

A peer-reviewed version of this preprint was published in PeerJ on 4 April 2018.

[View the peer-reviewed version](https://peerj.com/articles/4603) (peerj.com/articles/4603), which is the preferred citable publication unless you specifically need to cite this preprint.

Escobar-Flores JG, Lopez-Sanchez CA, Sandoval S, Marquez-Linares MA, Wehenkel C. 2018. Predicting *Pinus monophylla* forest cover in the Baja California Desert by remote sensing. PeerJ 6:e4603
<https://doi.org/10.7717/peerj.4603>

Detection of *Pinus monophylla* forest in the Baja California desert by remote sensing

Jonathan G Escobar-Flores¹, Carlos A Lopez-Sanchez², Sarahi Sandoval³, Marco A Marquez-Linares¹, Christian Wehenkel^{Corresp.}⁴

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, Durango, Mexico

³ CONACYT - Instituto Politécnico Nacional. CIIDIR. Unidad Durango, CONACYT - Instituto Politécnico Nacional. CIIDIR. Unidad Durango, Durango, Durango, México

⁴ Instituto de Silvicultura e Industria de la Madera, Universidad Juárez del Estado de Durango, Durango, Mexico

Corresponding Author: Christian Wehenkel
Email address: wehenkel@ujed.mx

The Californian single-leaf pinyon (*Pinus monophylla* var. *californiarum*), a subspecies of the single-leaf pinyon (the world's only 1-needled pine), inhabits semi-arid zones of the Mojave Desert in southern Nevada and southeastern California (US) and also of northern Baja California (Mexico). This subspecies is distributed as a relict in the geographically isolated arid Sierra La Asamblea, between 1,010 and 1,631 m, with mean annual precipitation levels of between 184 and 288 mm. The aim of this research was i) to establish the distribution of *Pinus monophylla* var. *californiarum* in Sierra La Asamblea, Baja California (Mexico) using Sentinel-2 images, and ii) to test and describe the relationship between this distribution of *Pinus monophylla* and five topographic and 18 climate variables. We hypothesized that i) the Sentinel-2 images can be used to accurately detect the *P. monophylla* distribution in the study site due to higher resolution (x3) and increased number of bands (x2) relative to Landsat-8, and ii) the topographical variables aspect, ruggedness and slope are particularly influential because they represent important microhabitat factors that can affect where conifers can become established and persist.

Methods. It was used an atmospherically corrected a 12-bit Sentinel-2A MSI image with eleven spectral bands in the visible, near infrared, and short-wave infrared light region combined with the normalized differential vegetation index (NDVI). Supervised classification of this image was carried out using a backpropagation-type artificial neural network algorithm. Stepwise multivariate binominal logistical regression and Random Forest regression including cross valuation (10 fold) were used to model the associations between presence/absence of pines and the five topographical and 18 climate variables.

Results. Probably, *P. monophylla* covers 4,955 hectares in the isolated in Sierra La Asamblea, Baja California (Mexico) via supervised classification of Sentinel-2 satellite images. The NDVI was one of the variables that contributed to the detection and clearly separated the forest cover (NDVI > 0.35) from the other vegetation cover (NDVI < 0.20). The ruggedness was the best environmental predictor variable and indicated that the probability of *P. monophylla* occurrence was higher than 50% when the degree of ruggedness was greater than 17.5 m. When average temperature in the warmest month increased from 23.5 to 25.2 °C, the probability of occurrence of *P. monophylla* decreased.

Discussion. The classification accuracy (Kappa) was similar to other studies using Sentinel-2A MSI images. Ruggedness is known to generate microclimates and provides shade that decreases evapotranspiration from pines in desert environments. Identification of *P. monophylla* 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.

1 **Detection of *Pinus monophylla* Forest in the Baja California** 2 **Desert by Remote Sensing**

3 Jonathan. G. Escobar-Flores ¹, Carlos Antonio López-Sánchez ², Sarahi Sandoval ³, Marco A. Márquez-
4 Linares ¹, Christian Wehenkel ²

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

7 ² Instituto de Silvicultura e Industria de la Madera, Universidad Juárez del Estado de Durango, Durango,
8 México

9 ³ CONACYT - Instituto Politécnico Nacional. CIIDIR. Unidad Durango. Durango, México

10

11 Corresponding Author:

12 Christian Wehenkel ²

13 Km 5.5 Carretera Mazatlán, Durango, 34120 Durango, México

14 Email address: wehenkel@ujed.mx

15

16 **ABSTRACT**

17 **Background.** The Californian single-leaf pinyon (*Pinus monophylla* var. *californiarum*), a
18 subspecies of the single-leaf pinyon (the world's only 1-needled pine), inhabits semi-arid zones
19 of the Mojave Desert in southern Nevada and southeastern California (US) and also of northern
20 Baja California (Mexico). This subspecies is distributed as a relict in the geographically isolated
21 arid Sierra La Asamblea at elevations of between 1,010 and 1,631 m, with mean annual
22 precipitation levels of between 184 and 288 mm. The aim of this research was i) to establish the
23 distribution of *P. monophylla* var. *californiarum* in Sierra La Asamblea, Baja California
24 (Mexico) using Sentinel-2 images, and ii) to test and describe the relationship between this
25 distribution of *P. monophylla* and five topographic and 18 climate variables. We hypothesized
26 that i) the Sentinel-2 images can be used to accurately detect the *P. monophylla* distribution in

27 the study site due to higher resolution (x3) and increased number of bands (x2) relative to
28 Landsat-8, and ii) the topographical variables aspect, ruggedness and slope are particularly
29 influential because they represent important microhabitat factors that can affect where conifers
30 can become established and persist. **Methods.** It was used an atmospherically corrected a 12-bit
31 Sentinel-2A MSI image with eleven spectral bands in the visible, near infrared, and short-wave
32 infrared light region combined with the normalized differential vegetation index (NDVI).
33 Supervised classification of this image was carried out using a backpropagation-type artificial
34 neural network algorithm. Stepwise multivariate binominal logistical regression and Random
35 Forest regression including cross valuation (10 fold) were used to model the associations
36 between presence/absence of *P. monophylla* and the five topographical and 18 climate variables.
37 **Results.** Probably, *P. monophylla* covers 4,955 hectares in the isolated in Sierra La Asamblea via
38 supervised classification of Sentinel-2 satellite images. The NDVI was one of the variables that
39 contributed to the detection and clearly separated the forest cover (NDVI > 0.35) from the other
40 vegetation cover (NDVI < 0.20). The ruggedness was the best environmental predictor variable
41 and indicated that the probability of *P. monophylla* occurrence was higher than 50% when the
42 degree of ruggedness was greater than 17.5 m. When average temperature in the warmest month
43 increased from 23.5 to 25.2 °C, the probability of occurrence of *P. monophylla* decreased.
44 **Discussion.** The classification accuracy was similar to other studies using Sentinel-2A MSI
45 images. Ruggedness is known to generate microclimates and provides shade that decreases
46 evapotranspiration from pines in desert environments. Identification of *P. monophylla* in the
47 Sierra La Asamblea as the most southern populations represents an opportunity for research on
48 climatic tolerance and community responses to climatic variation and change.

