

Predicting *Pinus monophylla* forest cover in the Baja California Desert by remote sensing

Jonathan G. Escobar-Flores¹, Carlos A. Lopez-Sanchez², Sarahi Sandoval³, Marco A. Marquez-Linares¹ and Christian Wehenkel²

¹ Centro Interdisciplinario De Investigación para el Desarrollo Integral Regional, Unidad Durango, Instituto Politécnico Nacional, Durango, Durango, México

² Instituto de Silvicultura e Industria de la Madera, Universidad Juárez del Estado de Durango, Durango, Mexico

³ CONACYT—Instituto Politécnico Nacional, CIIDIR Unidad Durango, Durango, Durango, México

ABSTRACT

The Californian single-leaf pinyon (*Pinus monophylla* var. *californiarum*), a subspecies of the single-leaf pinyon (the world's only one-needled pine), inhabits semi-arid zones of the Mojave Desert (southern Nevada and southeastern California, US) and also of northern Baja California (Mexico). This tree is distributed as a relict subspecies, at elevations of between 1,010 and 1,631 m in the geographically isolated arid Sierra La Asamblea, an area characterized by mean annual precipitation levels of between 184 and 288 mm. The aim of this research was (i) to estimate the distribution of *P. monophylla* var. *californiarum* in Sierra La Asamblea by using Sentinel-2 images, and (ii) to test and describe the relationship between the distribution of *P. monophylla* and five topographic and 18 climate variables. We hypothesized that (i) Sentinel-2 images can be used to predict the *P. monophylla* distribution in the study site due to the finer resolution ($\times 3$) and greater number of bands ($\times 2$) relative to Landsat-8 data, which is publically available free of charge and has been demonstrated to be useful for estimating forest cover, and (ii) the topographical variables aspect, ruggedness and slope are particularly important because they represent important microhabitat factors that can determine the sites where conifers can become established and persist. An atmospherically corrected a 12-bit Sentinel-2A MSI image with 10 spectral bands in the visible, near infrared, and short-wave infrared light region was used in combination with the normalized differential vegetation index (NDVI). Supervised classification of this image was carried out using a backpropagation-type artificial neural network algorithm. Stepwise multiple linear binominal logistical regression and Random Forest classification including cross validation were used to model the associations between presence/absence of *P. monophylla* and the five topographical and 18 climate variables. Using supervised classification of Sentinel-2 satellite images, we estimated that *P. monophylla* covers $6,653 \pm 319$ ha in the isolated Sierra La Asamblea. The NDVI was one of the variables that contributed most to the prediction and clearly separated the forest cover ($\text{NDVI} > 0.35$) from the other vegetation cover ($\text{NDVI} < 0.20$). Ruggedness was the most influential environmental predictor variable, indicating that the probability of occurrence of *P. monophylla* was greater than 50% when the degree of ruggedness terrain ruggedness index was greater than 17.5 m.

Submitted 29 November 2017

Accepted 21 March 2018

Published 4 April 2018

Corresponding author

Christian Wehenkel,
wehenkel@ujed.mx

Academic editor

Robert Winkler

Additional Information and
Declarations can be found on
page 13

DOI 10.7717/peerj.4603

© Copyright

2018 Escobar-Flores et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

The probability of occurrence of the species decreased when the mean temperature in the warmest month increased from 23.5 to 25.2 °C. Ruggedness is known to create microclimates and provides shade that minimizes evapotranspiration from pines in desert environments. Identification of the *P. monophylla* stands in Sierra La Asamblea as the most southern populations represents an opportunity for research on climatic tolerance and community responses to climate variability and change.

Subjects Ecology, Plant Science

Keywords DEM, Sentinel-2, Ruggedness, Remote sensing, Neural net, Forest, Baja California, NDVI, Kappa, Classification

INTRODUCTION

The Californian single-leaf pinyon (*Pinus monophylla* var. *californiarum*), a subspecies of the single-leaf pinyon (the world's only one-needled pine), inhabits semi-arid zones of the Mojave Desert (southern Nevada and southeastern California, US) and also of northern Baja California (BC) (Mexico). It is both cold-tolerant and drought-resistant and is mainly differentiated from the typical subspecies *P. monophylla* var. *monophylla* by a larger number of leaf resin canals and longer fascicle-sheath scales (Bailey, 1987). This subspecies was first reported in BC in 1767 (Bullock & Heath, 2006). The southernmost record of *P. monophylla* var. *californiarum* in America was previously in BC, 26–30 miles north of Punta Prieta, at an elevation of 1,280 m (longitude $-114^{\circ}.155$; latitude $29^{\circ}.070$, catalogue number ASU 0000235), and the type specimen is held in the Arizona State University Vascular Plant Herbarium.

This species is distributed as a relict subspecies in the geographically isolated Sierra La Asamblea, at a distance of 196 km from the Southern end of the Sierra San Pedro Martir and at elevations of between 1,010 and 1,631 m (Moran, 1983) in areas with mean annual precipitation levels of between 184 and 288 mm (Roberts & Ezcurra, 2012). The Californian single-leaf pinyon grows together with up to about 86 endemic plant species, although the number of species decreases from north to south (Bullock et al., 2008).

Adaptation of *P. monophylla* var. *californiarum* to arid ecosystems enables the species to survive annual precipitation levels of less than 150 mm. In fact, seeds of this variety survive well under shrubs such as *Quercus* spp. and *Arctostaphylos* spp., a strategy that enables the pines to widen their distribution, as has occurred in the great basin in California (Callaway et al., 1996; Chambers, 2001), and for them to occupy desert zones such as Sierra de la Asamblea. Despite the importance of this relict pine species, its existence is not considered in most forest inventories in Mexico (CONABIO, 2017).

Remote sensing with Landsat images has been demonstrated to be useful for estimating forest cover; the Landsat-8 satellite has sensors (seven bands) that can be predicted vegetation attributes at a spatial resolution of 30 m (Madonsela et al., 2017). However, the European Space Agency's Copernicus program has made Sentinel-2 satellite images available to the public free of charge. The spatial resolution (10 m is pixel) of the images is three times finer than that of Landsat images, thus increasing their potential for predicting and differentiating types of vegetation cover (Drusch et al., 2012; Borràs et al., 2017).

The Sentinel-2 has 13 bands, of which 10 provide greater-quality radiometric images of spatial resolution 10–20 m in the visible and infrared regions of the electromagnetic spectrum. These images are therefore ideal for land classification (ESA, 2017).

The aim of this research was (i) to estimate the distribution of *P. monophylla* var. *californiarum* in Sierra La Asamblea, Baja California (Mexico) by using Sentinel-2 images, and (ii) to test and describe the relationship between this distribution of *P. monophylla* and five topographic and 18 climate variables. We hypothesized that (i) the Sentinel-2 images can be used to accurately predict the *P. monophylla* distribution in the study site due to finer resolution ($\times 3$) and greater number of bands ($\times 2$) than in Landsat-8 data, and (ii) the topographical variables aspect, ruggedness and slope are particularly influential because they represent important microhabitat factors that can determine where conifers can become established and persist (Marston, 2010).

