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

Jonathan G Escobar-Flores [1], Carlos A Lopez-Sanchez [2], Sarahi Sandoval [3], Marco A Marquez-Linares [1], Christian Wehenkel Corresp. [2]

[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

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 (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 (x3) and greater number of bands (x2) 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 ten 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 (BPNN). 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 ± 319 hectares 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 (NDVI > 0.35) from the other vegetation cover (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 TRI was greater than 17.5 m. 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.

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

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

11 Corresponding author:

12 Christian Wehenkel [2]

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

14 Email address: wehenkel@ujed.mx

ABSTRACT

18 **Background.** The Californian single-leaf pinyon (*Pinus monophylla* var. *californiarum*), a

19 subspecies of the single-leaf pinyon (the world's only 1-needled pine), inhabits semi-arid zones

20 of the Mojave Desert (southern Nevada and southeastern California, US) and also of northern

21 Baja California (Mexico). This tree is distributed as a relict subspecies, at elevations of between

22 1,010 and 1,631 m in the geographically isolated arid Sierra La Asamblea (Baja California,

23 Mexico), an area characterized by mean annual precipitation levels of between 184 and 288 mm.

24 The aim of this research was i) to estimate the distribution of *P. monophylla* var. *californiarum*

25 in Sierra La Asamblea by using Sentinel-2 images, and ii) to test and describe the relationship

26 between the distribution of *P. monophylla* and five topographic and 18 climate variables. We

27 hypothesized that i) Sentinel-2 images can be used to predict the *P. monophylla* distribution in

28   the study site due to the finer resolution (x3) and greater number of bands (x2) relative to

29   Landsat-8 data, which is publically available free of charge and has been demonstrated to be

30   useful for estimating forest cover, and ii) the topographical variables aspect, ruggedness and

31   slope are particularly important because they represent important microhabitat factors that can

32   determine the sites where conifers can become established and persist. **Methods.** An

33   atmospherically corrected a 12-bit Sentinel-2A MSI image with ten spectral bands in the visible,

34   near infrared, and short-wave infrared light region was used in combination with the normalized

35   differential vegetation index (NDVI). Supervised classification of this image was carried out

36   using a backpropagation-type artificial neural network algorithm (BPNN). Stepwise multiple

37   linear binominal logistical regression and Random Forest classification including cross

38   validation (10-fold) were used to model the associations between presence/absence of *P.*

39   *monophylla* and the five topographical and 18 climate variables. **Results.** Using supervised

40   classification of Sentinel-2 satellite images, we estimated that *P. monophylla* covers 6,653 ± 319

41   (standard error) hectares in the isolated Sierra La Asamblea. The NDVI was one of the variables

42   that contributed most to the prediction and clearly separated the forest cover (NDVI > 0.35) from

43   the other vegetation cover (NDVI < 0.20). Ruggedness was the most influential environmental

44   predictor variable, indicating that the probability of occurrence of *P. monophylla* was greater

45   than 50% when the degree of ruggedness TRI was greater than 17.5 m. The probability of

46   occurrence of the species decreased when the mean temperature in the warmest month increased

47   from 23.5 to 25.2 °C. **Discussion.** The accuracy of classification was similar to that reported in

48   other studies using Sentinel-2A MSI images. Ruggedness is known to create microclimates and

49   provides shade that minimizes evapotranspiration from pines in desert environments.

50   Identification of the *P. monophylla* stands in Sierra La Asamblea as the most southern

51 populations represents an opportunity for research on climatic tolerance and community

52 responses to climate variability and change.

INTRODUCTION

54 The Californian single-leaf pinyon (*Pinus monophylla* var. *californiarum*), a subspecies of the

55 single-leaf pinyon (the world's only 1-needled pine), inhabits semi-arid zones of the Mojave

56 Desert (southern Nevada and southeastern California, US) and also of northern Baja California

57 (BC) (Mexico). It is both cold-tolerant and drought-resistant and is mainly differentiated from

58 the typical subspecies *Pinus monophylla* var. *monophylla* by a larger number of leaf resin canals

59 and longer fascicle-sheath scales (Bailey, 1987). This subspecies was first reported in BC in

60 1767 (Bullock et al., 2006). The southernmost record of *P. monophylla* var. *californiarum* in

61 America was previously in BC, 26-30 miles north of Punta Prieta, at an elevation of 1,280 m

62 (longitude -114 °.155; latitude 29 °.070, catalogue number ASU 0000235), and the type

63 specimen is held in the Arizona State University Vascular Plant Herbarium.

64 This species is distributed as a relict subspecies in the geographically isolated Sierra La

65 Asamblea, at a distance of 196 km from the Southern end of the Sierra San Pedro Martir and at

66 elevations of between 1,010 and 1,631 m (Moran, 1983) in areas with mean annual precipitation

67 levels of between 184 and 288 mm (Roberts & Ezcurra, 2012). The Californian single-leaf

68 pinyon grows together with up to about 86 endemic plant species, although the number of

69 species decreases from north to south (Bullock et al., 2008).

70 Adaptation of *P. monophylla* var. *californiarum* to arid ecosystems enables the species to survive

71 annual precipitation levels of less than 150 mm. In fact, seeds of this variety survive well under

72   shrubs such as *Quercus spp.* and *Arctostaphylus spp.,* a strategy that enables the pines to widen

73   their distribution, as has occurred in the great basin in California (Callaway et al., 1996;

74   Chambers, 2001), and for them to occupy desert zones such as Sierra de la Asamblea. Despite

75   the importance of this relict pine species, its existence is not considered in most forest

76   inventories in Mexico (CONABIO, 2017).

77   Remote sensing with Landsat images has been demonstrated to be useful for estimating forest

78   cover; the Landsat-8 satellite has sensors (7 bands) that can be predicted vegetation attributes at a

79   spatial resolution of 30 m (Madonsela et al., 2017). However, the European Space Agency's

80   Copernicus program has made Sentinel-2 satellite images available to the public free of charge.

81   The spatial resolution (10 m is pixel) of the images is three times finer that of Landsat images,

82   thus increasing their potential for predicting and differentiating types of vegetation cover (Drush

83   et al., 2012; Borras et al., 2017). The Sentinel-2 has 13 bands, of which 10 provide greater-

84   quality radiometric images of spatial resolution 10 to 20 m in the visible and infrared regions of

85   the electromagnetic spectrum. These images are therefore ideal for land classification (ESA,

86   2017).

87   The aim of this research was i) to estimate the distribution of *Pinus monophylla* var.

88   *californiarum* in Sierra La Asamblea, Baja California (Mexico) by using Sentinel-2 images, and

89   ii) to test and describe the relationship between this distribution of *P. monophylla* and five

90   topographic and 18 climate variables. We hypothesized that i) the Sentinel-2 images can be used

91   to accurately predict the *P. monophylla* distribution in the study site due to finer resolution (x3)

92   and greater number of bands (x2) than in Landsat-8 data, and ii) the topographical variables

93   aspect, ruggedness and slope are particularly influential because they represent important

94   microhabitat factors that can determine where conifers can become established and persist

95   (Marston, 2010).