49

50 **INTRODUCTION**

51 The Californian single-leaf pinyon (*Pinus monophylla* var. *californiarum*), a subspecies of the
52 single-leaf pinyon (the world's only 1-needled pine), inhabits semi-arid zones of the Mojave
53 Desert in southern Nevada and southeastern California (US) and also of northern Baja California
54 (BC) (Mexico). It is cold-tolerant, drought resistant and is mainly differentiated from the typical
55 subspecies *Pinus monophylla* var. *monophylla* by a larger number of leaf resin canals and longer
56 fascicle-sheath scales (Bailey, 1987). This subspecies was first reported in BC in 1767 (Bullock
57 et al. 2006). The southernmost record of *P. monophylla* var. *californiarum* in America was
58 previously in BC, 26-30 miles north of Punta Prieta, at an elevation of 1,280 m (longitude -
59 114°.155; latitude 29°.070, catalogue number ASU 0000235), and the type specimen is held in
60 the Arizona State University Vascular Plant Herbarium.

61 This subspecies is distributed as a relict in the geographically isolated Sierra La Asamblea, at a
62 distance of 196 km from the Southern end of the Sierra San Pedro Martir and at elevations of
63 between 1,010 and 1,631 m (Moran, 1983, Table 2), with mean annual precipitation levels of
64 between 184 and 288 mm (Roberts and Ezcurra, 2012, Table 2). The Californian single-leaf
65 pinyon grows together with up to about 86 endemic plant species, although the number of
66 species decreases from north to south (Bullock et al. 2008).

67 Adaptation of *P. monophylla* var. *californiarum* to arid ecosystems enables the species to survive
68 annual precipitation levels below 150 mm. In fact, seeds of this variety display a high survival
69 rate under shrubs such as *Quercus* spp. and *Arctostaphylos* spp., a strategy that enables the pines
70 to widen their distribution, as has occurred in the great basin in California (Callaway et al. 1996;
71 Chambers, 2001) and for them to occupy desert zones such as Sierra de la Asamblea. Despite the

72 importance of this relict pine species, its existence is not considered in most forest inventories in
73 Mexico, and its distribution is generalized in vegetation cover maps (CONABIO, 2017).

74 Remote sensing techniques facilitate analysis of the temporal-space dynamics of the vegetation
75 in isolated sites, as with the piñon pine in Sierra de la Asamblea. Tree species distribution is
76 generally modulated by hydroclimatic variables and topographies (Elliot et al. 2005), and it is
77 therefore possible to determine the spatial conditions that favour the presence of forests,
78 especially by using digital terrain models (DTMs). Such models have shown, e.g., that tropical
79 and temperate forests tend to grow faster and more densely in sites with variable elevation and
80 slope (Decastilho et al. 2006; Spasojevic et al. 2016). Another attribute that can be analyzed
81 using DTMs is the slope, aspect and terrain ruggedness index (Riley et al. 1999), used to express
82 the difference in elevation of adjacent cells in a digital elevation grid. The less intense solar
83 radiation to which northern orientations are exposed is known to promote the growth and
84 productivity of vegetation (Osem et al. 2009). In addition, forests in less rugged sites (flatlands
85 and valleys) are known to be particularly susceptible to fire, while heterogeneous and highly
86 rugged forest sites are more dispersed and less likely to be affected by fire (Ganteaume and
87 Jappiot, 2013). Geomorphic attributes such as ruggedness and profile convexity have recently
88 been used to classify vegetation types from satellite images (Franklin et al. 2000; Carler &
89 Wolff, 2004; Waser et al. 2011) and have been used along with multi-temporal analysis of
90 images to identify forest types (Zhu & Liu, 2014).

91 Remote sensing with Landsat images has proved useful for detecting forest cover; the Landsat-8
92 satellite has sensors (7 bands) that can be used to analyze vegetation in spatial resolutions of 30
93 m (Johansen and Phinn, 2006). However, the European Space Agency's Copernicus program has

94 made Sentinel-2 satellite images available to the public free of charge. The spatial resolution (10
95 meters per pixel) is three times higher than that of Landsat images, thus increasing their potential
96 for detecting and differentiating between types of vegetation cover (Drush et al. 2012; Borrás et
97 al. 2017). The Sentinel-2 has 13 bands which provide high-quality radiometric images of spatial
98 resolution 10 - 20 m in the visible and infrared regions of the electromagnetic spectrum. These
99 images are therefore ideal for land classification (ESA, 2017).

100 The aim of this research was i) to establish the distribution of *Pinus monophylla* var.
101 *californiarum* in the Sierra La Asamblea, Baja California (Mexico) using Sentinel-2 images, and
102 ii) to test and describe the relationship between this distribution of *Pinus monophylla* and five
103 topographic and 18 climate variables. We hypothesized that i) the Sentinel-2 images can be used
104 to accurately detect the *P. monophylla* distribution in the study site due to higher resolution (x3)
105 and increased number of bands (x2) relative to Landsat-8, and ii) the topographical variables
106 aspect, ruggedness and slope are particularly influential because they represent important
107 microhabitat factors that can affect where conifers can become established and persist (Marston,
108 2010).

109 MATERIALS AND METHODS

110 *Study area*

111 Sierra La Asamblea is located in Baja California's central desert (-114° W 29° 19' N, range of
112 altitude 280-1,662 m, Fig. 1). The climate is arid, with maximum temperatures of 40° C in the
113 summer (García, 1998). The Sierra is steeper on the western slopes, with an average incline of
114 35°, and with numerous canyons with occasional springs and oases. The valleys and plateaus are
115 common in the proximity of the Gulf of California. Granite rocks occur south of the Sierra and

116 meta-sedimentary rocks along the north and southeast of the slopes. The predominant types of
117 vegetation are the xerophilous scrub, which is distributed at elevations ranging from 200 to 1,000
118 m. Chaparral begins at an altitude of 800 m, and representative specimens of *Adenostoma*
119 *fasciculatum*, *Ambrosia ambrosioides*, *Dalea bicolor orcuttiana* *Quercus tuberculata*, *Juniperus*
120 *california* and *Pinus monophylla* are also present at elevations higher than 1,000 m. Populations
121 of the endemic palm tree *Brahea armata* also occur in the lower parts of the canyons with
122 superficial water flow and through the rocky granite slopes (Bullock et al. 2006).

123 **Figure 1.** Map of Sierra La Asamblea. The black circles indicate georeferenced sites occupied by
124 *Pinus monophylla*.

125 *Establishing the distribution of P. monophylla var. californiarum using Sentinel-2 images*

126 Sentinel-2A multispectral instrument (MSI) L1C dataset acquired on 11 October 2016, in the
127 trajectory of coordinates latitude 29°814, longitude 114°93, was downloaded from the US
128 Geological Survey (USGS) Gloval Visualizaton Viewer at <http://glovis.usgs.gov/>. The 12-bit
129 Sentinel-2A MSI image has 13 spectral bands in the visible, NIR, and SWIR wavelength region
130 with spatial resolutions of 10-60 m. However, the band one used for studies of coastal aerosol
131 and the band ten applied for cirrus were not used in this study (ESA, 2017). Hence, the data
132 preparation involved the resampling of the seven S2 bands acquired at 20 m and 60 m to obtain a
133 layer stack of 11 spectral bands at 10 m (Table 1) using the ESA's Sentinel-toolbox ESA
134 Sentinel Application Platform (SNAP) and then converted to ENVI format.