MATERIALS AND METHODS

Study area

Sierra La Asamblea is located in Baja California's central desert ($-114^{\circ}9'W$ $29^{\circ}19'N$, elevation range 280–1,662 m, Fig. 1). The climate in the area is arid, with maximum temperatures of $40^{\circ}C$ in the summer (García, 1998). The sierra is steeper on the western slopes, with an average incline of 35° , and with numerous canyons with occasional springs and oases. Valleys and plateaus are common in the proximity of the Gulf of California. Granite rocks occur south of the sierra and meta-sedimentary rocks along the north and southeast of the slopes. The predominant type of vegetation is xerophilous scrub, which is distributed at elevations ranging from 200 to 1,000 m. Chaparral begins at an altitude of 800 m, and representative specimens of *Adenostoma fasciculatum*, *Ambrosia ambrosioides*, *Dalea bicolor orcuttiana*, *Quercus tuberculata*, *Juniperus californica*, and *P. monophylla* are also present at elevations above 1,000 m. Populations of the endemic palm tree *Brahea armata* also occur in the lower parts of the canyons with superficial water flow and through the rocky granite slopes (Bullock & Heath, 2006).

Datasets

Sentinel-2

The Sentinel-2A multispectral instrument (MSI) L1C dataset, acquired on October 11, 2016, in the trajectory of coordinates latitude $29^{\circ}.814$, longitude $114^{\circ}.93$, was downloaded from the US Geological Survey (USGS) Global Visualization Viewer at <http://glovis.usgs.gov/>. The 12-bit Sentinel-2A MSI image has 13 spectral bands in the visible, NIR, and SWIR wavelength regions with spatial resolutions of 10–60 m. However, band one, used for studies of coastal aerosols, and bands nine and 10, applied for, respectively, water vapor correction and cirrus detection, were not used in this study (ESA, 2017). Hence, the data preparation involved four bands at 10 m and the resampling of the six S2 bands acquired at 20 m to obtain a layer stack of 10 spectral bands at 10 m (Table 1) using the ESA's Sentinel-toolbox ESA Sentinel Application Platform (SNAP) and then converted to ENVI format.

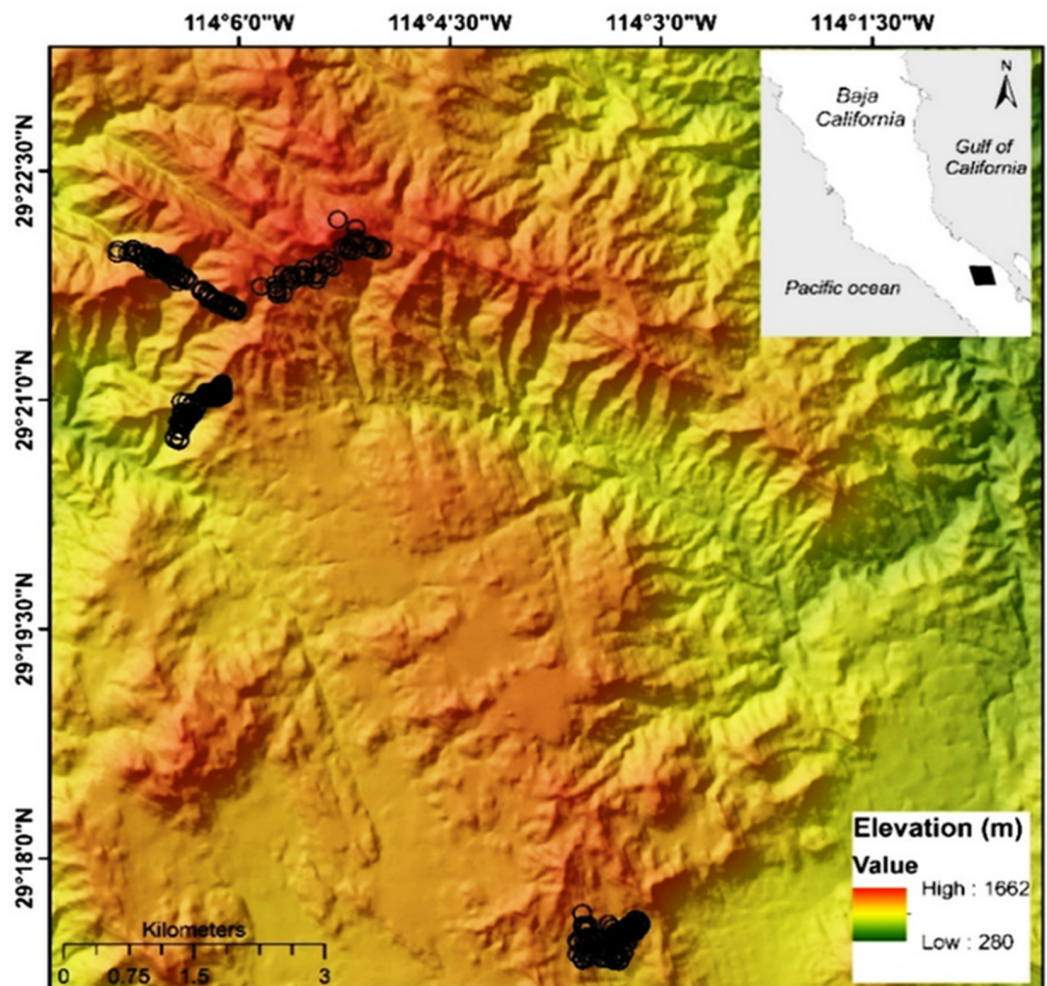


Figure 1 Map of Sierra La Asamblea. The black circles indicate georeferenced sites occupied by *Pinus monophylla*. [Full-size](#) DOI: 10.7717/peerj.4603/fig-1

Because atmospherically improved images are essential to enable assessment of spectral indices with spatial reliability and product comparison, Level-1C data were converted to Level-2A (bottom of atmosphere reflectance) by taking into account the effects of aerosols and water vapor on reflectance (Radoux *et al.*, 2016). The corrections were made using the Sen2Cor tool (Telespazio VEGA Deutschland GmbH, 2016) for Sentinel-2 images.

The following equation was used to calculate the normalized difference vegetation index (NDVI): $NDVI = (NIR - R) / (NIR + R)$, where NIR is the near infrared light (band) reflected by the vegetation, and R is the visible red light reflected by the vegetation (Rouse *et al.*, 1974). The NDVI is useful for discriminating the layers of temperate forest from scrub and chaparral. Areas occupied by large amounts of unstressed green vegetation will have values much greater than 0 and areas with no vegetation will have values close to 0 and, in some cases, negative values (Pettorelli, 2013). The NDVI image was combined with the previously described multi spectral bands.

Table 1 Sentinel-2 spectral bands used to predict the *Pinus monophylla* forest cover.