MATERIALS AND METHODS

Study area

98   Sierra La Asamblea is located in Baja California's central desert (-114 ° 9' W 29 ° 19′ N,

99   elevation range 280-1,662 m, Fig. 1). The climate in the area is arid, with maximum

100   temperatures of 40 °C in the summer (Garcia, 1998). The sierra is steeper on the western slopes,

101   with an average incline of 35 °, and with numerous canyons with occasional springs and oases.

102   Valleys and plateaus are common in the proximity of the Gulf of California. Granite rocks occur

103   south of the sierra and meta-sedimentary rocks along the north and southeast of the slopes. The

104   predominant type of vegetation is xerophilous scrub, which is distributed at elevations ranging

105   from 200 to 1,000 m. Chaparral begins at an altitude of 800 m, and representative specimens of

106   *Adenostoma fasciculatum, Ambrosia ambrosioides, Dalea bicolor orcuttiana, Quercus*

107   *tuberculata, Juniperus california* and *Pinus monophylla* are also present at elevations above

108   1,000 m. Populations of the endemic palm tree *Brahea armata* also occur in the lower parts of

109   the canyons with superficial water flow and through the rocky granite slopes (Bullock et al.,

110   2006).

111   **Figure 1.** Map of Sierra La Asamblea. The black circles indicate georeferenced sites occupied by

112   *Pinus monophylla*.

113

114 **Datasets**

115 **Sentinel-2**

116 The Sentinel-2A multispectral instrument (MSI) L1C dataset, acquired on 11 October 2016, in

117 the trajectory of coordinates latitude 29 °.814, longitude 114 °.93, was downloaded from the US.

118 Geological Survey (USGS) Global Visualization Viewer at http://glovis.usgs.gov/. The 12-bit

119 Sentinel-2A MSI image has 13 spectral bands in the visible, NIR, and SWIR wavelength regions

120 with spatial resolutions of 10-60 m. However, band one, used for studies of coastal aerosols, and

121 bands nine and ten, applied for respectively water vapour correction and cirrus detection, were

122 not used in this study (ESA, 2017). Hence, the data preparation involved four bands at 10 m and

123 the resampling of the six S2 bands acquired at 20 m to obtain a layer stack of 10 spectral bands at

124 10 m (Table 1) using the ESA´s Sentinel-toolbox ESA Sentinel Application Platform (SNAP)

125 and then converted to ENVI format.

126 Because atmospherically improved images are essential to enable assessment of spectral indices

127 with spatial reliability and product comparison, Level-1C data were converted to Level-2A

128 (Bottom of Atmosphere -BOA- reflectance) by taking into account the effects of aerosols and

129 water vapour on reflectance (Radoux et al., 2016). The corrections were made using the Sen2Cor

130 tool (Telespazio VEGA Deutschland GmbH, 2016) for Sentinel-2 images.

131

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

133

134 The following equation was used to calculate the normalized difference vegetation index

135 (NDVI): NDVI = (NIR − R) / (NIR + R), where NIR is the near infrared light (band) reflected by

136 the vegetation, and R is the visible red light reflected by the vegetation (Rouse et al., 1974). The

137 NDVI is useful for discriminating the layers of temperate forest from scrub and chaparral. Areas

138 occupied by large amounts of unstressed green vegetation will have values much greater than 0

139 and areas with no vegetation will have values close to 0 and, in some cases, negative values

140 (Pettorelli, 2013). The NDVI image was combined with the previously described multi spectral

141 bands.

142 **Environmental variables**

143 Tree species distribution is generally modulated by hydroclimate and topographical variables

144 (Elliot et al., 2005; Decastilho et al., 2006), which can be estimated from digital terrain models

145 (DTM) (Osem et al., 2005; Spasojevic et al., 2016). A DTM was obtained by using tools

146 available   from   the   Instituto   Nacional   de   Estadistica   y   Geografía

147 (http:www.inegi.org.mx/geo/contenidos/datosrelieve) with a spatial resolution of 15 m. The

148 DTM was processed with the QGIS (QGIS Development Team, 2016), using *Terrain analysis*

149 tools, elevation, slope and aspect (Table 2).

150 The ruggedness was estimated using two indexes: i) the terrain ruggedness index (TRI) of Riley

151 et al. (1999) and ii) a vector ruggedness measure (VRM), both implemented in QGIS (QGIS

152 Development Team, 2016). The TRI computes the values for each grid cell of a DEM. This

153 calculates the sum change in elevation between a grid cell and its eight-neighbor grid cell. VRM

154 incorporates the heterogeneity of both slope and aspect. This measure of ruggedness uses 3-

155 dimensional dispersion of vectors normal to planar facets on landscape. This index lacks units

156 and ranges from 0 (indicating a totally flat area) to 1 (indicating maximum ruggedness)

157 (Sappington et al., 2007).

158 In addition, 18 climate variables with a 30-arc second resolution (approximate 800 meters)

159 (Table 2) were obtained from a national database managed by the University of Idaho

160 (http://charcoal.cnre.vt.edu/climate) and which requires point coordinates (latitude, longitude and

161 elevation) as the main inputs (Rehfeldt, 2006; Rehfeldt et al., 2006). These variables are

162 frequently used to study the potential effects of global warming on forests and plants in Western

163 North America and Mexico (Sáenz-Romero et al., 2010; Silva-Flores et al., 2014).

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

165

166 **Pixel-based classification**

167 **Classification method**

168 Pixel-based classification was carried out in order to predict four different types of land cover in

169 the study area (*P. monophylla*, scrub, chaparral and no apparent vegetation). A supervised

170 classification approach with a backpropagation-type artificial neural network (BPNN) (Tan and

171 Smeins, 1996) was applied. BPNN is widely used because of its structural simplicity and

172 robustness in modelling non-linear relationships. In this study, the BPNN comprises a set of

173 three layers (raster): an input layer, a hidden layer and an output layer (Richards, 1999). Each

174 layer consists of a series of parallel processing elements (neurons or nodes). Each node in a layer

175 is linked to all nodes in the next layer (Guo et al., 2013).

176 The first step in BPNN supervised classification is to enter the input layer, which in this study

177 corresponded to the values of the pixels of ten Sentinel-2 bands and of the NDVI image. Weights

178 were then assigned to the BPNN to produce analytical predictions from the input values. These

179 data were contrasted with the category to which each training pixel belongs, corresponding to

180 Georeferenced sites (Datum WGS-84, 11N) obtained in the field in October 2014 and October

181 2015.