135 Because atmospherically improved images are crucial to assess spectral indices with spatial
136 reliability and products comparison, level-1C data have been converted to level-2A (top-of-
137 canopy) taking into account the effects of aerosols and water vapor on reflectance (Radoux et al.,

138 2016). These corrections have been realized using the Sea2Cor tool (Telespazio VEGA
139 Deutschland GmbH, 2016) for the Sentinel-2 images.

140 **Table 1.** Sentinel-2 spectral bands used to detect the *Pinus monophylla* forest

141 The following equation was used to calculate the normalized difference vegetation index
142 (NDVI): $NDVI = (NIR - R) / (NIR + R)$, where NIR is the near infrared light (band) reflected by
143 the vegetation, and R is the visible red light reflected by the vegetation. The NDVI is useful for
144 discriminating the layers of temperate forest from scrub and chaparral. Areas occupied by large
145 amounts of unstressed green vegetation will have values much higher than 0 and areas with no
146 vegetation will have values close to 0 and, in some cases, negative values (Pettorelli, 2013).

147 The NDVI image was combined with the previously described multi spectral bands. Supervised
148 classification of this image was carried out using a backpropagation-type artificial neural
149 network (ANN) algorithm. The input weights corresponded to the values of the pixels twelve of
150 Sentinel-2 and of the NDVI image. A logistic activation function was used with a training rate of
151 0.20 and 40 interactive processes. The network also calculates the error at each iteration (RMS)
152 (Braspenning & Thuijisman, 1995). Additionally, the support vector machine algorithm (SVM)
153 was used to classify the MSI (Mountrakis et al., 2011).

154 The training sites corresponded to two groups of georeferenced sites (Datum WGS-84, 11N)
155 obtained during a project entitled "Oasis evaluation in Baja California". The coordinates of the
156 first group of sites whose were obtained in the field in October 2014 and October 2015. Four
157 classes were defined with the object improving the discrimination between vegetation cover. The
158 following classes were considered: i) pines, 502 sites, ii) scrub, 563 sites, iii) chaparral, 419 site

159 and iv) no apparent vegetation, 419 sites. The second group comprised control sites and the
160 coordinates were obtained by systematic sampling of the vegetation layers of the V series of land
161 use and cover vegetation of the National Institute of Statistics and Geography (INEGI, 2015).
162 The group included the following classes: i) pines, 596 sites, ii) scrub, 619 sites, iii) chaparral,
163 481 sites, and iv) no apparent vegetation, 418 sites.

164 The classification was validated using a confusion matrix, which is a table that compares the real
165 values with the classification results. The confusion matrix was also used to determine the user
166 accuracy, which refers to the total number of correct pixels/total number of reference pixels \times
167 100% (Congalton, 1991). The accuracy of classification was calculated using the Kappa (K)
168 coefficient. The K coefficient is a statistic used in accuracy assessment to measure whether one
169 error matrix is significantly different from another. This statistic takes values of between -1 and
170 +1, where values close to one indicate a high degree of agreement between classes and
171 observations, and a value of 0 suggests that the observed agreement is random (Abraira, 2001).

172 *Relationship between the distribution of *P. monophylla* and topographic and climate variables*

173 To test and model the association between presence/absence of *P. monophylla* in the study area
174 and topographical or climate variables, points estimates of the topographical variables
175 ruggedness, slope, aspect, elevation and convexity and 18 climate variables (Table 2) were
176 obtained from a national database managed by the University of Idaho
177 (<http://forest.moscowfsl.wsu.edu/climate/>) and which requires point coordinates (latitude,
178 longitude, and elevation) as the main inputs (Rehfeldt et al. 2006; Rehfeldt et al. 2006).

179 **Table 2.** Topographical and climatic variables considered in the study

180 For each variable in Table 2, a Kruskal-Wallis test was used to determine the difference in the
181 median values in relation to presence and absence of *P. monophylla*. All variables for which no
182 significant difference between the medians was obtained after Bonferroni correction ($\alpha = 0.0005$)
183 were excluded from further analysis. The colinearity between the significant variables was
184 measured using the Spearman correlation coefficient (r_s). When the r_s value for the difference
185 between two significant variables was larger than 0.7, only the variable with the lowest p value
186 in the Kruskal-Wallis test was used in the multivariate regressions. Finally, stepwise multivariate
187 binomial logistical regression and Random Forest regression including cross valuation (10 fold)
188 were used to model the associations between presence/absence of *P. monophylla* and the most
189 important topographical and climate variables (Shirk et al., 2017).

190 Regressions including cross valuation were carried out using the `trainControl`, `train`, `glm` (family
191 = "binomial") and `rf` functions, as well as the "randomForest" and "caret" packages (Venables
192 and Ripley, 2002) in R (version 3.3.2) (Development Core Team, 2017). The goodness-of-fit of
193 the regression models was evaluated using Akaike information criterion (*AIC*), root-mean-square
194 error (*RMSE*) and pseudo coefficient of determination (R^2).

195 RESULTS

196 Our model showed a potential *P. monophylla* cover of 4,955 hectares in the in Sierra de la
197 Asamblea, Baja California, Mexico. The supervised classification with ANN indicated an overall
198 accuracy of identification 89.78%. This level of accuracy was obtained in the 32 interactions
199 with 0.04 RMS training. The proportion of omission errors in the pine class was only 12.42%,
200 *i.e.* 87.58% of the pixels were correctly classified. The chaparral class had the highest proportion
201 of omission errors (27.65%) (Table 3, Fig. 2; Fig. 3). The value of NDVI in the pine forest

202 fluctuated between 0.30 - 0.41, and in chaparral between 0.24 - 0.28. The lowest values of NDVI
203 occurred in the scrub vegetation with values between 0.10 - 0.15. The analysis using the SVM
204 classifier only showed overall accuracy of 72%.

205 **Table 3.** Results of the classification monitored by neural network. The overall accuracy of
206 classification was 89.78%.

207 **Figure 2.** (A) Detection of *Pinus monophylla* by neural network classification. The light yellow
208 shading polygon represents pine forest published in the V series of INEGI (2013). (B)
209 Distribution of pines in the rugged sites in the Sierra La Asamblea (Photograph by Jonathan
210 Escobar).

211 **Figure 3.** Spectral signatures of cover vegetation in the Sierra La Asamblea, Baja California.

212 The Kruskal-Wallis test indicated that the median values for ruggedness ($p < 2.1e-16$), slope ($p <$
213 $2.2e-16$), MTWM ($p = 0.000014$), MMAX ($p = 0.000048$) and SPRP ($p = 0.00037$) were most
214 different between sites with presence and absence of *P. monophylla*. The variable slope was
215 closely correlated with ruggedness as well as with MMAX and MTWM ($r_s > 0.7$). The p_{slope} of
216 the Kruskal-Wallis test was larger than $p_{\text{ruggedness}}$ and p_{MMAX} larger than p_{MTWM} . Slope and
217 MMAX were therefore excluded from the multivariate regression analysis. The stepwise
218 multivariate binominal logistical and Random Forest regression showed that the best “presence
219 of pines” model included the independent variables ruggedness and average temperature in the
220 warmest month (MTWM) (Table 4).

221 **Table 4.** Results obtained with the best multivariate binomial logistic regression model (AIC =
222 611.96).

223 The ruggedness factor was the best predictor variable and indicated that the probability of *P.*
224 *monophylla* occurrence was higher than 50% when the degree of ruggedness was greater than
225 17.5 m (Fig. 4). When MTWM increased from 23.5 to 25.2 °C, the probability of occurrence of
226 *Pinus monophylla* decreased (Fig. 5). After cross validation (tenfold), the Random Forest model
227 revealed that the variables ruggedness and MTWM explained the variation in the presence of *P.*
228 *monophylla*, with $R^2 = 0.371$ and RMSE = 0.403.