Band	Central wavelength (μm)	Resolution (m)
Band 2—Blue	0.490	10
Band 3—Green	0.560	10
Band 4—Red	0.665	10
Band 5—Vegetation red edge	0.705	20
Band 6—Vegetation red edge	0.740	20
Band 7—Vegetation red edge	0.783	20
Band 8—NIR	0.842	10
Band 8A—Vegetation red edge	0.865	20
Band 11—SWIR	1.610	20
Band 12—SWIR	2.190	20

Environmental variables

Tree species distribution is generally modulated by hydroclimate and topographical variables (Elliott et al., 2005; DeCastilho et al., 2006), which can be estimated from digital terrain models (DTM) (Osem et al., 2009; Spasojevic et al., 2016). A DTM was obtained by using tools available from the Instituto Nacional de Estadística y Geografía (<http://www.inegi.org.mx/geo/contenidos/datosrelieve>) with a spatial resolution of 15 m. The DTM was processed with the QGIS (QGIS Development Team, 2016), using *Terrain analysis* tools, elevation, slope and aspect (Table 2).

The ruggedness was estimated using two indexes: (i) the terrain ruggedness index (TRI) of Riley, Degloria & Elliot (1999) and (ii) a vector ruggedness measure (VRM), both implemented in QGIS (QGIS Development Team, 2016). The TRI computes the values for each grid cell of a DEM. This calculates the sum change in elevation between a grid cell and its eight-neighbor grid cell. VRM incorporates the heterogeneity of both slope and aspect. This measure of ruggedness uses three-dimensional dispersion of vectors normal to planar facets on landscape. This index lacks units and ranges from 0 (indicating a totally flat area) to 1 (indicating maximum ruggedness) (Sappington, Longshore & Thompson, 2007).

In addition, 18 climate variables with a 30-arc second resolution (approximate 800 m) (Table 2) were obtained from a national database managed by the University of Idaho (<http://charcoal.cnre.vt.edu/climate>) and which requires point coordinates (latitude, longitude and elevation) as the main inputs (Rehfeldt, 2006; Rehfeldt et al., 2006). These variables are frequently used to study the potential effects of global warming on forests and plants in Western North America and Mexico (Sáenz-Romero et al., 2010; Silva-Flores, Pérez-Verdín & Wehenkel, 2014).

Pixel-based classification

Classification method

Pixel-based classification was carried out in order to predict four different types of land cover in the study area (*P. monophylla*, scrub, chaparral, and no apparent vegetation). A supervised classification approach with a backpropagation-type artificial neural network (BPNN) (Tan & Smeins, 1996) was applied. BPNN is widely used because of

Table 2 Topographical and climatic variables considered in the study.

Variable	Abbreviation	Units	Mean	SD	Max	Min
Terrain ruggedness index	TRI	m	20.33	6.66	35.90	4.69
Vector ruggedness measure	VRM	NA	0.005	0.007	0.13	0
Slope	S	°	28.38	8.92	48.34	3.42
Aspect*	A	°	190.51	68.72	350.44	20.55
Elevation*	E	m	1,302.41	124.96	1,631	1,010
Mean annual temperature*	MAT	°C	16.57	0.38	17.4	15.5
Mean annual precipitation*	MAP	mm	229.56	19.95	288	184
Growing season precipitation, April–September*	GSP	mm	79.08	9.60	108	57
Mean temperature in the coldest month*	MTCM	°C	10.85	0.37	11.7	9.8
Minimum temperature in the coldest month*	MMIN	°C	3.42	0.41	4.3	2.3
Mean temperature in the warmest month	MTWM	°C	24.52	0.31	25.2	23.5
Maximum temperature in the warmest month	MMAX	°C	34.10	0.31	34.7	33.1
Julian date of the last freezing data of spring*	SDAY	Days	82.57	7.86	106	60
Julian date of the first freezing data of autumn*	FDAY	Days	331.28	2.62	339	324
Length of the frost-free period*	FFP	Days	259.22	8.36	285	240
Degree days >5 °C*	DD5	Days	4,245.26	137.52	4,550	3,852
Degree days >5 °C accumulating within the frost-free period*	GSDD5	Days	3,491.82	164.76	3,944	2,995
Julian date when the sum degree days >5 °C reaches 100*	D100	Days	17.07	1.10	20	15
Degree days <0 °C*	DD0	Days	0	0	0	0
Minimum degree days <0 °C*	MMINDD0	Days	8.07	20.29	145	45
Spring precipitation	SPRP	mm	7.54	0.71	10	6
Summer precipitation*	SMRP	mm	43.74	6.29	62	29
Winter precipitation*	WINP	mm	110.93	7.93	133	93

Note:

* Variables for which no significant difference between the medians was obtained after Bonferroni correction ($\alpha = 0.0005$) were excluded from further analysis.

its structural simplicity and robustness in modeling non-linear relationships. In this study, the BPNN comprises a set of three layers (raster): an input layer, a hidden layer, and an output layer (*Richards, 1999*). Each layer consists of a series of parallel processing elements (neurons or nodes). Each node in a layer is linked to all nodes in the next layer (*Guo et al., 2013*).

The first step in BPNN supervised classification is to enter the input layer, which in this study corresponded to the values of the pixels of 10 Sentinel-2 bands and of the NDVI image. Weights were then assigned to the BPNN to produce analytical predictions from the input values. These data were contrasted with the category to which each training pixel belongs, corresponding to Georeferenced sites (Datum WGS-84, 11N) obtained in the field in October 2014 and October 2015.

A stratified random sampling method (*Olofsson et al., 2013*) was used to generate the reference data in QGIS software (*QGIS Development Team, 2016*). A total of 2,143 random points were sampled, with at least 400 points for each class (*Goodchild, 1994*). The following classes were considered: (i) *P. monophylla*, 536 sites, (ii) scrub, 764 sites, (iii) chaparral, 405 sites, and (iv) no apparent vegetation, 438 sites (*INEGI, 2013*).

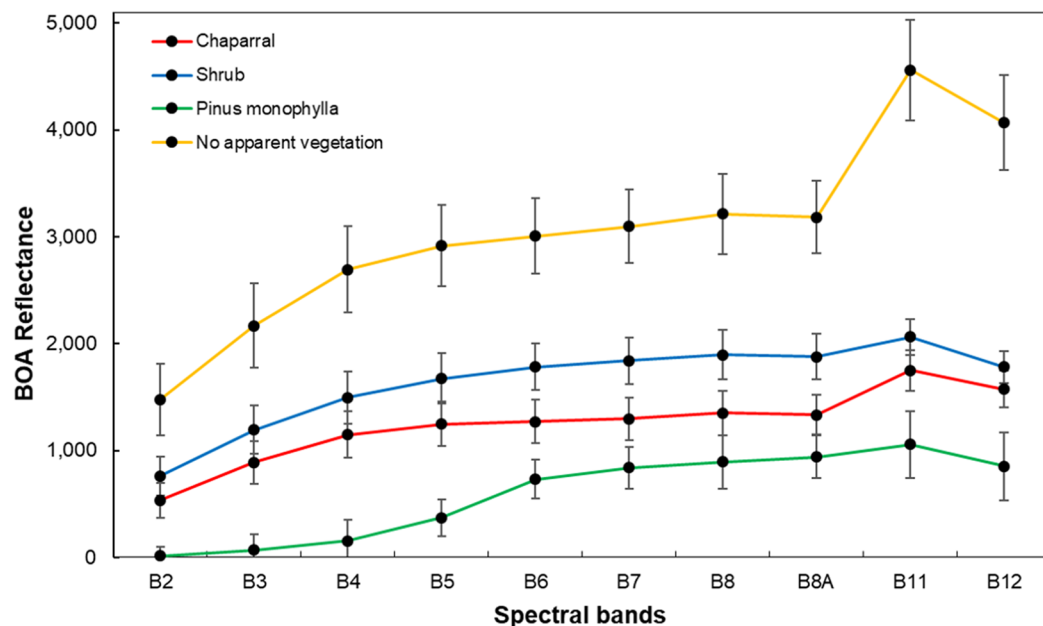


Figure 2 Average spectral signatures of cover vegetation in Sierra La Asamblea, Baja California.

