínez, F. M., García-Moya, E., Flores-Flores, J. L., & Pimienta Barrios, E. (1999) Ordenación de las poblaciones silvestres de pitayo y cardón en la Cuenca de Sayula, Jalisco. *Boletin de la Sociedad Botánica de México*, 1(64), 11–24. https://doi.org/10.17129/botsci.1578

Huerta-Martinez, F., Neri-Luna, C., Muñoz-Urias, A., Castruita-Dominguez, J. P., & Sahagún-Sánchez, F. J. (2020). Assessing the Relative Importance of Climate and Soil for Vegetation Patterns in a Semiarid Land of Central Mexico. *Egyptian Journal of Botany*, 60(3), 825-835. https://doi.org/10.21608/ejbo.2020.18673.1374

Instituto Nacional de Estadística, Geografía e Informática-INEGI (2014) "*Conjunto de datos vectoriales edafológico, escala 1:250000 Serie II. (Continuo Nacional)', escala: 1:250000*". edición: 2. Aguascalientes, Ags., México.

Instituto Nacional de Estadística, Geografía e Informática-INEGI (2021) "*Conjunto de Datos Vectoriales de Uso de Suelo y Vegetación. Escala 1:250 000, Serie VII*". Conjunto Nacional.', escala: 1:250 000. edición: 1. Aguascalientes, México.

Illoldi-Rangel, P., Ciarleglio, M., Sheinvar, L., Linaje, M., Sánchez-Cordero, V., & Sarkar, S. (2012) *Opuntia* in méxico: Identifying priority areas for conserving biodiversity in a multiuse landscape. *PLoS ONE*, 7(5). https://doi.org/10.1371/journal.pone.0036650

Kozak, A., & Kozak, R. (2003) Does cross-validation provide additional information in the evaluation of regression models? *Canadian Journal of Forest Research*, 33(6), 976-987.

Kravchenko, A., & Bullock, D. G. (1999) A comparative study of interpolation methods for mapping soil properties. *Agronomy Journal*, 91(3), 393–400. https://doi.org/10.2134/agronj1999.00021962009100030007x

Lastiri-Hernández, M. A., Cruz-Cárdenas, G., Álvarez-Bernal, D., Vázquez-Sánchez, M., & Bermúdez-Torres, K. (2020) Ecological niche modeling for halophyte species with possible anthropogenic use in agricultural saline soils. *Environmental Modeling & Assessment*, 25, 429-440. https://doi.org/10.1007/s10666-020-09690-1

Mandal, S., Choudhury, B. U., & Satpati, L. (2020) Soil site suitability analysis using geo-statistical and visualization techniques for selected winter crops in Sagar Island, India. *Applied Geography*, 122. https://doi.org/10.1016/j.apgeog.2020.102249

QGIS.org. (2023) "*QGIS Geographic Information System*". QGIS Association. http://www.qgis.org

R Core Team (2023) "*R: A language and environment for statistical computing*". R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ranjan, P., Ranjan, J. K., Misra, R. L., Dutta, M., & Singh, B. (2016) Cacti: notes on their uses and potential for climate change mitigation. *Genetic Resources and Crop Evolution*, 63, 901-917. https://doi.org/10.1007/s10722-016-0394-z

Ruán-Tejeda, I., Santerre, A., Huerta-Martínez, F. M., Iñiguez-Dávalos, L. I., & Castro-Félix, P. (2014). Genetic diversity and relationships among wild and cultivated *Stenocereus queretaroensis* populations in western Mexico. *Biochemical Systematics and Ecology*, 55, 125-130. https://doi.org/10.1016/j.bse.2014.03.005

SEMARNAT-Secretaría de Medio Ambiente y Recursos Naturales (2002) "*Norma Oficial Mexicana NOM-021-RECNAT-2000. Que establece las especificaciones de fertilidad, salinidad y clasificación de suelos. Estudios, muestreo y análisis*". In Diario Oficial de la Federación. SEGOB. http://dof.gob.mx/nota_detalle.php?codigo=717582&fecha=31/12/2002

Shapiro, S. S., & Wilk, M. B. (1965) An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591-611.

Václavík, T., & Meentemeyer, R. K. (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling*, 220(23), 3248–3258. https://doi.org/10.1016/j.ecolmodel.2009.08.013

Vicente-Silva, L., Faggioni, G. P., & Paggi, G. M. (2022) Predicting potential distribution and evaluating biotic interactions of threatened species: a case study of *Discocactus ferricola* (Cactaceae). *Biota Neotropica*, 22(1). https://doi.org/10.1590/1676-0611-BN-2020-1145

Wolf, B. O., & Rio, C. M. D. (2003) How important are columnar cacti as sources of water and nutrients for desert consumers? A review. *Isotopes in Environmental and Health Studies*, 39(1), 53-67. https://doi.org/10.1080/1025601031000102198.