229 **Figure 4.** The relationship between the probability (*P*) of occurrence of *Pinus monophylla* and
230 the ruggedness (m) of the terrain in Sierra La Asamblea, Baja California, Mexico.

231 **Figure 5.** The relationship between the probability (*P*) of occurrence of *Pinus monophylla* and
232 the average temperature in the warmest month (MTWM) in Sierra La Asamblea, Baja California,
233 Mexico.

234 DISCUSSION

235 Detection of pine forest by using ANN proved efficient. The NDVI was one of the variables that
236 contributed to the detection and clearly separated the forest cover (NDVI > 0.35) from the other
237 vegetation cover (NDVI < 0.20). The presence of the blue palm *Brahea armata* and fan palm
238 *Washingtonia filifera*, which grow in the canyons and had values of NDVI greater than 0.30,
239 may have confused the classification. However, these species are restricted to an elevation of less
240 than 1,000 m, and were therefore excluded from the classification (Bullock et al. 2008).

241 The overall classification accuracy in this study ($K = 0.90$) was similar to other studies using
242 Sentinel-2A MSI images. Immitzer et al. (2016) reported a *K* of 0.85 in tree detection in Europa
243 using five classes and random forest classifier in Europa. Vieira et al. (2003) found a $K = 0.77$ in

244 eastern Amazonia using seven classes and 1999 Landsat 7 ETM imagery. However, Sothe et al.
245 (2017) reported a $K = 0.98$ and $K = 0.90$, respectively evaluating three successional forest stages
246 and field in a subtropical forest in Southern Brazil by Sentinel-2 and Landsat-8 Data associated
247 with the support vector machine algorithm. Kun et al. (2014) showed a K of 0.70 to 0.85 in land-
248 use type detection including forests in China using the support vector machine algorithm
249 classifier and Landsat-8 images providing lower spatial resolution than Sentinel. The cause of
250 this very good accuracy of Kun et al. was probably the large-scale and clearly differentiated
251 land-use types used as classes.

252 Ruggedness of the terrain was the most important topographic variable, significantly explaining
253 the presence of pines in Sierra La Asamblea (Table 3). Ruggedness, which is strongly positively
254 correlated with slope, may reduce solar radiation, air temperature and evapotranspiration due to
255 increased shading (Di Castri et al. 1981; Tsujino et al. 2006; Bullock et al. 2008).

256 The pines were expected to colonize north facing slopes, which are exposed to less solar
257 radiation than slopes facing other directions. However, the topographical variable aspect was not
258 important in determining the presence of *P. monophylla* var. *californiarum* in the study site,
259 possibly because of physiological adaptations regarding water-use efficiency and photosynthetic
260 nitrogen-use efficiency (DeLucia and Schlesinger, 1991), as reported for the *Pinus monophylla*,
261 *P. halepensis*, *P. edulis*, *P. remota*, in arid zones (Lanner & Van Devender, 2000; Helman et al.
262 2017). The Mediterranean climate, with wet winters and dry summers, is another characteristic
263 factor in this mountain range. In the winter in this part of the northern hemisphere, the lower
264 position of the sun, which normally affects stronger the southern aspect by radiation could not
265 show to advantage due to clouds, rainfall and occasional snowfall (León-Portilla, 1988). During

266 the summer, the level of solar radiation is high, but similar in all directions because the sun is
267 closest to its highest point (Stage and Salas, 2007).

268 The above-mentioned finding contrasts with those of other studies reporting that north-eastern
269 facing slopes in the northern hemisphere receive less direct solar radiation, thus providing more
270 favourable microclimatic conditions (air temperature, soil temperature, soil moisture) for forest
271 development, permanence and productivity than southwest-facing sites (Astrom et al. 2007;
272 Stage & Salas, 2007; Hang et al 2009; Marston et al. 2010; Klein et al. 2014). DeLucia &
273 Schleinger (1991) reported for the *P. monophylla* populations in the Great Basin California
274 desert with summer rainfall (monsoon) that this tree species preferred an east-southeast aspect
275 with lower solar radiation and evapotranspiration (DeLucia & Schleinger, 1991).

276 The probability of occurrence of *P. monophylla* was also related to the climatic variable MTWM.
277 In the Sierra La Asamblea, this pine species was found in a narrow range of MTWM of between
278 23.5° and 25.2° (Table 1), which, however, is a wider temperature range than reported for the
279 other pine species (Tapias et al., 2004; Roberts & Ezcurra, 2012). Therefore, this species should
280 adapt well to high temperatures in the summer (Lanner et al., 1998), which is usually a very dry
281 period in the study site (León-Portilla, 1988). However, the probability of occurrence was the
282 highest for an MTWM of 23.5°C (Fig. 5, which occurred at the top of the Sierra La Asamblea, at
283 an elevation of about 1,660 m). We therefore conclude that this species can also grow well when
284 the MTWM is below 23.5°C. On the other hand, considering MTWM as factor yielded a
285 probability of occurrence of 25-80%. The spatial resolution of the climatic data by the national
286 database run by the University of Idaho is probably not adequate to describe the microhabitat of
287 *P. monophylla* (Rehfeldt et al., 2006; Marston et al., 2010).

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

291

292 **Acknowledgements**

293 We are grateful to E. Espinoza, F. Macias and A. Guerrero for their support in the field.

294 **References**

295 Abraira, V. 2001. El índice kappa. *Semergen* 27: 247-249. DOI:10.1016/S1138- 3593(01)73955-
296 X.

297 Allen, C.D., Macalady, A. .K., Chenchouni, H., Bachelet, D., Vennetier, M., Kitzberger, G.,
298 Rigling, H., Breshears, D., Hoog, T, Gonzalez, P. K., Fensham, R., Zhangm, Z, Castro, J.,
299 Demidova, N., Jong-Hwan, L., Allard, G., Running S., Semerci, A., Cobbt, N. 2010. A
300 global overview of drought and heat-induced tree mortality reveals emerging climatic
301 change risks for forest. *Forest ecology and management* 259:660–684. DOI:
302 10.1016/j.foreco.2009.09.001

303 Bailey, D. K. 1987. A study of Pinus subsection Cembroides. The single-needle pinyons of the
304 California and the Great Basin. Notes from the Royal Botanic Garden, Edinburgh. 44:275–
305 310.