Full-size DOI: 10.7717/peerj.4603/fig-2

Class discrimination processes occurred in the hidden layer and the synapses between the layers were estimated by an activation function. We used a logistic function and training rate of 0.20, previously applied to land cover classification (Hepner *et al.*, 1990; Richards, 1999; Braspenning & Thuijsman, 1995). Learning occurs by adjusting the weights in the node to minimize the difference between the output node activation, and BPNN then calculates the error at each iteration with root mean square (RMS) error. The output layer comprised four neurons representing the four target classes of land cover (*P. monophylla*, Scrub, Chaparral, and no apparent vegetation). Average spectral signatures for the four different types of land cover are shown in Fig. 2.

Validation

The BPNN classification was cross-validated (10-fold) using a confusion matrix, which is a table that compares the reference data and the classification results. We estimated the uncertainty of the classification using estimated error matrix in terms of proportion of area and estimates of overall map accuracy (\hat{O}), user's accuracy (\hat{U}_i) (or commission error), and producer's accuracy (\hat{P}_j) (or omission error) recommended by Olofsson *et al.* (2013): p_{ij} is defined as a cell entry of error matrix of i map classes. A poststratified estimator of p_{ij} is:

$$\hat{p}_{ij} = W_i \frac{n_{ij}}{n_i} \quad (1)$$

where W_i is the proportion of the area mapped as class i . n_i is the total number of sample units in map class i . n_{ij} is the sample count at cell (i, j) in the error matrix.

$\hat{p}_{.j}$ is a poststratified estimator for simple random and systematic sampling:

$$\hat{p}_{.j} = \sum_{i=1}^q W_i \frac{n_{ij}}{n_i} \quad (2)$$

where q is the class number.

An unbiased estimator of the total area of class j is then

$$\hat{A}_j = A \cdot \hat{p}_{.j} \quad (3)$$

where A is the total map area. For $\hat{p}_{.j}$, the standard error is estimated by (Cochran, 1977):

$$S(\hat{p}_{.j}) = \sqrt{\sum_{i=1}^q W_i^2 \frac{\frac{n_{ij}}{n_i} \left(1 - \frac{n_{ij}}{n_i}\right)}{n_i - 1}} \quad (4)$$

The standard error of the error-adjusted estimated area is

$$S(\hat{A}_j) = A \cdot S(\hat{p}_{.j}) \quad (5)$$

Finally,

$$\hat{A}_j \pm 1.96 \cdot S(\hat{A}_j) \quad (6)$$

presents an approximate 95% confidence interval.

The \hat{O} , \hat{U}_i , and \hat{P}_j were calculated with Eqs. (7)–(9) (Congalton, 1991). \hat{U}_i of class i is the proportion of the area mapped as class i that has reference class i . \hat{P}_j of class j is the proportion of the area of reference class j that is mapped as class j .

$$\hat{O} = \sum_{j=1}^q \hat{p}_{jj} \quad (7)$$

$$\hat{U}_i = \frac{\hat{p}_{ii}}{\hat{p}_{.i}} \quad (8)$$

$$\hat{P}_j = \frac{\hat{p}_{jj}}{\hat{p}_{.j}} \quad (9)$$

We then generated a map from the results of the probability of class assignment. The accuracy of classification was also estimated using the Kappa (K) coefficient. The K coefficient is often used as an overall measure of accuracy (Abraira, 2001). This coefficient takes values of between 0 and 1, where values close to 1 indicate a greater degree of agreement between classes and observations, and a value of 0 suggests that the observed agreement is random. However, the use of K is controversial because (i) K would underestimate the probability that a randomly selected pixel is correctly classified, (ii) K is greater correlated with overall accuracy so reporting Kappa is redundant for overall accuracy (Olofsson et al., 2014).

Relationship between presence of *P. monophylla* and environmental variables

To model and test the association between presence/absence of *P. monophylla* in the study area and topographical or climate variables, a Kruskal–Wallis test was used to estimate the difference in the median values in relation to presence and absence of *P. monophylla*. All variables for which no significant difference between the median values was predicted after Bonferroni correction ($\alpha = 0.0005$) were excluded from further analysis. The collinearity between the variables with a significant difference between the medians of presence and absence was estimated using the Spearman correlation coefficient (r_s). When the r_s value for the difference between two variables was greater than 0.7, only the variable with the lowest p value in the Kruskal–Wallis test was used in the models (as reported by [Salas, Valdez & Michel, 2017](#) and [Shirk et al., 2018](#)). Finally, stepwise multiple linear binominal logistical regression and Random Forest classification including cross valuation (10-fold) were used to model the associations between presence/absence of *P. monophylla* and the most important topographical and climate variables ([Shirk et al., 2018](#)).

Regression and classification including cross-validations were carried out using the `trainControl`, `train`, `glm` (family = “binomial”), and `rf` functions, as well as the “`randomForest`” and “`caret`” packages ([Venables & Ripley, 2002](#)) in R (version 3.3.2) ([R Development Core Team, 2017](#)). The goodness-of-fit of the logistical regression model was evaluated using the Akaike information criterion, root mean square error, and residual deviance. Validation of the `randomForest` model was performed using under the curve (AUC; [Fawcett, 2006](#)), true skill statistic (TSS; [Allouche, Tsoar & Kadmon, 2006](#)), Kappa ([Abraira, 2001](#)), specificity, and sensitivity.

RESULTS

Pixel-based classification

We estimated the area of *P. monophylla* cover with a margin of error (at approximate 95% confidence interval) of $6,653 \pm 319$ (standard error) ha in Sierra de la Asamblea, Baja California, Mexico ([Fig. 3](#)). The confusion matrix of the accuracy assessment is listed in [Table 3](#) including user’s and producer’s accuracy for each class. The supervised classification with BPNN yielded predictions with an overall accuracy of identification of 87.74% ([Table 4](#)). This level of accuracy was estimated in the 32 interactions with 0.04 RMS training. The proportion of omission errors in the *P. monophylla* class was only 2.62%, i.e., 97.39% of the pixels were correctly classified. The shrub class had the larger proportion of omission errors (18.98%). The value of NDVI in the *P. monophylla* forest fluctuated between 0.30 and 0.41, and in chaparral between 0.24 and 0.28. The smallest values of NDVI corresponded to scrub vegetation, with values between 0.10 and 0.15.