182 A stratified random sampling method (Olofsson et al., 2013) was used to generate the reference

183 data in QGIS software (QGIS Development Team, 2016). A total of 2143 random points were

184 sampled, with at least 400 points for each class (Goodchild et al., 1994). The following classes

185 were considered: i) *P. monophylla*, 536 sites, ii) scrub, 764 sites, iii) chaparral, 405 sites, and iv)

186 no apparent vegetation, 438 sites. Class discrimination processes occurred in the hidden layer

187 and the synapses between the layers were estimated by an activation function. We used a logistic

188 function and training rate of 0.20, previously applied to land cover classification (Hepner et al.,

189 1990; Richards, 1999; Braspenning & Thuijisman, 1995). Learning occurs by adjusting the

190 weights in the node to minimize the difference between the output node activation, and BPNN

191 then calculates the error at each iteration with root mean square error (RMS). The output layer

192 comprised four neurons representing the four target classes of land cover (*P. monophylla*, Scrub,

193 Chaparral and no apparent vegetation). Average spectral signatures for the four different types of

194 land cover are shown in Figure 2.

195

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

197 California.

198 **Validation**

199 The BPNN classification was cross-validated (10-fold) using a confusion matrix, which is a table

200 that compares the reference data and the classification results. We estimated the uncertainty of

201 the classification using estimated error matrix in terms of proportion of area and estimates of

202 overall map accuracy ($\hat{O}$), user's accuracy ($\hat{U}_i$) (or commission error) and producer's accuracy ($\hat{P}_j$

203) (or omission error) recommended by Olofsson et al. (2013): p_{ij} is defined as a cell entry of error

204 matrix of i map classes. A poststratified estimator of p_{ij} is:

$$\hat{p}_{ij} = W_i \frac{n_{ij}}{n_{i.}} \tag{1}$$

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.

205 $\hat{p}_{.j}$ is a poststratified estimator for simple random and systematic sampling:

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

206 where q is the class number.

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

$$\hat{A}_j = A \cdot \hat{p}_{.j} \tag{3}$$

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

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

209 The standard error of the error-adjusted estimated area is

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

210   Finally,

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

211   presents an approximate 95% confidence interval.

212   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

213   proportion of the area mapped as class i that has reference class i. $\hat{P}_j$ of class j is the proportion

214   of the area of reference class j that is mapped as class j.

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

$$\hat{U}_i = \frac{\hat{p}_{ii}}{\hat{p}_{i.}} \tag{8}$$

$$\hat{P}_j = \frac{\hat{p}_{jj}}{\hat{p}_{.j}} \tag{9}$$

215   We then generated a map from the results of the probability of class assignment. The accuracy of

216   classification was also estimated using the Kappa (K) coefficient. The K coefficient is often used

217   as an overall measure of accuracy (Abraira, 2001). This coefficient takes values of between 0

218   and 1, where values close to one indicate a greater degree of agreement between classes and

219   observations, and a value of 0 suggests that the observed agreement is random. However, the use

220   of K is controversial because i) K would underestimate the probability that a randomly selected

221   pixel is correctly classified, ii) K is greater correlated with overall accuracy so reporting Kappa is

222   redundant for overall accuracy (Olofsson et al., 2014).

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

224 To model and test the association between presence/absence of *P. monophylla* in the study area

225 and topographical or climate variables, a Kruskal-Wallis test was used to estimate the difference

226 in the median values in relation to presence and absence of *P. monophylla*. All variables for

227 which no significant difference between the median values was predicted after Bonferroni

228 correction (α = 0.0005) were excluded from further analysis. The collinearity between the

229 variables with a significant difference between the medians of presence and absence was

230 estimated using the Spearman correlation coefficient (r_s). When the r_s value for the difference

231 between two variables was greater than 0.7, only the variable with the lowest p value in the

232 Kruskal-Wallis test was used in the models (as reported by Salas et al., 2017 and Shirk et al.,

233 2018). Finally, stepwise multiple linear binominal logistical regression and Random Forest

234 classification including cross valuation (10-fold) were used to model the associations between

235 presence/absence of *P. monophylla* and the most important topographical and climate variables

236 (Shirk et al., 2018).

237 Regression and classification including cross-validations were carried out using the trainControl,

238 train, glm (family = "binomial") and rf functions, as well as the "randomForest" and "caret"

239 packages (Venables and Ripley, 2002) in R (version 3.3.2) (Development Core Team, 2017).

240 The goodness-of-fit of the logistical regression model was evaluated using the Akaike

241 information criterion (AIC), root-mean-square error (RMSE) and residual deviance. Validation

242 of the randomForest model was performed using under the curve (AUC; Fawcett, 2006), True

243 Skill Statistic (TSS; Allouche et al., 2006), Kappa (Abraira, 2001), specificity and sensitivity.

244 **RESULTS**

245 **Pixel-based classification**

246 We estimated the area of *P. monophylla* cover with a margin of error (at approximate 95%

247 confidence interval) of 6,653 ± 319 (standard error) hectares in Sierra de la Asamblea, Baja

248 California, Mexico (Fig. 3). The confusion matrix of the accuracy assessment is listed in Table 3

249 including user´s and producer´s accuracy for each class. The supervised classification with

250 BPNN yielded predictions with an overall accuracy of identification of 87.74% (Table 4). This

251 level of accuracy was estimated in the 32 interactions with 0.04 RMS training. The proportion of

252 omission errors in the *P. monophylla* class was only 2.62%, *i.e.* 97.39% of the pixels were

253 correctly classified. The shrub class had the larger proportion of omission errors (18.98%). The

254 value of NDVI in the *P. monophylla* forest fluctuated between 0.30 and 0.41, and in chaparral

255 between 0.24 and 0.28. The smallest values of NDVI corresponded to scrub vegetation, with

256 values between 0.10 and 0.15.

257

258 **Figure 3**. (A) Estimated land cover classes using BPNN classification in Sierra La Asambla. (B)

259 Probability map of class assignment.

260

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

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

263

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

265 terms of proportion of area. Accuracy measures are presented with a 95% confidence interval.

266 Map classes (rows), reference classes (columns).

267

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

269 The Kruskal-Wallis test indicated that the median values for ruggedness TRI (p < 2.1×10^{-16}),

270 slope ($p < 2.2 \times 10^{-16}$), ruggedness VRM ($p = 4.9 \times 10^{-9}$), MTWM ($p = 0.000014$), MMAX ($p =$

271 0.000048) and SPRP ($p = 0.00037$) were most variable between sites in which *P. monophylla*

272 was present and absent. The variable slope was closely correlated with ruggedness as well as

273 with MMAX and MTWM ($r_s > 0.7$). The p_{slope} of the Kruskal-Wallis test was larger than

274 $p_{ruggedness}$ and p_{MMAX} was larger than p_{MTWM}. Slope and MMAX were therefore excluded from the

275 model analysis. The stepwise multiple linear binominal logistical and Random Forest models

276 showed that the "presence of *P. monophylla*" model included the independent variables

277 ruggedness, ruggedness VRM and average temperature in the warmest month (MTWM) (Table

278 5).