- 306 Borràs, J., Delegido, J., Pezzola, A., Pereira, M., Morassi, G., Camps-Valls, G. 2017. Land use
307 classification from Sentinel-2 imagery. *Revista de Teledetección* 48: 55–66. DOI:
308 10.4995/raet.2017.7133.
- 309 Braspenning, P. J., Thuijsman, F. 1995. Artificial neural networks: an introduction to ANN
310 theory and practice. Springer Science & Business Media. USA. 295 p.
- 311 Brockmann Consult, 2017. Sentinel Application Platform (SNAP). Available at:
312 <http://step.esa.int/main/> / (accessed 18 April 2017).
- 313 Bullock, S. H., Heath, D. 2006. Growth rates and age of native palms in the Baja California
314 desert. *Journal of arid environments* 67:391–402- DOI: 10.1016/j.jaridenv.2006.03.002.
- 315 Bullock, S. H., Salazar Ceseña, J. M., Rebman, J. P., Riemann, H. 2008. Flora and vegetation of
316 an isolated mountain range in the desert of Baja California. *The Southwestern Naturalist*
317 53:61–73. DOI: 10.1894/0038-4909(2008)53[61:FAVOAI]2.0.CO;2
- 318 Callaway. R. M., DeLucia E. H., Nowak, R., Schlesinger W. H. 1996. Competition and
319 facilitation: contrasting effects of *Artemisia tridentata* on desert vs. montane pines.
320 *Ecology* 77:2130–2141. DOI: 10.2307/2265707
- 321 Carleer, A., Wolff, E. 2004. Exploitation of very high resolution satellite data for tree species
322 identification. *Photogrammetric Engineering & Remote Sensing* 70: 135–140. DOI:
323 10.14358/PERS.70.1.135

- 324 Chambers, J. C. 2001. *Pinus monophylla* establishment in an expanding Pinus-Juniperus
325 woodland: Environmental conditions, facilitation and interacting factors. *Journal of*
326 *Vegetation Science* 12: 27–40.
- 327 CONABIO. 2017. Comisión Nacional para el Conocimiento y uso de la Biodiversidad.
328 Geoportal de información. Sistema Nacional de información sobre Biodiversidad.
329 Available at: <http://www.conabio.gob.mx/informacion/gis/> (accessed 12 February 2017).
- 330 Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed
331 data. *Remote sensing of environment* 37:35–46.
- 332 DeCastilho, C. V., Magnusson, W. E., de Araújo, R. N. O., Luizao, R. C., Luizao, F. J., Lima, A.
333 P., Higuchi, N. 2006. Variation in aboveground tree live biomass in a central Amazonian
334 Forest: Effects of soil and topography. *Forest ecology and management* 234: 85–96. DOI:
335 10.1016/j.foreco.2006.06.024
- 336 Development Core Team. 2017. A language and environment for statistical computing. R
337 foundation for statistical computing, Vienna Austria. Available at: [http://www.R-](http://www.R-project.org)
338 [project.org](http://www.R-project.org). (accessed 8 September 2017).
- 339 Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernández, V., Gascón, F., Hoersch, B., Isola,
340 C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P.
341 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational
342 Services. *Remote sensing environment* 120: 25–36. DOI: 10.1016/j.rse.2011.11.026

- 343 Elliott, K.J., Miniati, C.F., Pederson, N., Laseter, S.H. 2005. Forest tree growth response to
344 hydroclimate variability in the southern Appalachians. *Global Change Biology* 21: 4627–
345 4641. DOI: 10.1111/gcb.13045
- 346 ESA, 2017. EuropeanSpaceAgency. Copernicus, Sentinel-2. *Available At:* <http://www.esa.int>
347 (accessed 21 March 2016).
- 348 Franklin, S.E., Hall, R.J., Moskal, L.M., Maudie, A.J., Lavigne, M.B. 2000. Incorporating
349 texture into classification of forest species composition from airborne multispectral images.
350 *International Journal of Remote Sensing* 21: 61–79. DOI: 10.1080/014311600210993
- 351 Ganteaume, A., Jappiot, M. 2013. What causes large fires in Southern France? *Forest Ecology*
352 *and Management* 294: 76–85. DOI: 10.1016/j.foreco.2012.06.055
- 353 García, E. 1998. Clasificación de Köppen, modificado por García, E. Comisión Nacional para el
354 Conocimiento y Uso de la Biodiversidad (CONABIO), 1998. Available at:
355 <http://www.conabio.gob.mx/informacion/gis/> (accessed 2 June 2017).
- 356 Heithecker, T. D., Halpern, C. B. 2007. Edge-related gradients in microclimate in forest
357 aggregates following structural retention harvests in western Washington. *Forest Ecology*
358 *and Management* 48: 163–173. DOI: 10.1016/j.foreco.2007.05.003
- 359 Helman, D., Osem, Y., Yakir, D., and Lensky, I. M. 2017. Relationships between climate,
360 topography, water use and productivity in two key Mediterranean forest types with
361 different water-use strategies. *Agricultural and Forest Meteorology* 232: 319–330. DOI:
362 10.1016/j.agrformet.2016.08.018

- 363 Immitzer, M., Vuolo, F., Atzberger, C. 2016. First Experience with Sentinel-2 Data for Crop and
364 Tree Species Classifications in Central Europe. *Remote Sensing* 8:1-27. DOI:
365 10.3390/rs8030166
- 366 INEGI. 2015. Conjunto de datos vectoriales de uso de suelo y vegetación escala 1:250 000, serie
367 V. Instituto Nacional de Estadística y Geografía. Aguascalientes. *Available at:*
368 <http://www.conabio.gob.mx/informacion/gis/> (accessed 10 September 2015).
- 369 Johansen, K., Phinn, S. 2006. Mapping structural parameters and species composition of riparian
370 vegetation using IKONOS and Landsat ETM+ data in Australian tropical
371 savannahs. *Photogrammetric Engineering & Remote Sensing* 72: 71–80. DOI:
372 10.14358/PERS.72.1.71
- 373 Klein, T., Hoch, G., Yakir, D., Körner, C. 2014. Drought stress, growth and nonstructural
374 carbohydrate dynamics of pine trees in a semi-arid forest. *Tree physiology* 34: 981–992.
375 DOI: 10.1093/treephys/tpu071
- 376 Kun, J., Xiangqin W., Xingfa G., Yunjun, J. Xianhong, X, Bin, L. 2014. Land cover
377 classification using Landsat 8 Operational Land Imager data in Beijing, China. *Geocarto*
378 *International* 29:941-951. DOI:10.1080/10106049.2014.894586
- 379 Lanner, R. M., Van Devender, T. R. 2000. The recent history of pinyon pines. In: Richardson, D.
380 M. (eds). *The American Southwest*, Cambridge University Press. 171–182
- 381 Léon-Portilla. 1988. Miguel del Barco, Historia natural y crónica de la antigua California.
382 Universidad Nacional Autónoma de México, México. 483 p.

- 383 Marston, R. A. 2010. Geomorphology and vegetation on hillslopes: Interactions, dependencies,
384 and feedback loops. *Geomorphology* 116: 206–217. DOI:
385 10.1016/j.geomorph.2009.09.028
- 386 Mountrakis, G., Im, J. and Ogole, C. 2011. Support vector machines in remote sensing: A
387 review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(3): 247-259. DOI:
388 10.1016/j.isprsjprs.2010.11.001
- 389 Moran, R. V. 1983. Relictual northern plants on peninsular mountain tops. In: *Biogeography of*
390 *the Sea of Cortez*; University of California Press, Berkeley, USA. 408–410.
- 391 Osem, Y., Zangy, E., Bney-Moshe, E., Moshe, Y., Karni, N., Nisan, Y. 2009. The potential of
392 transforming simple structured pine plantations into mixed Mediterranean forests through
393 natural regeneration along a rainfall gradient. *Forest Ecology Management* 259:14–23.
394 DOI:10.1016/j.foreco.2009.09.034
- 395 Pettorelli, N. 2013. *The Normalized Difference Vegetation Index*. Oxford, University Press.
396 United Kingdom. 194 p.
- 397 Radoux, J., Chomé, G., Jacques, D.C., Waldner, F., Bellemans, N., Matton, N., Lamarche, C.,
398 d'Andrimont, R. and Defourny, P., 2016. Sentinel-2's potential for sub-pixel landscape
399 feature detection. *Remote Sensing* 8(6): 488. DOI:10.3390/rs8060488
- 400 Rehfeldt G. A spline model of climate for the Western United States. 2006. Gen Tech Rep.
401 RMRS-GTR-165. U.S. Department of Agriculture, Forest Service, Rocky Mountain
402 Research Station, Fort Collins, Colorado, USA.