Relationship between presence of *P. monophylla* and environmental variables

The Kruskal–Wallis test indicated that the median values for ruggedness TRI ($p < 2.1 \times 10^{-16}$), slope ($p < 2.2 \times 10^{-16}$), ruggedness VRM ($p = 4.9 \times 10^{-9}$), mean temperature in the warmest month (MTWM) ($p = 0.000014$), Maximum temperature in the warmest

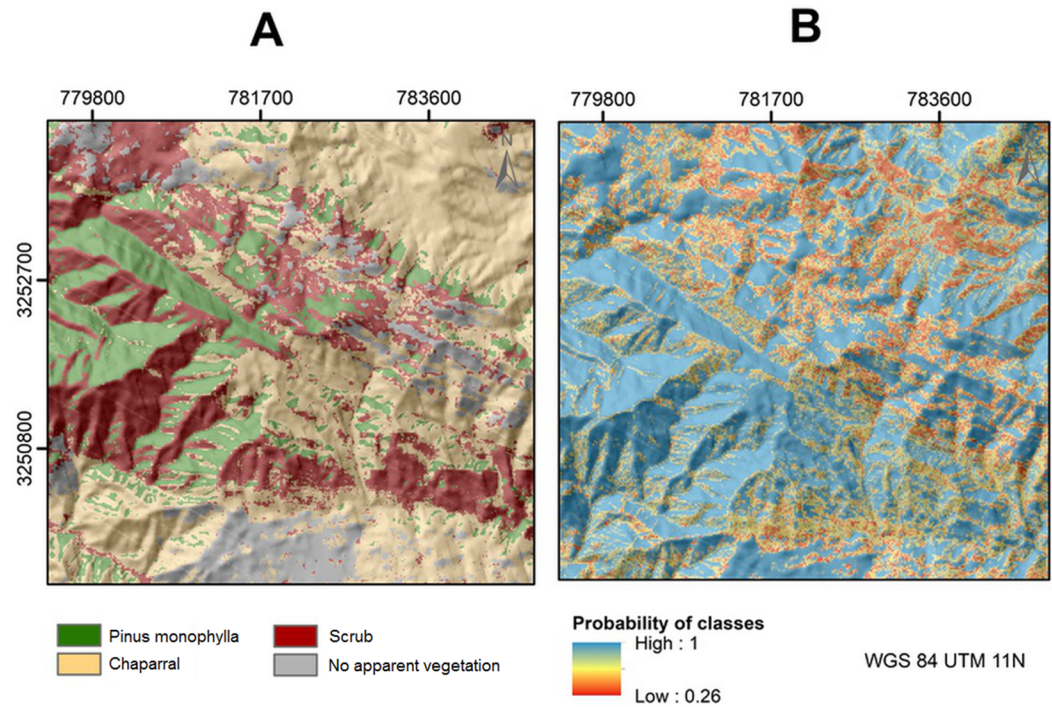


Figure 3 (A) Estimated land cover classes using BPNN classification in Sierra La Asamblea. (B) Probability map of class assignment. [Full-size !\[\]\(1663bb69f307a960345edb0e712f8c02_img.jpg\) DOI: 10.7717/peerj.4603/fig-3](https://doi.org/10.7717/peerj.4603/fig-3)

Table 3 Estimated error matrix based of sample counts (n_{ij}) from the accuracy assessment sample.

	Classes	Reference				Total	Map area (ha)	W_i
		P	S	C	WV			
Map	P	522	0	14	0	536	5,395	0.169
	S	24	619	119	2	764	12,309	0.387
	C	50	0	348	7	405	8,206	0.258
	NAP	0	0	20	418	438	5,913	0.186
	Total	596	619	501	427	2,143	31,823	1

Notes:

Map classes are the rows while the reference classes are the columns.

P, *Pinus monophylla*; S, shrub; C, chaparral; NAP, no apparent vegetation; W_i , proportion of the area mapped as class i .

Table 4 Error matrix of four classes with cell entries (p_{ij}) based on Table 3 and expressed in terms of proportion of area.

	Classes	References				Accuracy		
		P	S	C	WV	User's	Producer's	Overall
Map	P	0.1651	0.0000	0.0044	0.0000	0.974 ± 0.07	0.790 ± 0.04	0.877 ± 0.01
	S	0.0122	0.3134	0.0602	0.0010	0.810 ± 0.02	1.000	
	C	0.0318	0.0000	0.2216	0.0045	0.859 ± 0.01	0.752 ± 0.07	
	NAP	0.0000	0.0000	0.0085	0.1773	0.954 ± 0.002	0.970 ± 0.02	
	Total	0.2091	0.3134	0.2947	0.1828			

Notes:

Accuracy measures are presented with a 95% confidence interval. Map classes (rows), reference classes (columns).

P, *Pinus monophylla*; S, shrub; C, chaparral; NAP, no apparent vegetation.

Table 5 Results of the multiple linear binomial logistic regression model (AIC = 601.8; residual deviance = 593.85 on 588 degrees of freedom), TRI, terrain ruggedness index; VRM, vector ruggedness measure; MTWM, mean temperature in the warmest month.

Variable	Estimate	Std. Error	Z value	Pr(> z)
Intercept	25.351	8.895	2.850	0.0044
MTWM	-1.159	0.362	-3.201	0.0014
TRI	0.178	0.015	11.200	<2e-16
VRM	28.476	13.847	2.056	0.0397

month (MMAX) ($p = 0.000048$), and SPRP ($p = 0.00037$) were most variable between sites in which *P. monophylla* was present and absent. The variable slope was closely correlated with ruggedness as well as with MMAX and MTWM ($r_s > 0.7$). The p_{slope} of the Kruskal–Wallis test was larger than $p_{\text{ruggedness}}$ and p_{MMAX} was larger than p_{MTWM} . Slope and MMAX were therefore excluded from the model analysis. The stepwise multiple linear binomial logistical and Random Forest models showed that the “presence of *P. monophylla*” model included the independent variables ruggedness, ruggedness VRM, and average temperature in the warmest month (MTWM) (Table 5).

The ruggedness factor was the most influential predictor variable and indicated that the probability of *P. monophylla* occurrence was larger than 50% when the degree of ruggedness TRI was greater than 17.5 m (Table 5). The ruggedness VRM also indicated that a minimum change in roughness increases the probability of presence of the pine. The probability of occurrence of *P. monophylla* decreased when MTWM increased from 23.5 to 25.2 °C (Table 5). After cross validation (10-fold), the Random Forest model revealed that the variables ruggedness TRI, ruggedness VRM, and MTWM yielded a greater correlation for their ability to predict presence of the *P. monophylla* (AUC = 0.920, TSS = 0.690, Kappa = 0.691). The sensitivity was 0.812 and specificity was 0.878.

DISCUSSION

Pixel-based classification

Predicting the presence of pine forest by using BPNN proved feasible. The NDVI was one of the variables that contributed to the prediction and clearly separated forest cover (NDVI > 0.35) from the other types of vegetation cover (NDVI < 0.20). The overall accuracy of classification ($K = 0.87$) was similar to that reported in other studies using Sentinel-2A MSI images. For example, *Immitzer, Vuolo & Atzberger (2016)* reported a K of 0.85 for tree prediction in Europe by using five classes and a random forest classifier. *Vieira et al. (2003)* reported a $K = 0.77$ in eastern Amazon using seven classes and 1999 Landsat 7 ETM imagery. However, *Sothe et al. (2017)* reported K values of 0.98 and 0.90 for, respectively, three successional forest stages and field in a subtropical forest in Southern Brazil by using Sentinel-2 and Landsat-8 data associated with the support vector machine algorithm. *Kun et al. (2014)* estimated K values of 0.70 to 0.85 for land-use type prediction (including forest) in China by using the support vector machine algorithm classifier and Landsat-8 images of rougher spatial resolution than Sentinel images.