279 The ruggedness factor was the most influential predictor variable and indicated that the

280 probability of *P. monophylla* occurrence was larger than 50% when the degree of ruggedness

281 TRI was greater than 17.5 m (Table 5). The ruggedness VRM also indicated that a minimum

282 change in roughness increases the probability of presence of the pine. The probability of

283 occurrence of *Pinus monophylla* decreased when MTWM increased from 23.5 to 25.2 °C (Table

284 5). After cross validation (10-fold), the Random Forest model revealed that the variables

285 ruggedness TRI, ruggedness VRM and MTWM yielded a greater correlation for their ability to

286 predict presence of the *P. monophylla* (AUC = 0.920, TSS = 0.690, Kappa = 0.691). The

287 sensitivity was 0.812 and specificity was 0.878.

288 **Table 5.** Results of the multiple linear binomial logistic regression model (AIC = 601.8; residual

289 deviance= 593.85 on 588 degrees of freedom), TRI = terrain ruggedness index, VRM = vector

290 ruggedness measure, MTWM = mean temperature in the warmest month.

DISCUSSION

Pixel-based classification

293 Predicting the presence of pine forest by using BPNN proved feasible. The NDVI was one of the

294 variables that contributed to the prediction and clearly separated forest cover (NDVI > 0.35)

295 from the other types of vegetation cover (NDVI < 0.20). The overall accuracy of classification

296 (K = 0.87) was similar to that reported in other studies using Sentinel-2A MSI images. For

297 example, Immitzer et al. (2016) reported a K of 0.85 for tree prediction in Europe by using five

298 classes and a random forest classifier. Vieira et al. (2003) reported a K = 0.77 in eastern Amazon

299 using seven classes and 1999 Landsat 7 ETM imagery. However, Sothe et al. (2017) reported K

300 values of 0.98 and 0.90 for respectively three successional forest stages and field in a subtropical

301 forest in Southern Brazil by using Sentinel-2 and Landsat-8 data associated with the support

302 vector machine algorithm. Kun et al. (2014) estimated K values of 0.70 to 0.85 for land-use type

303 prediction (including forest) in China by using the support vector machine algorithm classifier

304 and Landsat-8 images of rougher spatial resolution than Sentinel images. The very greater

305 accuracy of predictions by Kun et al. (2014) was probably due to the large-scale of the study and

306 the clearly differentiated types of land considered.

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

308 Ruggedness of the terrain was the most important topographic variable, significantly explaining

309 the presence of pines in Sierra La Asamblea (Table 5). Ruggedness, which is strongly positively

310 correlated with slope, may reduce solar radiation, air temperature and evapotranspiration due to

311 increased shading (Tsujino et al., 2006; Bullock et al., 2008). The ruggedness indicated by the

312 TRI index explains the presence of the pines because Sierra La Asamblea is heterogeneous in

313 terms of elevation. The VRM index was less important partly because the index is strongly

314 dependent on the vector aspect (Gisbert & Martí, 2010) and in the case of Sierra Asamblea the

315 aspect is very homogeneous and the index values therefore tend to be very low (Table 5), as also

316 reported by Wu et al. (2018). The pines were expected to colonize north facing slopes, which are

317 exposed to less solar radiation than slopes facing other directions. However, the topographical

318 variable aspect was not important in determining the presence of *P. monophylla* var.

319 *californiarum* in the study site, possibly because of physiological adaptations regarding water-

320 use efficiency and photosynthetic nitrogen-use efficiency (DeLucia & Schlesinger, 1991), as

321 reported for the *Pinus monophylla, P. halepensis, P. edulis* and *P. remota* in arid zones (Lanner

322 & Van Devender, 2000; Helman et al., 2017). The Mediterranean climate, with wet winters and

323 dry summers, is another characteristic factor in this mountain range. In the winter in this part of

324 the northern hemisphere, the sun (which is in a lower position and usually affects the southern

325 aspect by radiation) is masked by clouds, rainfall and occasional snowfall (León-Portilla, 1988).

326 During the summer, the solar radiation is more intense, but similar in all directions because the

327 sun is closest to its highest point (Stage & Salas, 2007).

328 The above-mentioned finding contrasts with those of other studies reporting that north-eastern

329 facing slopes in the northern hemisphere receive less direct solar radiation, thus providing more

330   favourable microclimatic conditions (air temperature, soil temperature, soil moisture) for forest

331   development, permanence and productivity than southwest-facing sites (Astrom et al., 2007;

332   Stage & Salas, 2007; Hang et al 2009; Marston et al., 2010; Klein et al., 2014). DeLucia &

333   Schleinger (1991) reported that *P. monophylla* populations in the Great Basin California desert

334   with summer rainfall (monsoon) preferred an east-southeast aspect with less intense solar

335   radiation and evapotranspiration.

336   The probability of occurrence of *P. monophylla* was also related to the climatic variable MTWM.

337   In Sierra La Asamblea, this pine species was found in a narrow range of MTWM of between

338   23.5° and 25.2° (Table 1), which, however, is a smaller range than reported for the other pine

339   species (Tapias et al., 2004; Roberts & Ezcurra, 2012). Therefore, this species should adapt well

340   to greater temperatures in the summer (Lanner et al., 2000), which is usually a very dry period in

341   the study site (León-Portilla, 1988). However, the probability of occurrence was greatest for an

342   MTWM of 23.5 °C (Table 5), which occurred at the top of Sierra La Asamblea, at an elevation

343   of about 1,660 m). We therefore conclude that this species can also grow well when the MTWM

344   is below 23.5 °C. On the other hand, considering MTWM as factor yielded a probability of

345   occurrence of 25-80%. The spatial resolution of the climatic data by the national database run by

346   the University of Idaho is probably not adequate for describing the microhabitat of *P.*

347   *monophylla* (Rehfeldt et al., 2006; Marston et al., 2010).

348   Identification of the *P. monophylla* stands in Sierra La Asamblea as the most southern

349   populations represents an opportunity for research on climatic tolerance and community

350   responses to climatic variation and change.

ACKNOWLEDGEMENTS

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

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). J. Appl. Ecol. 43, 1223–1232. DOI:10.1111/j.1365-2664.2006.01214.x

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 P J, Thuijsman F. 1995. Artificial neural networks: an introduction to ANN theory and practice. Springer Science & Business Media. USA. 295 p.

371 Brockmann Consult, 2017. Sentinel Application Platform (SNAP). Available at:

372 http://step.esa.int/main. / (accessed 18 April 2017).