- 403 Rehfeldt G. E, Crookston N. L, Warwell M. V, Evans J. S. 2006. Empirical analyses of plant-
404 climate relationships for the Western United States. *International journal plant science*
405 167, 1123–1150. DOI: 1058-5893/2006/16706-0005
- 406 Riley, S. J., S. D. Degloria, and R. Elliot. 1999. A terrain ruggedness index that quantifies
407 topographic heterogeneity. *Intermountain Journal of Sciences* 5: 23–27
408 (<http://arcscripsts.esri.com/details.asp?dbid=12435>).
- 409 Roberts, N., y E. Ezcurra. Desert Climate. 2012. In: Rebman, J. P, Roberts. NC, eds. Baja
410 California Plant Field Guide; San Diego Natural History Museum. San Diego, USA. 1–23
- 411 Satage, A.R., Salas, C. 2007. Interactions of Elevation, Aspect, and Slope in Models of Forest
412 Species Composition and Productivity. *Forest Science* 53:486-492. Available at:
413 <http://www.ingentaconnect.com/>
- 414 Shirk A. J., Waring K., Cushman S., Wehenkel C., Leal-Sáenz A., Toney C., Lopez-Sanchez, C.
415 A. 2017. Southwestern white pine (*Pinus strobiformis*) species distribution models predict
416 large range shift and contraction due to climate change. *Forest Ecology Management* (in
417 review).
- 418 Sothe, C., Almeida, C.M.D., Liesenberg, V. and Schimalski, M.B. 2017. Evaluating Sentinel-2
419 and Landsat-8 Data to Map Sucessional Forest Stages in a Subtropical Forest in Southern
420 Brazil. *Remote Sensing* 9(8): 838. DOI:10.3390/rs9080838
- 421 Spasojevic, M. J., Bahlai, C. A., Bradley, B. A., Butterfield, B.J., Tuanmu, M.N., Sistla, S.,
422 Wiederholt, R., Suding, K. N. 2016. Scaling up the diversity-resilience relationship with

- 423 trait databases and remote sensing data: the recovery of productivity after wildfire. *Global*
424 *Change Biology* 1421–1432. DOI: 10.1111/gcb.13174
- 425 Tapias R., Climent, J., Pardos., Gil, L. 2004. Life histories of Mediterranean pines. *Plant*
426 *Ecology* 171: 53-68. DOI:10.1023/B:VEGE.0000029383.72609.f0
- 427 Telespazio VEGA Deutschland GmbH 2016. Sentinel-2 MSI-Level-2A. Prototype Processor
428 Installation and User Manual. Available at:
429 <http://step.esa.int/thirdparties/sen2cor/2.2.1/S2PAD-VEGA-SUM-0001-2.2.pdf>
- 430 Tsujino, R., Takafumi, H., Agetsuma, N., Yumoto, T. 2006. Variation in tree growth, mortality
431 and recruitment among topographic positions in a warm temperate forest. *Journal of*
432 *Vegetation Science* 17:281-290 DOI:10.1658/1100-9233(2006)17[281:VITGMA]2.0.CO;2
- 433 Venables W. N, Ripley B. D. 2002. Modern Applied Statistics with S-Plus. Fourth Edition.
434 New York, Springer.
- 435 Vieira, I.C.G., de Almeida, A.S., Davidson, E.A., Stone, T.A., de Carvalho, C.J.R. and Guerrero,
436 J.B. 2003. Classifying successional forests using Landsat spectral properties and ecological
437 characteristics in eastern Amazonia. *Remote Sensing of Environment* 87(4): 470-481.
438 DOI:10.1016/j.rse.2002.09.002
- 439 Waser, L. T., Ginzler, C., Kuechler, M., Baltsavias, E., Hurni, L. 2011. Semi-automatic
440 classification of tree species in different forest ecosystems by spectral and geometric
441 variables derived from Airborne Digital Sensor (ADS40) and RC30 data. *Remote Sensing*
442 *of Environment* 115: 76–85. DOI: 10.1016/j.rse.2010.08.006

443 Zhu, X., Liu, D. 2014. Accurate mapping of forest types using dense seasonal Landsat time-
444 series. *Journal of Photogrammetry and Remote Sensing* 96:1–11. DOI:
445 10.1016/j.isprsjprs.2014.06.012

446

447

Figure 1 (on next page)

Map of Sierra La Asamblea.