The very greater accuracy of predictions by [Kun et al. \(2014\)](#) was probably due to the large-scale of the study and the clearly differentiated types of land considered.

Relationship between presence of *P. monophylla* and environmental variables

Ruggedness of the terrain was the most important topographic variable, significantly explaining the presence of pines in Sierra La Asamblea ([Table 5](#)). Ruggedness, which is strongly positively correlated with slope, may reduce solar radiation, air temperature, and evapotranspiration due to increased shading ([Tsujino et al., 2006](#); [Bullock et al., 2008](#); [Allen et al., 2010](#)). The ruggedness indicated by the TRI index explains the presence of the pines because Sierra La Asamblea is heterogeneous in terms of elevation. The VRM index was less important partly because the index is strongly dependent on the vector aspect ([Gisbert & Martí, 2010](#)) and in the case of Sierra Asamblea the aspect is very homogeneous and the index values therefore tend to be very low ([Table 5](#)), as also reported by [Wu et al. \(2018\)](#). The pines were expected to colonize north facing slopes, which are exposed to less solar radiation than slopes facing other directions. However, the topographical variable aspect was not important in determining the presence of *P. monophylla* var. *californiarum* in the study site, possibly because of physiological adaptations regarding water-use efficiency and photosynthetic nitrogen-use efficiency ([DeLucia & Schlesinger, 1991](#)), as reported for the *P. monophylla*, *P. halepensis*, *P. edulis*, and *P. remota* in arid zones ([Lanner & Van Devender, 2000](#); [Helman et al., 2017](#)). The Mediterranean climate, with wet winters and dry summers, is another characteristic factor in this mountain range. In the winter in this part of the northern hemisphere, the sun (which is in a lower position and usually affects the southern aspect by radiation) is masked by clouds, rainfall and occasional snowfall ([León-Portilla, 1988](#)). During the summer, the solar radiation is more intense, but similar in all directions because the sun is closest to its highest point ([Stage & Salas, 2007](#)).

The above-mentioned finding contrasts with those of other studies reporting that north-eastern facing slopes in the northern hemisphere receive less direct solar radiation, thus providing more favorable microclimatic conditions (air temperature, soil temperature, soil moisture) for forest development, permanence, and productivity than southwest-facing sites ([Åström et al., 2007](#); [Stage & Salas, 2007](#); [Huang et al., 2009](#); [Marston, 2010](#); [Klein et al., 2014](#)). [DeLucia & Schlesinger \(1991\)](#) reported that *P. monophylla* populations in the Great Basin California desert with summer rainfall (monsoon) preferred an east-southeast aspect with less intense solar radiation and evapotranspiration.

The probability of occurrence of *P. monophylla* was also related to the climatic variable MTWM. In Sierra La Asamblea, this pine species was found in a narrow range of MTWM of between 23.5° and 25.2° ([Table 1](#)), which, however, is a smaller range than reported for the other pine species ([Tapias et al., 2004](#); [Roberts & Ezcurra, 2012](#)). Therefore, this species should adapt well to greater temperatures in the summer ([Lanner & Van Devender, 2000](#)), which is usually a very dry period in the study site ([León-Portilla, 1988](#)). However, the probability of occurrence was greatest for an MTWM of 23.5 °C ([Table 5](#)), which occurred at the top of Sierra La Asamblea, at an elevation of about 1,660 m). We therefore

conclude that this species can also grow well when the MTWM is below 23.5 °C. On the other hand, considering MTWM as factor yielded a probability of occurrence of 25%–80%. The spatial resolution of the climatic data by the national database run by the University of Idaho is probably not adequate for describing the microhabitat of *P. monophylla* (Rehfeldt et al., 2006; Marston, 2010).

Identification of the *P. monophylla* stands in Sierra La Asamblea as the most southern populations represents an opportunity for research on climatic tolerance and community responses to climatic variation and change.

ACKNOWLEDGEMENTS

We are grateful to E. Espinoza, F. Macias, and A. Guerrero for support with the fieldwork.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the National Council of Science and Technology (CONACYT) by grant #205965. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
National Council of Science and Technology (CONACYT): #205965.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Jonathan G. Escobar-Flores conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft, calculation of data and review of articles.
- Carlos A. Lopez-Sanchez conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft, calculation data.
- Sarahi Sandoval analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft, review of articles.
- Marco A. Marquez-Linares analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft, review of manuscript.
- Christian Wehenkel conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared

figures and/or tables, authored or reviewed drafts of the paper, approved the final draft, calculation of data and review of articles.

Data Availability

The following information was supplied regarding data availability:

The raw data are provided as [Supplemental Dataset Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.4603#supplemental-information>.

REFERENCES

- Abraira V.** 2001. El índice kappa. *Semergen* 27:247–249 DOI 10.1016/S1138-3593(01)73955-X.
- Allen CD, Macalady AK, Chenchouni H, Bachelet D, Vennetier M, Kitzberger G, Rigling H, Breshears D, Hoog T, Gonzalez PK, Fensham R, Zhangm Z, Castro J, Demidova N, Jong-Hwan L, Allard G, Running S, Semerci A, Cobbt N.** 2010. A global overview of drought and heat-induced tree mortality reveals emerging climatic change risks for forest. *Forest Ecology and Management* 259:660–684 DOI 10.1016/j.foreco.2009.09.001.
- Allouche O, Tsoar A, Kadmon R.** 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43(6):1223–1232 DOI 10.1111/j.1365-2664.2006.01214.x.
- Åström M, Dynesius M, Hylander K, Nilsson C.** 2007. Slope aspect modifies community responses to clear-cutting in boreal forests. *Ecology* 88(3):749–758 DOI 10.1890/06-0613.
- Bailey DK.** 1987. A study of *Pinus* subsection *Cembroides*. The single-needle pinyons of the Californias and the Great Basin. *Notes from the Royal Botanic Garden, Edinburgh* 44:275–310.
- Borràs J, Delegido J, Pezzola A, Pereira M, Morassi G, Camps-Valls G.** 2017. Land use classification from Sentinel-2 imagery. *Revista de Teledetección* 48:55–66 DOI 10.4995/raet.2017.7133.
- Braspenning PJ, Thuijsman F.** 1995. *Artificial Neural Networks: An Introduction to ANN Theory and Practice*. New York: Springer Science & Business Media, 295.
- Bullock SH, Heath D.** 2006. Growth rates and age of native palms in the Baja California desert. *Journal of Arid Environments* 67(3):391–402 DOI 10.1016/j.jaridenv.2006.03.002.
- Bullock SH, Salazar Ceseña JM, Rebman JP, Riemann H.** 2008. Growth rates and age of native palms in the Baja California desert Flora and vegetation of an isolated mountain range in the desert of Baja California. *The Southwestern Naturalist* 53(1):61–73 DOI 10.1894/0038-4909(2008)53[61:FAVOAI]2.0.CO;2.
- Callaway RM, DeLucia EH, Nowak R, Schlesinger WH.** 1996. Competition and facilitation: contrasting effects of *Artemisia tridentata* on desert vs. montane pines. *Ecology* 77(7):2130–2141 DOI 10.2307/2265707.
- Chambers JC.** 2001. *Pinus monophylla* establishment in an expanding *Pinus-Juniperus* woodland: environmental conditions, facilitation and interacting factors. *Journal of Vegetation Science* 12(1):27–40 DOI 10.2307/3236671.
- Cochran WG.** 1977. *Sampling Techniques*. New York: Wiley.
- CONABIO.** 2017. Comisión Nacional para el Conocimiento y uso de la Biodiversidad. Geoportal de información. Sistema Nacional de información sobre Biodiversidad. Available at <http://www.conabio.gob.mx/informacion/gis/> (accessed 12 February 2017).