373 Bullock SH, Heath D. 2006. Growth rates and age of native palms in the Baja California desert.

374 *Journal of Arid Environments* 67(3):391-402. DOI: 10.1016/j.jaridenv.2006.03.002.

375 Bullock SH, Salazar Ceseña JM, Rebman JP, Riemann H. 2008. Flora and vegetation of an

376 isolated mountain range in the desert of Baja California. *The Southwestern Naturalist*

377 53:61-73. DOI: 10.1894/0038-4909(2008)53[61:FAVOAI]2.0.CO;2.

378 Callaway RM, DeLucia EH, Nowak R, Schlesinger WH. 1996. Competition and facilitation:

379 contrasting effects of *Artemisia tridentata* on desert vs. montane pines. *Ecology* 77:2130-

380 2141. DOI: 10.2307/2265707.

381 Chambers JC. 2001. *Pinus monophylla* establishment in an expanding *Pinus-Juniperus*

382 woodland: Environmental conditions, facilitation and interacting factors. *Journal of*

383 *Vegetation Science* 12:27-40.

384 Cochran, W. G., 1977. Sampling techniques. New York, NY: Wiley.

385 CONABIO. 2017. Comisión Nacional para el Conocimiento y uso de la Biodiversidad.

386 Geoportal de información. Sistema Nacional de información sobre Biodiversidad.

387 *Available at*: http://www.conabio.gob.mx/informacion/gis/ (accessed 12 February 2017).

388 Congalton RG. 1991. A review of assessing the accuracy of classifications of remotely sensed

389 data. Remote sensing of environment 37:35-46. DOI: 10.1016/0034-4257(91)90048-B

390   DeCastilho CV, Magnusson WE, de Araújo RNO, Luizao RC, Luizao FJ, Lima AP, Higuchi N.

391   2006. Variation in aboveground tree live biomass in a central Amazonian Forest: Effects of

392   soil and topography. *Forest ecology and management* 234:85-96. DOI:

393   10.1016/j.foreco.2006.06.024.

394   DeLucia, EH, & Schlesinger, WH. 1991. Resource-use efficiency and drought tolerance in

395   adjacent Great Basin and sierran plants. *Ecology*, *72*(1), 51-58. DOI: 10.2307/1938901

396   Development Core Team. 2017. A language and environment for statistical computing. R

397   foundation for statistical computing, Vienna Austria. *Available at*: http://www.R-

398   project.org. (accessed 8 September 2017).

399   Drusch M, Del Bello U, Carlier S, Colin O., Fernández V, Gascón F, Hoersch B, Isola C,

400   Laberinti, P, Martimort P, Meygret A, Spoto F, Sy O, Marchese F, Bargellini P. 2012.

401   Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services.

402   *Remote sensing environment* 120:25-36. DOI: 10.1016/j.rse.2011.11.026.

403   Elliott KJ, Miniat CF, Pederson N, Laseter SH. 2005. Forest tree growth response to

404   hydroclimate variability in the southern Appalachians. *Global Change Biology*

405   21(12):4627-4641. DOI: 10.1111/gcb.13045.

406   ESA, 2017. European Space Agency. Copernicus, Sentinel-2. *Available At*: http://www.esa.int

407   (accessed 21 March 2016).

408   Fawcett, T. 2006. An introduction to ROC analysis. Pattern Recognition Letters 27:861–874.

409   DOI: 10.1016/j.patrec.2005.10.010

410 García E. 1998. Clasificación de Köppen, modificado por García, E. Comisión Nacional para el

411   Conocimiento y Uso de la Biodiversidad (CONABIO), 1998. Available at:

412   http://www.conabio.gob.mx/informacion/gis/ (accessed 2 June 2017).

413 Gisbert FJG, Martí IC. 2010. Un índice de rugosidad del terreno a escala municipal a partir de

414   Modelos de Elevación Digital de acceso público. *Documento de Trabajo*. Available at:

415   https://wheui3.grupobbva.com/TLFU/dat/DT_7_2010.pdf

416 Goodchild MF. 1994. Integrating GIS and remote sensing for vegetation analysis and modeling:

417   methodological issues. *Journal of Vegetation Science* 5:615-626. DOI: 10.2307/3235878.

418 Guo PT, Wu W, Sheng QK, Li MF, Liu HB, Wang ZY. 2013. Prediction of soil organic matter

419   using artificial neural network and topographic indicators in hilly areas. *Nutrient cycling in*

420   *agroecosystems* 95:333344. DOI: 10.1007/s10705-013-9566-9.

421 Helman D, Osem Y, Yakir D, Lensky IM. 2017. Relationships between climate, topography,

422   water use and productivity in two key Mediterranean forest types with different water-use

423   strategies. *Agricultural and Forest Meteorology* 232:319-330. DOI:

424   10.1016/j.agrformet.2016.08.018.

425 Hepner G, Logan T, Ritter N, Bryant N. 1990. Artificial neural network classification using a

426   minimal training set. Comparison to conventional supervised classification.

427   *Photogrammetric Engineering and Remote Sensing* 56(4):469-473.

428   Immitzer M,Vuolo F, Atzberger C. 2016. First Experience with Sentinel-2 Data for Crop and

429   Tree Species Classifications in Central Europe. *Remote Sensing* 8:1-27. DOI:

430   10.3390/rs8030166.

431   INEGI. 2013. Conjunto de datos vectoriales de uso de suelo y vegetación escala 1:250 000, serie

432   V. Instituto Nacional de Estadística y Geografía. Aguascalientes. *Available at*:

433   http://www.conabio.gob.mx/informacion/gis/ (accessed 10 September 2015).

434   Klein T, Hoch G, Yakir D, Körner C. 2014. Drought stress, growth and nonstructural

435   carbohydrate dynamics of pine trees in a semi-arid forest. *Tree physiology* 34:981-992.

436   DOI: 10.1093/treephys/tpu071.

437   Kun J, Xiangqin W, Xingfa G, Yunjun J. Xianhong X, Bin L. 2014. Land cover classification

438   using Landsat 8 Operational Land Imager data in Beijing, China. *Geocarto International*

439   29:941-951. DOI:10.1080/10106049.2014.894586.

440   Lanner RM, Van Devender TR. 2000. The recent history of pinyon pines. In: Richardson, D. M.

441   (eds). The American Southwest, Cambridge University Press. 171–182

442   Léon-Portilla. 1988. Miguel del Barco, Historia natural y crónica de la antigua California.

443   Universidad Nacional Autónoma de México, México. 483 p.

444   Madonsela S, Cho MA., Ramoelo A, Mutanga O. 2017. Remote sensing of species diversity

445   using Landsat 8 spectral variables. *ISPRS Journal of Photogrammetry and Remote Sensing*

446   133: 116–127. DOI: 10.1016/j.isprsjprs.2017.10.008.