The black circles indicate georeferenced sites occupied by *Pinus monophylla*

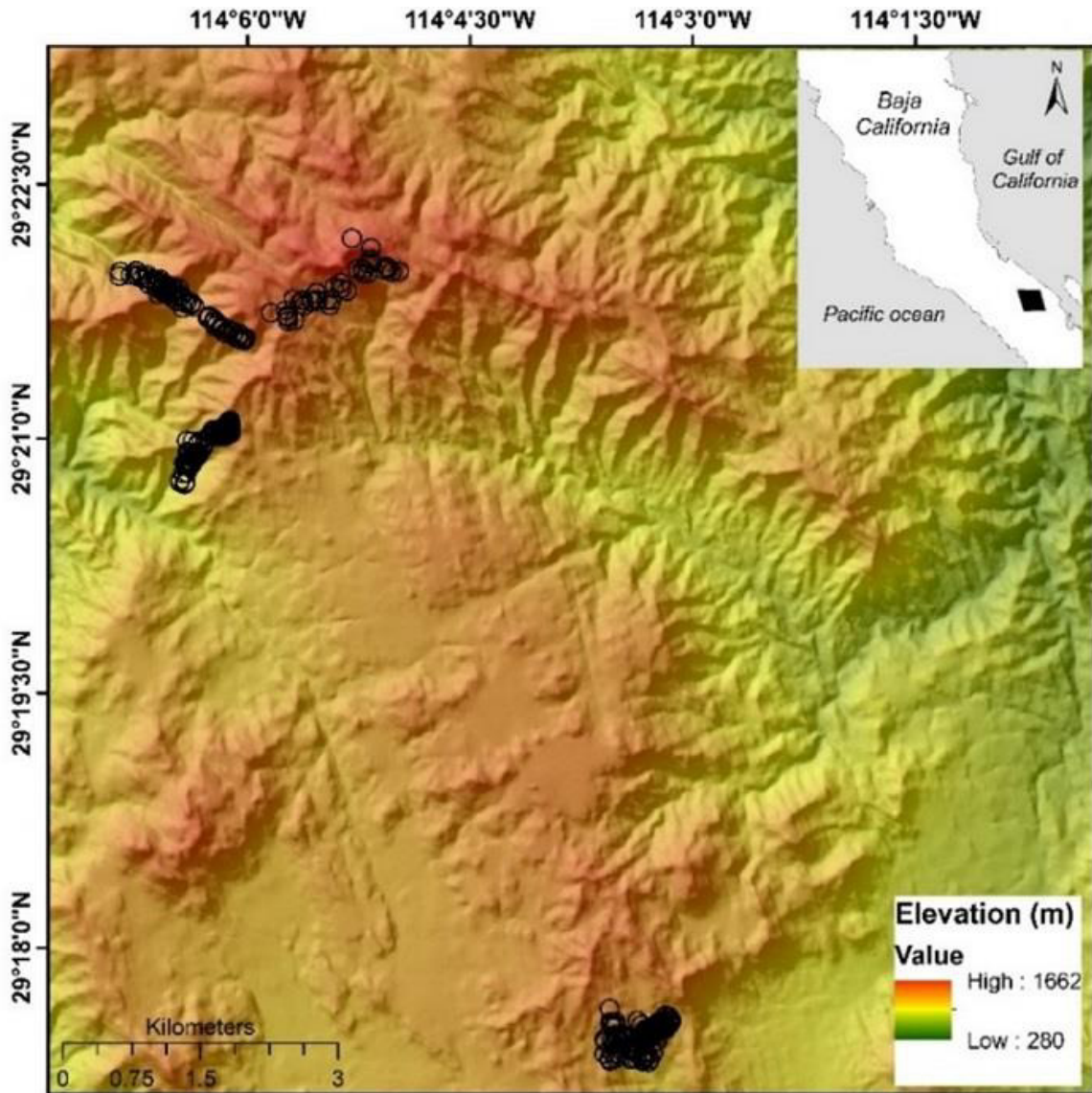


Figure 2 (on next page)

Detection of *Pinus monophylla*

(A) Detection of *Pinus monophylla* by neural network classification. The light yellow shading polygon represents pine forest published in the V series of INEGI (2013). (B) Distribution of pines in the rugged sites in the Sierra La Asamblea (Photograph by Jonathan Escobar)

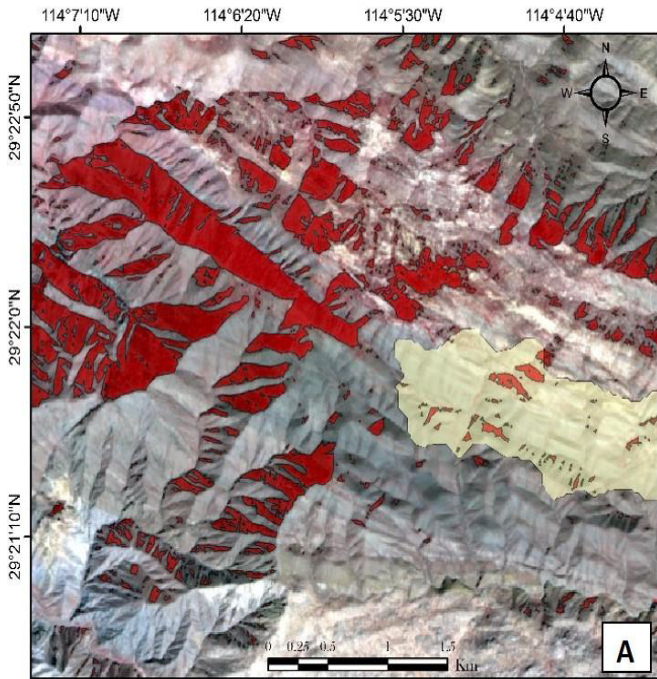


Figure 3 (on next page)

Spectral signatures

Spectral signatures of cover vegetation in the Sierra La Asamblea, Baja California

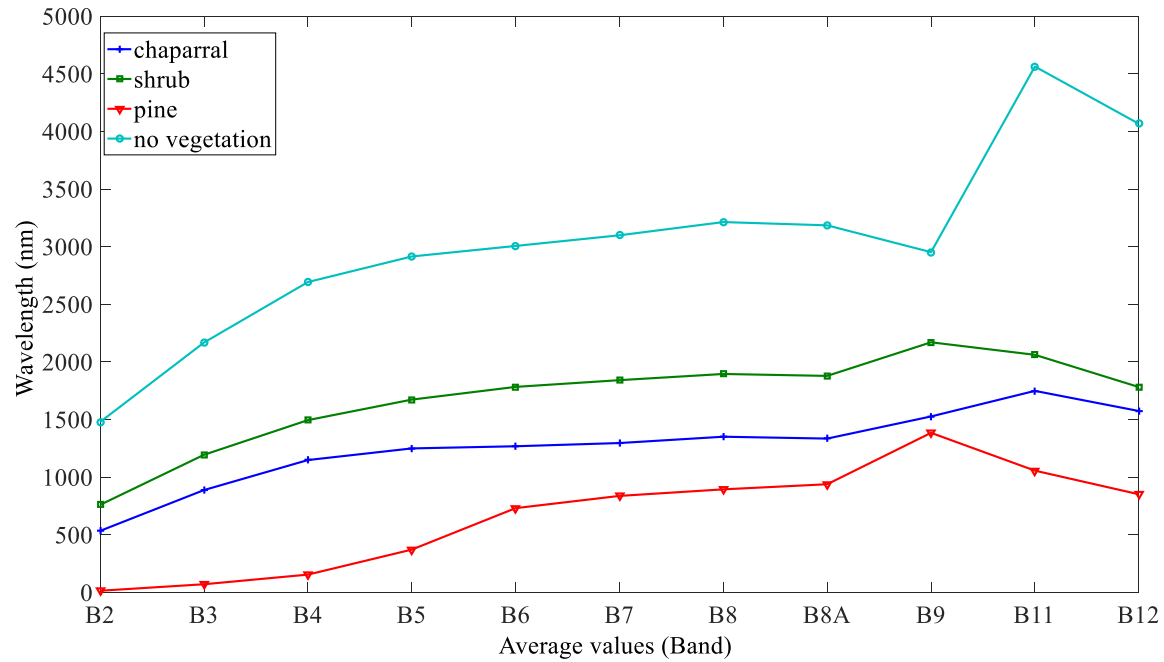


Figure 4(on next page)

The relationship between the probability (P) of occurrence of *Pinus monophylla* and the ruggedness

The relationship between the probability (P) of occurrence of *Pinus monophylla* and the ruggedness (m) of the terrain in Sierra La Asamblea, Baja California, Mexico