- Congalton RG. 1991.** A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37(1):35–46 DOI 10.1016/0034-4257(91)90048-B.
- DeCastilho CV, Magnusson WE, de Araújo RNO, Luizao RC, Luizao FJ, Lima AP, Higuchi N. 2006.** Variation in aboveground tree live biomass in a central Amazonian Forest: effects of soil and topography. *Forest Ecology and Management* 234(1–3):85–96 DOI 10.1016/j.foreco.2006.06.024.
- DeLucia EH, Schlesinger WH. 1991.** Resource-use efficiency and drought tolerance in adjacent Great Basin and sierran plants. *Ecology* 72(1):51–58 DOI 10.2307/1938901.
- R Development Core Team. 2017.** *A Language and Environment for Statistical Computing*. Vienna: R foundation for statistical computing. Available at <http://www.R-project.org> (accessed 8 September 2017).
- Drusch M, Del Bello U, Carlier S, Colin O, Fernández V, Gascón F, Hoersch B, Isola C, Laberinti P, Martimort P, Meygret A, Spoto F, Sy O, Marchese F, Bargellini P. 2012.** Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing Environment* 120:25–36 DOI 10.1016/j.rse.2011.11.026.
- Elliott KJ, Miniati CF, Pederson N, Laseter SH. 2005.** Forest tree growth response to hydroclimate variability in the southern Appalachians. *Global Change Biology* 21(12):4627–4641 DOI 10.1111/gcb.13045.
- European Space Agency (ESA). 2017.** European Space Agency. Copernicus, Sentinel-2. Available at <http://www.esa.int> (accessed 21 March 2016).
- Fawcett T. 2006.** An introduction to ROC analysis. *Pattern Recognition Letters* 27(8):861–874 DOI 10.1016/j.patrec.2005.10.010.
- García E. 1998.** Clasificación de Köppen, modificado por García, E. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), 1998. Available at <http://www.conabio.gob.mx/informacion/gis/> (accessed 2 June 2017).
- Gisbert FJG, Martí IC. 2010.** Un índice de rugosidad del terreno a escala municipal a partir de Modelos de Elevación Digital de acceso público. *Documento de Trabajo*. Available at https://wheui3.grupobbva.com/TLFU/dat/DT_7_2010.pdf.
- Goodchild MF. 1994.** Integrating GIS and remote sensing for vegetation analysis and modeling: methodological issues. *Journal of Vegetation Science* 5(5):615–626 DOI 10.2307/3235878.
- Guo PT, Wu W, Sheng QK, Li MF, Liu HB, Wang ZY. 2013.** Prediction of soil organic matter using artificial neural network and topographic indicators in hilly areas. *Nutrient Cycling in Agroecosystems* 95(3):333–344 DOI 10.1007/s10705-013-9566-9.
- Helman D, Osem Y, Yakir D, Lensky IM. 2017.** Relationships between climate, topography, water use and productivity in two key Mediterranean forest types with different water-use strategies. *Agricultural and Forest Meteorology* 232:319–330 DOI 10.1016/j.agrformet.2016.08.018.
- Hepner G, Logan T, Ritter N, Bryant N. 1990.** Artificial neural network classification using a minimal training set. Comparison to conventional supervised classification. *Photogrammetric Engineering and Remote Sensing* 56(4):469–473.
- Huang C, Kim S, Song K, Townshend JRG, Davis P, Altstatt A, Rodas O, Yanoski A, Clay R, Tucker CJ, Musinsky J. 2009.** Assessment of Paraguay's forest cover change using Landsat observations. *Global and Planetary Change* 67:1–12 DOI 10.1016/j.gloplacha.2008.12.009.
- Immitzer M, Vuolo F, Atzberger C. 2016.** First experience with sentinel-2 data for crop and tree species classifications in Central Europe. *Remote Sensing* 8(3):1–27 DOI 10.3390/rs8030166.
- Instituto Nacional de Estadística y Geografía (INEGI). 2013.** Conjunto de datos vectoriales de uso de suelo y vegetación escala 1:250 000, serie V. Aguascalientes: Instituto Nacional de Estadística y Geografía. Available at <http://www.conabio.gob.mx/informacion/gis/> (accessed 10 September 2015).

- Klein T, Hoch G, Yakir D, Körner C. 2014.** Drought stress, growth and nonstructural carbohydrate dynamics of pine trees in a semi-arid forest. *Tree Physiology* **34(9)**:981–992 DOI [10.1093/treephys/tpu071](https://doi.org/10.1093/treephys/tpu071).
- Kun J, Xiangqin W, Xingfa G, Yunjun J, Xianhong X, Bin L. 2014.** Land cover classification using Landsat 8 Operational Land Imager data in Beijing, China. *Geocarto International* **29(8)**:941–951 DOI [10.1080/10106049.2014.894586](https://doi.org/10.1080/10106049.2014.894586).
- Lanner RM, Van Devender TR. 2000.** The recent history of pinyon pines. In: Richardson DM, ed. *The American Southwest*. Madrid: Cambridge University Press, 171–182.
- León-Portilla M. 1988.** *Miguel del Barco, Historia Natural y Crónica de la Antigua California*. México City: Universidad Nacional Autónoma de México, 483.
- Madonsela S, Cho MA, Ramoelo A, Mutanga O. 2017.** Remote sensing of species diversity using Landsat 8 spectral variables. *ISPRS Journal of Photogrammetry and Remote Sensing* **133**:116–127 DOI [10.1016/j.isprsjprs.2017.10.008](https://doi.org/10.1016/j.isprsjprs.2017.10.008).
- Marston RA. 2010.** Geomorphology and vegetation on hillslopes: interactions, dependencies, and feedback loops. *Geomorphology* **116(3–4)**:206–217 DOI [10.1016/j.geomorph.2009.09.028](https://doi.org/10.1016/j.geomorph.2009.09.028).
- Moran RV. 1983.** Relictual northern plants on peninsular mountain tops. In: Case TJ, Cody ML, eds. *Biogeography of the Sea of Cortez*. Berkeley: University of California Press, 408–410.
- Olofsson P, Foody GM, Herold M, Stehman SV, Woodcock CE, Wulder MA. 2014.** Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment* **148**:42–57 DOI [10.1016/j.rse.2014.02.015](https://doi.org/10.1016/j.rse.2014.02.015).
- Olofsson O, Foody GM, Stehman SV, Woodcock CE. 2013.** Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sensing of Environment* **129**:122–131 DOI [10.1016/j.rse.2012.10.031](https://doi.org/10.1016/j.rse.2012.10.031).
- Osem Y, Zangy E, Bney-Moshe E, Moshe Y, Karni N, Nisan Y. 2009.** The potential of transforming simple structured pine plantations into mixed Mediterranean forests through natural regeneration along a rainfall gradient. *Forest Ecology and Management* **259(1)**:14–23 DOI [10.1016/j.foreco.2009.09.034](https://doi.org/10.1016/j.foreco.2009.09.034).
- Pettorelli N. 2013.** *The Normalized Difference Vegetation Index*. New York: Oxford University Press, 194.
- QGIS Development Team. 2016.** QGIS Geographic Information System. Open source Geospatial Foundation. Available at <https://www.qgis.org>.
- Radoux J, Chomé G, Jacques DC, Waldner F, Bellemans N, Matton N, Lamarche C, d’Andrimont R, Defourny P. 2016.** Sentinel-2’s potential for sub-pixel landscape feature detection. *Remote Sensing* **8(6)**:488 DOI [10.3390/rs8060488](https://doi.org/10.3390/rs8060488).
- Rehfeldt GE. 2006.** A spline model of climate for the Western United States. General Technical Report. RMRS-GTR-165. Fort Collins: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Rehfeldt GE, Crookston NL, Warwell MV, Evans JS. 2006.** Empirical analyses of plant-climate relationships for the western United States. *International Journal Plant Science* **167(6)**:1123–1150 DOI [10.1086/507711](https://doi.org/10.1086/507711).
- Richards JA. 1999.** *Remote Sensing Digital Image Analysis*. Berlin: Springer-Verlag, 240.
- Riley SJ, Degloria SD, Elliot R. 1999.** A terrain ruggedness index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences* **5**:23–27.
- Roberts N, Ezcurra E. 2012.** Desert climate. In: Rebman JP, Roberts NC, eds. *Baja California Plant Field Guide*. San Diego: San Diego Natural History Museum, 1–23.