447 Marston, RA. 2010. Geomorphology and vegetation on hillslopes: interactions, dependencies,

448 and feedback loops. *Geomorphology*, *116*(3-4), 206-217.

449 Moran RV. 1983. Relictual northern plants on peninsular mountain tops. In: Biogeography of the

450 Sea of Cortez; University of California Press, Berkeley, USA. 408–410.

451 Olofsson O, Foody GM, Stehman SV, Woodcock CE. 2013. Making better use of accuracy data

452 in land change studies: Estimating accuracy and area and quantifying uncertainty using

453 stratified estimation. *Remote Sensing of Environment* 129:122–131. DOI:

454 10.1016/j.rse.2012.10.031

455 Olofsson, P, Foody, GM, Herold, M, Stehman, SV, Woodcock, CE, Wulder, MA. 2014. Good

456 practices for estimating area and assessing accuracy of land change. *Remote Sensing of*

457 *Environment* 148, 42-57. DOI: 10.1016/j.rse.2014.02.015

458 Osem Y, Zangy E, Bney-Moshe E., Moshe Y, Karni N, Nisan Y. 2009. The potential of

459 transforming simple structured pine plantations into mixed Mediterranean forests through

460 natural regeneration along a rainfall gradient. *Forest Ecology Management* 259:14–23.

461 DOI:10.1016/j.foreco.2009.09.034.

462 Pettorelli N. 2013. The Normalized Difference Vegetation Index. Oxford, University Press.

463 United Kingdom. 194 p.

464 QGIS Development, 2016. QGIS Geographic Information System. Open source Geospatial

465 Foundation. Available at: http//qgis.osgeo.org

466   Radoux J, Chomé G, Jacques DC, Waldner F, Bellemans N, Matton N, Lamarche C,

467   d'Andrimont R, Defourny P. 2016. Sentinel-2's potential for sub-pixel landscape feature

468   detection. *Remote Sensing* 8(6):488. DOI:10.3390/rs8060488.

469   Rehfeldt GE, 2006. A spline model of climate for the Western United States. Gen Tech Rep.

470   RMRS-GTR-165. U.S. Department of Agriculture, Forest Service, Rocky Mountain

471   Research Station, Fort Collins, Colorado, USA.

472   Rehfeldt GE, Crookston NL, Warwell MV, Evans JS. 2006. Empirical analyses of plant-climate

473   relationships for the western United States. *International journal plant science* 167:1123–

474   1150. DOI: 1058-5893/2006/16706-0005.

475   Richards JA. 1999. *Remote Sensing Digital Image Analysis*, Springer-Verlag, Berlin, p.240.

476   Riley SJ, Degloria SD, Elliot R. 1999. A terrain ruggedness index that quantifies topographic

477   heterogeneity. *Intermountain Journal of Sciences*5:23–27

478   (http://arcscripts.esri.com/details.asp?dbid=12435).

479   Roberts N, Ezcurra E. Desert Climate. 2012. In: Rebman, JP, Roberts NC, ed. *Baja California*

480   *Plant Field Guide*. San Diego Natural History Museum. San Diego, USA. 1-23.

481   Rouse JW, Haas RH, Schell A, Deering DW. 1974. Monitoring vegetation systems in the Great

482   Plains with ERTS. Proceedings of the Third Earth Resources Technology Satellite-1

483   Symposium, December 10–15 1974, Greenbelt, MD, NASA, Washington, DC, pp.301–

484   317.

485 Sáenz-Romero C, Rehfeldt GE, Crookston NL, Duval P, St-Amant R, Beaulieu J, Richardson

486   BA. 2010. Spline models of contemporary, 2030, 2060 and 2090 climates for Mexico and

487   their use in understanding climate-change impacts on the vegetation. *Climatic Change*,

488   102:595-623. DOI:10.1007/s10584-009-9753-5.

489 Salas EAL, Valdez R, Michel S. 2017. Summer and winter habitat suitability of Marco Polo

490   argali in southeastern Tajikistan: A modeling approach. *Heliyon* 3(11):e00445.

491   DOI:10.1016/j.heliyon.2017.e00445.

492 Sappington, JM., Longshore, KM., Thompson, D. B. 2007. Quantifying landscape ruggedness

493   for animal habitat analysis: a case study using bighorn sheep in the Mojave Desert. *Journal*

494   *of wildlife management*, *71*(5):1419-1426. DOI: 10.2193/2005-723

495 Satage AR, Salas C. 2007. Interactions of Elevation, Aspect, and Slope in Models of Forest

496   Species Composition and Productivity. *Forest Science* 53:486-492. Available at:

497   http://www.ingentaconnect.com/

498 Silva-Flores R, Pérez-Verdín G, Wehenkel C. 2014. Patterns of tree species diversity in relation

499   to climatic factors on the Sierra Madre Occidental, Mexico. *PloS one 9,* e105034. DOI:

500   10.1371/journal.pone.0105034.

501 Shirk AJ, Waring K, Cushman S, Wehenkel C, Leal-Sáenz A, Toney C, Lopez-Sanchez CA.

502   2017. Southwestern white pine (*Pinus strobiformis*) species distribution models predict

503   large range shift and contraction due to climate change. *Forest Ecology Management* (in

504   review).

505 Sothe C, Almeida CMD, Liesenberg V, Schimalski MB. 2017. Evaluating Sentinel-2 and

506   Landsat-8 Data to Map Sucessional Forest Stages in a Subtropical Forest in Southern

507   Brazil. *Remote Sensing* 9(8):838. DOI:10.3390/rs9080838.

508 Spasojevic MJ, Bahlai CA, Bradley BA, Butterfield BJ, Tuanmu MN, Sistla S, Wiederholt R,

509   Suding KN. 2016. Scaling up the diversity-resilience relationship with trait databases and

510   remote sensing data: the recovery of productivity after wildfire. *Global Change Biology*

511   22(4):1421–1432. DOI: 10.1111/gcb.13174.

512

513 Tan, S.S., Smeins, FE. 1996. Predicting grassland community changes with an artificial neural

514   network model. *Ecological Modelling* 84(1-3): 91-97. DOI: /10.1016/0304-3800(94)

515   00131-6.

516 Tapias R, Climent J, Pardos JA, Gil L. 2004. Life histories of Mediterranean pines. *Plant*

517   *Ecology* 171: 53-68. DOI:10.1023/B:VEGE.0000029383.72609.f0.

518 Telespazio VEGA Deutschland GmbH 2016. Sentinel-2 MSI-Level-2A. Prototype Processor

519   Installation and User Manual. Available at:

520   http://step.esa.int/thirdparties/sen2cor/2.2.1/S2PAD-VEGA-SUM-0001-2.2.pdf

521 Tsujino R, Takafumi H, Agetsuma N, Yumoto T. 2006. Variation in tree growth, mortality and

522   recruitment among topographic positions in a warm temperate forest. *Journal of*

523   *Vegetation Science* 17:281-290. DOI:10.1658/1100-

524   9233(2006)17[281:VITGMA]2.0.CO;2.