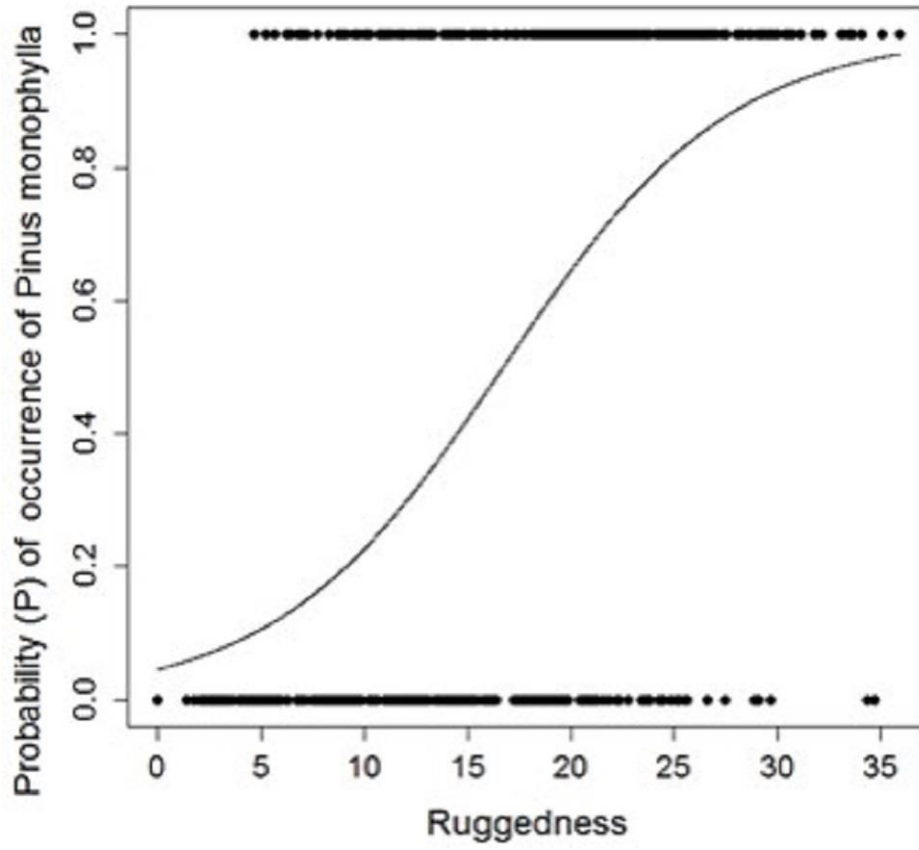


Figure 5(on next page)

The relationship between the probability (P) of occurrence of *Pinus monophylla* and the average temperature

The relationship between the probability (P) of occurrence of *Pinus monophylla* and the average temperature in the warmest month (MTWM) in Sierra La Asamblea, Baja California, Mexico

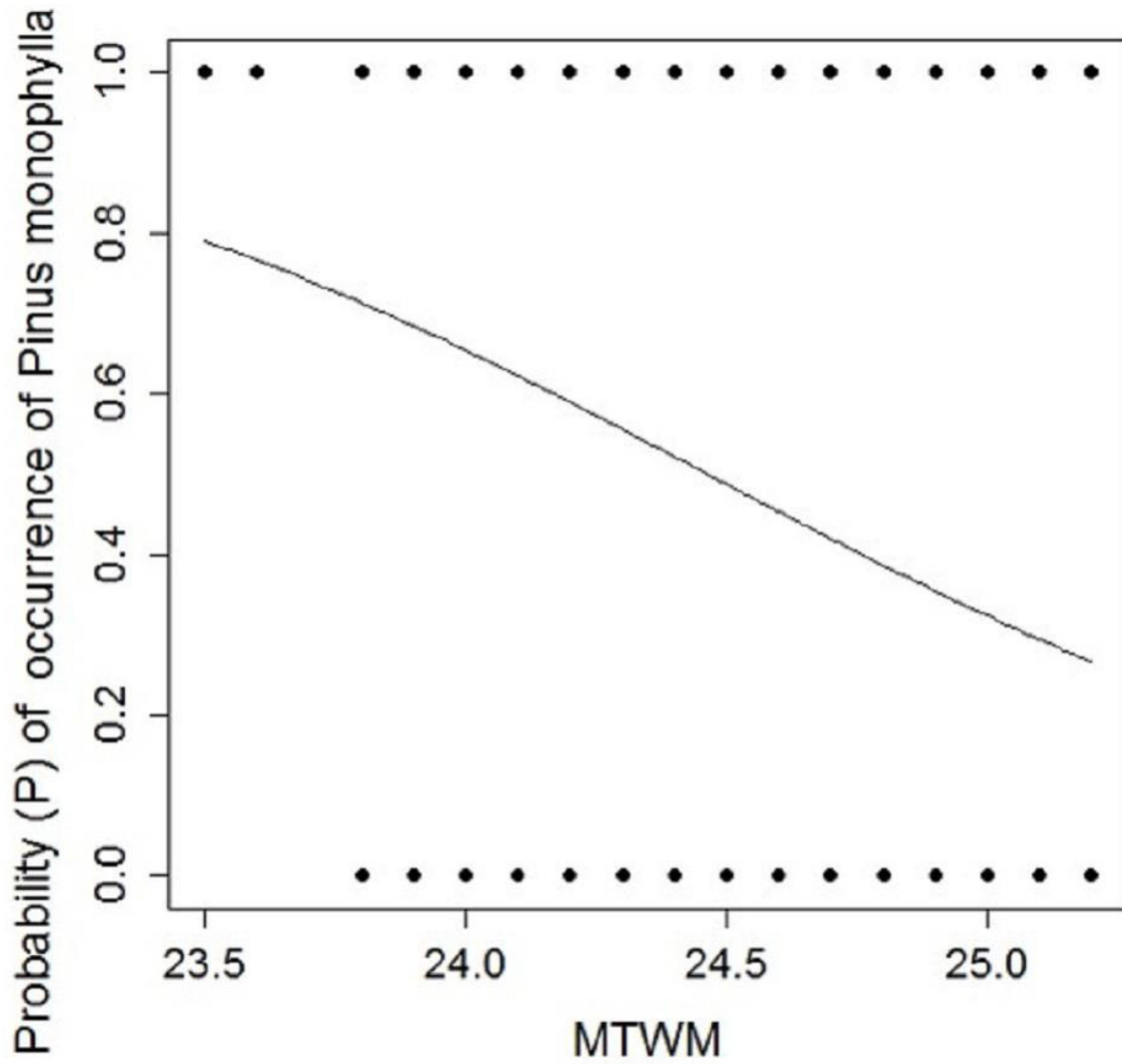


Table 1 (on next page)

Sentinel-2 spectral bands

Sentinel-2 spectral bands used to detect the *Pinus monophylla* forest

Bands	Central wave length (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 9 – Water vapour	0.945	60
Band 11 –SWIR	1.610	20
Band 12 –SWIR	2.190	20

1

Table 2 (on next page)

Topographical and climatic variables

Topographical and climatic variables considered in the study

Variable	Abbreviation	Units	Mean	SD	Max	Min
Ruggedness	IRT	m	20.33	6.66	35.90	4.69
Slope	S	°	28.38	8.92	48.34	3.42
Aspect	A	°	190.51	68.72	350.44	20.55
Elevation	E	m	1302.41	124.96	1631	1010
Convexity	C	°	-0.012	0.65	2.49	-2.44
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	4245.26	137.52	4550	3852
Degree days > 5°C accumulating within the frost-free period	GSDDD5	Days	3491.82	164.76	3944	2995
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	Smp	mm	43.74	6.29	62	29
Winter precipitation	Winp	mm	110.93	7.93	133	93

1

Table 3 (on next page)

Results of the classification

Results of the classification monitored by neural network. The value of the Kappa coefficient was 0.862. The accuracy of classification of pine forest was 89.78%

Training set data (Known Cover Types) *						
Classification data	P	S	C	WV	Total	User accuracy (%)
P	522	0	14	0	536	87.58
S	24	619	119	2	764	100
C	50	0	348	7	405	72.35
WV	0	0	20	418	409	97.85
Total	596	619	481	418	2,114	

1 * P = piñon pine; S = shrub; C = chaparral; WV= without vegetation

2

Table 4(on next page)

Results obtained with the best multivariate binomial

Results obtained with the best multivariate binomial logistic regression model (AIC = 611.96)

Factor	Estimate	RMSE	Z value	Pr(> z)
Intercept	26.38568	8.81813	2.992	0.00277
Ruggedness	0.18183	0.01579	11.519	<2e-16
MTWM	-1.19683	0.35920	-3.332	0.00086

1