- Rouse JW, Haas RH, Schell A, Deering DW. 1974. Monitoring vegetation systems in the Great Plains with ERTS. In: *Proceedings of the Third Earth Resources Technology Satellite-1 Symposium, December 10–15, 1974, Greenbelt, MD*, NASA: Washington, D.C., 301–317.
- Sáenz-Romero C, Rehfeldt GE, Crookston NL, Duval P, St-Amant R, Beaulieu J, Richardson BA. 2010. Spline models of contemporary, 2030, 2060 and 2090 climates for Mexico and their use in understanding climate-change impacts on the vegetation. *Climatic Change* **102**(3–4):595–623 DOI [10.1007/s10584-009-9753-5](https://doi.org/10.1007/s10584-009-9753-5).
- Salas EAL, Valdez R, Michel S. 2017. Summer and winter habitat suitability of Marco Polo argali in southeastern Tajikistan: a modeling approach. *Heliyon* **3**(11):e00445 DOI [10.1016/j.heliyon.2017.e00445](https://doi.org/10.1016/j.heliyon.2017.e00445).
- Sappington JM, Longshore KM, Thompson DB. 2007. Quantifying landscape ruggedness for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert. *Journal of Wildlife Management* **71**(5):1419–1426 DOI [10.2193/2005-723](https://doi.org/10.2193/2005-723).
- Silva-Flores R, Pérez-Verdín G, Wehenkel C. 2014. Patterns of tree species diversity in relation to climatic factors on the Sierra Madre Occidental, Mexico. *PLOS ONE* **9**(8):e105034 DOI [10.1371/journal.pone.0105034](https://doi.org/10.1371/journal.pone.0105034).
- Shirk AJ, Waring K, Cushman S, Wehenkel C, Leal-Sáenz A, Toney C, Lopez-Sanchez CA. 2018. Southwestern white pine (*Pinus strobiformis*) species distribution models project a large range shift and contraction due to regional climate changes. *Forest Ecology and Management* **411**:176–186 DOI [10.1016/j.foreco.2018.01.025](https://doi.org/10.1016/j.foreco.2018.01.025).
- Sothe C, Almeida CMD, Liesenberg V, Schimalski MB. 2017. Evaluating sentinel-2 and Landsat-8 data to map successional forest stages in a subtropical forest in Southern Brazil. *Remote Sensing* **9**(8):838 DOI [10.3390/rs9080838](https://doi.org/10.3390/rs9080838).
- Spasojevic MJ, Bahlai CA, Bradley BA, Butterfield BJ, Tuanmu MN, Sistla S, Wiederholt R, Suding KN. 2016. Scaling up the diversity-resilience relationship with trait databases and remote sensing data: the recovery of productivity after wildfire. *Global Change Biology* **22**(4):1421–1432 DOI [10.1111/gcb.13174](https://doi.org/10.1111/gcb.13174).
- Stage AR, Salas C. 2007. Interactions of elevation, aspect, and slope in models of forest species composition and productivity. *Forest Science* **53**:486–492.
- Tan SS, Smeins FE. 1996. Predicting grassland community changes with an artificial neural network model. *Ecological Modelling* **84**(1–3):91–97 DOI [10.1016/0304-3800\(94\)00131-6](https://doi.org/10.1016/0304-3800(94)00131-6).
- Tapias R, Climent J, Pardos JA, Gil L. 2004. Life histories of Mediterranean pines. *Plant Ecology* **171**(1/2):53–68 DOI [10.1023/B:VEGE.0000029383.72609.f0](https://doi.org/10.1023/B:VEGE.0000029383.72609.f0).
- Telespazio VEGA Deutschland GmbH. 2016. Sentinel-2 MSI-Level-2A. Prototype processor installation and user manual. Available at <http://step.esa.int/thirdparties/sen2cor/2.2.1/S2PAD-VEGA-SUM-0001-2.2.pdf>.
- Tsujino R, Takafumi H, Agetsuma N, Yumoto T. 2006. Variation in tree growth, mortality and recruitment among topographic positions in a warm temperate forest. *Journal of Vegetation Science* **17**(3):281–290 DOI [10.1658/1100-9233\(2006\)17\[281:VITGMA\]2.0.CO;2](https://doi.org/10.1658/1100-9233(2006)17[281:VITGMA]2.0.CO;2).
- Venables WN, Ripley BD. 2002. *Modern Applied Statistics with S-Plus*. Fourth Edition. New York: Springer.
- Vieira ICG, de Almeida AS, Davidson EA, Stone TA, de Carvalho CJR, Guerrero JB. 2003. Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazonia. *Remote Sensing of Environment* **87**(4):470–481 DOI [10.1016/j.rse.2002.09.002](https://doi.org/10.1016/j.rse.2002.09.002).
- Wu W, Li AD, He XH, Ma R, Liu HB, Lv JK. 2018. A comparison of support vector machines, artificial neural network and classification tree for identifying soil texture classes in southwest China. *Computers and Electronics in Agriculture* **144**:86–93 DOI [10.1016/j.compag.2017.11.037](https://doi.org/10.1016/j.compag.2017.11.037).