525   Venables WN, Ripley BD. 2002. Modern Applied Statistics with S-Plus. Fourth Edition. New

526   York, Springer.

527   Vieira ICG, de Almeida AS, Davidson EA, Stone TA, de Carvalho CJR, Guerrero JB. 2003.

528   Classifying successional forests using Landsat spectral properties and ecological

529   characteristics in eastern Amazonia. *Remote Sensing of Environment* 87(4):470-481.

530   DOI:10.1016/j.rse.2002.09.002.

531   Wu W, Li AD, He XH, Ma R, Liu HB., Lv JK. 2018. A comparison of support vector machines,

532   artificial neural network and classification tree for identifying soil texture classes in

533   southwest China. *Computers and Electronics in Agriculture* 144:86-93. DOI:

534   10.1016/j.compag.2017.11.037.

Figure 1

Map of Sierra La Asamblea.

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

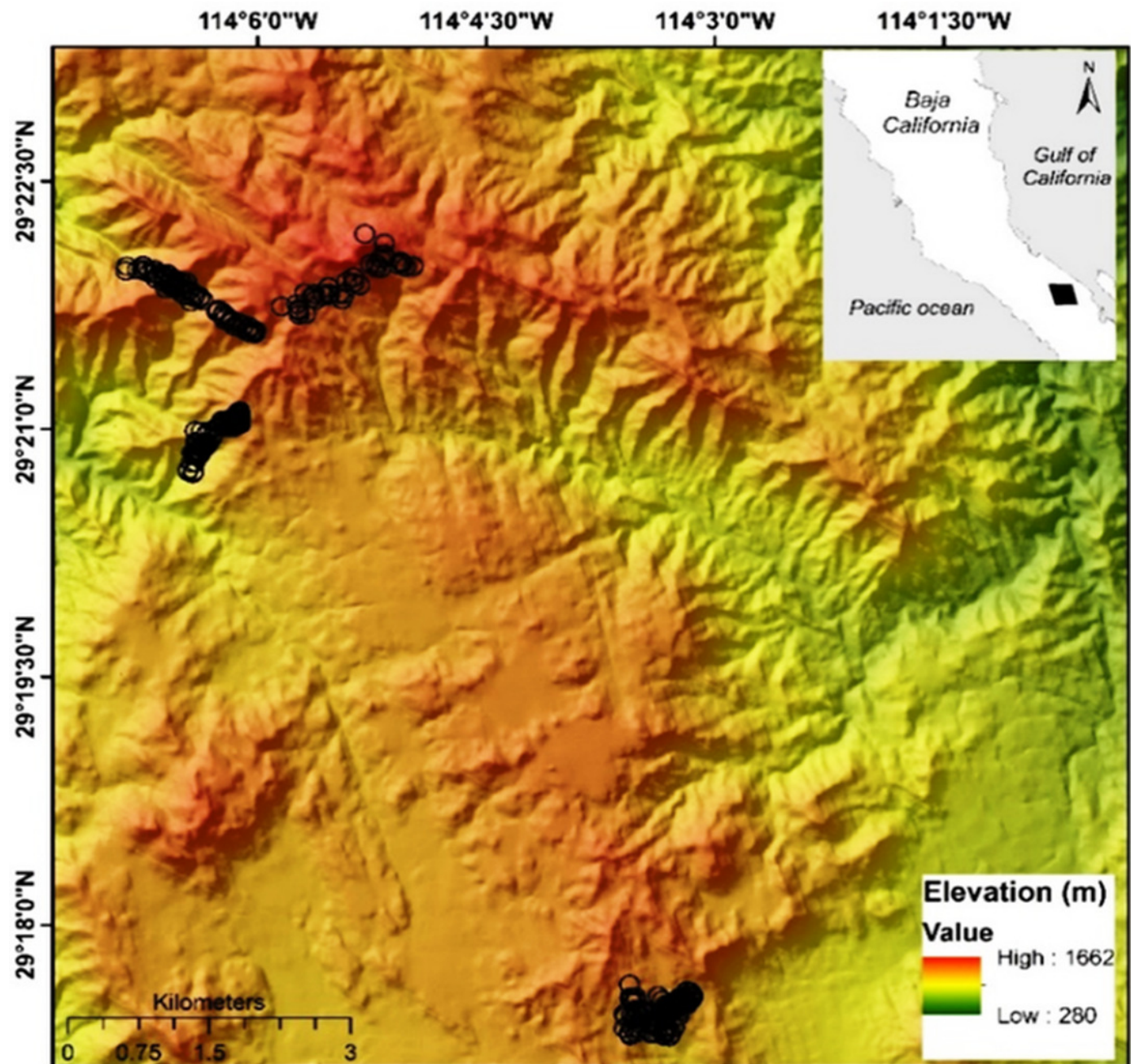



Figure 2

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

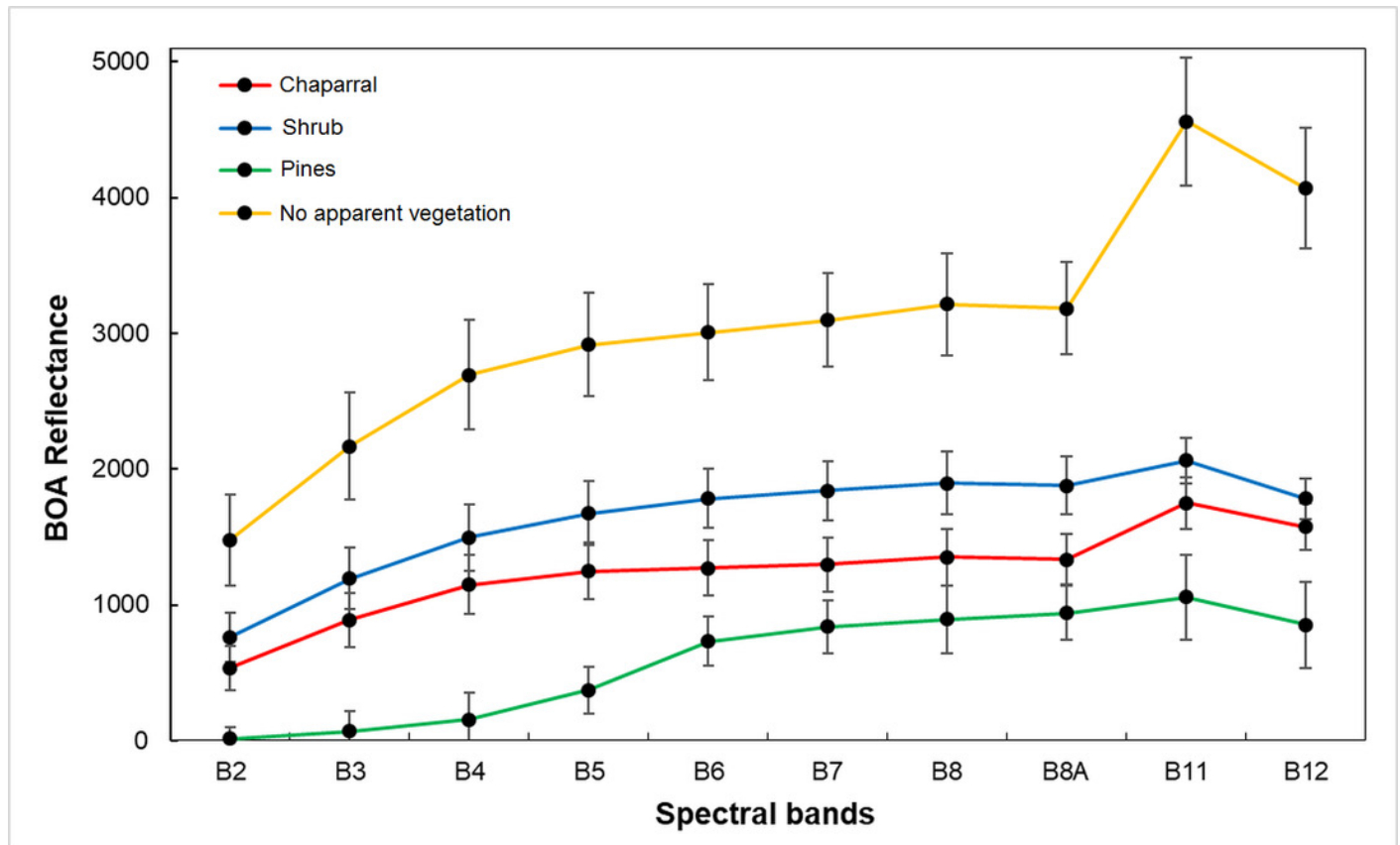
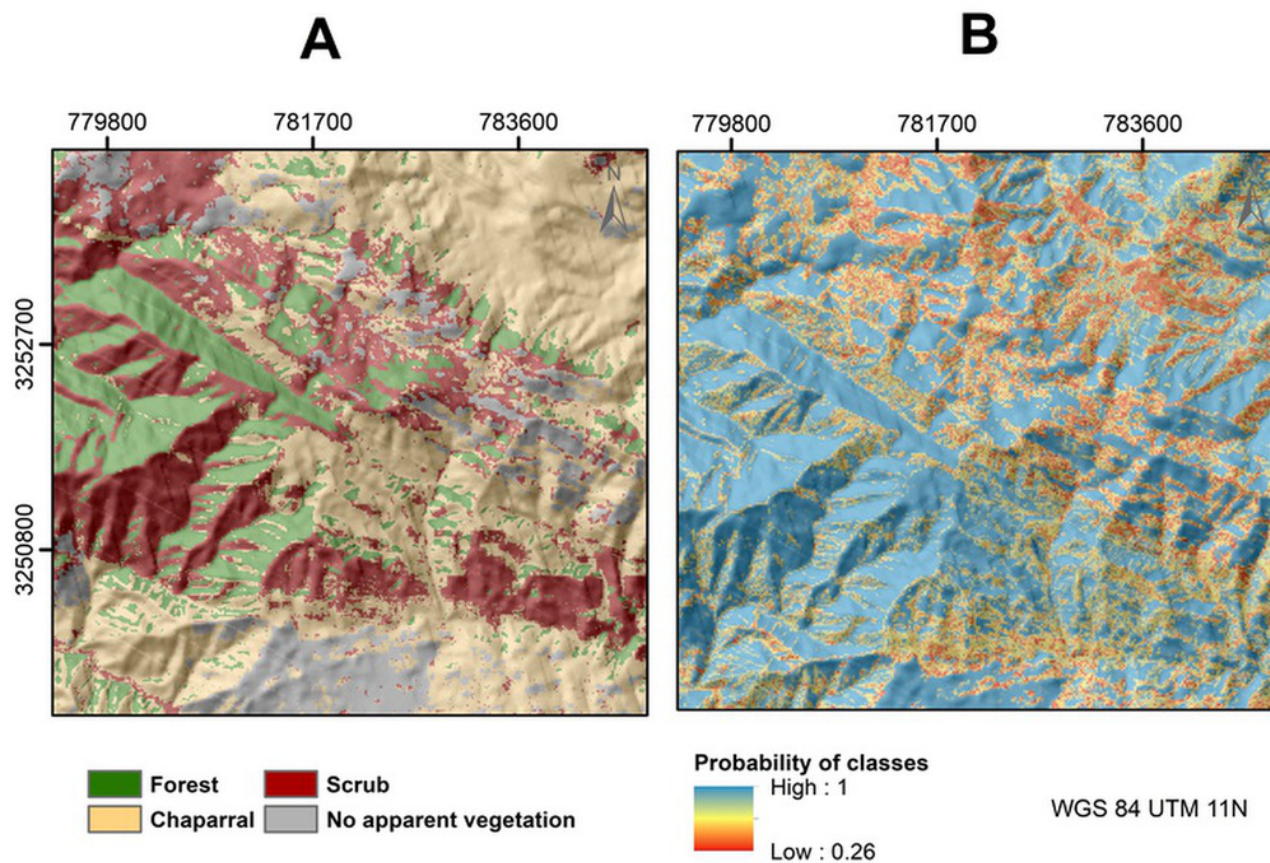



Figure 3

(A) Estimated land cover classes using BPNN classification in Sierra La Asambla. (B) Probability map of class assignment.

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

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
Ruggedness VRM	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	1302.41	124.96	1631	1010
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 *	GSDD5	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 *	Smrp	mm	43.74	6.29	62	29
Winter precipitation *	Winp	mm	110.93	7.93	133	93

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

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

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

	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
	WV	0	0	20	418	438	5,913	0.186
	Total	596	619	501	427	2,143	31,823	1

2 * Classes: P = *Pinus monophylla*; S = shrub; C = chaparral; WV= without vegetation; W_i =

3 proportion of the area mapped as class *i*.

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

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

	Classes	References				Accuracy		
		P	S	C	WV	User´s	Producer´s	Overall
Map	**P**	0.1651	0.0000	0.0044	0.0000	0.970±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	
	WV	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			

1   * Classes: P = *Pinus monophylla*; S = shrub; C = chaparral; WV= without vegetation.

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
Roughness TRI	0.178	0.015	11.200	< 2e-16
Roughness VRM	28.476	13.847	2.056	0.0397