Scientia Horticulturae 237 (2018) 287–295

Contents lists available at ScienceDirect

Scientia Horticulturae

journal homepage: www.elsevier.com/locate/scihorti

Potential geography and productivity of “Hass” avocado crops in Colombia estimated by ecological niche modeling

Joaquin Guillermo Ramírez-Gil[a,b,*], Juan Gonzalo Morales[a], A. Townsend Peterson[b,**]

[a] Departamento de Ciencias Agronómicas, Facultad de Ciencias Agrarias, Universidad Nacional de Colombia, sede Medellín, Calle 59ª #63-20, Bloque 11, Núcleo El Volador, Colombia

[b] Biodiversity Institute, University of Kansas, Lawrence, KS, 66045, USA

ARTICLE INFO

Keywords:
Sustainable agriculture
Avocado production
Model selection
Ecological niche models

ABSTRACT

Variation in environmental suitability across the geographic range of a species is common, with consequences for population health and production. Ecological niche modeling (ENM) is an effort to estimate environmental requirements of species based on associations between known geographic occurrences and environmental conditions at those sites, to allow estimation of the potential distribution of the species. We developed an ENM for Hass avocado production fields in Colombia, in which we related the distribution of avocado production lots to environmental information derived from multitemporal MODIS satellite imagery and digital elevation models. Models were calibrated using maximum entropy approaches, using likelihood and information content metrics to optimize model selection; model performance was measured in terms of omission error with respect to independent testing data. Two models were selected that presented best combinations of high information content, significance, and performance based on lower values of Akaike Information Criterion (AICc), partial receiver operating characteristic (partialROC), and low omission of independent testing data records. Suitability values from those models were compared to Hass avocado production; models were able to identify both novel areas with potential for cultivation, and sites where this crop has been planted in relatively unsuitable areas. A statistically significant positive relationship existed between suitability values from the models and avocado production, with different relationships for three technological levels. ENM approaches thus offer a useful and novel tool for exploring and designing sustainable agriculture systems, identifying areas with environmental potential in Colombia and currently not planted with Hass avocado.

1. Introduction

Avocado (*Persea americana* Mill) is a plant species of the family Lauraceae that includes three geographic races and cultivars resulting from hybridization of two or more races; these varieties are cultivated worldwide. The original distribution of this species was in tropical America, from Mexico through Central America to Colombia, Venezuela, Ecuador, and Peru (Galindo-Tovar et al., 2008). Hass avocadoes are the most common commercial cultivar in the world, and Mexico is the top producer country (FAO, 2017; Bernal and Díaz, 2014).

In Colombia, the avocado production system has grown greatly in recent years, thanks both to the excellent economic opportunities that it presents in foreign markets, and to a significant unsatisfied domestic demand (FAO, 2017; Bernal and Díaz, 2014; Ramírez-Gil et al., 2017a). Hass is the variety with the greatest cultivated area in moderately cold (temperate) climate zones in Colombia it is grown primarily for exportation (Bernal and Díaz, 2014; Ramírez-Gil et al., 2017a).

Despite the increased area planted with Hass avocadoes in Colombia in recent years, the production system has many limitations, owing to rapid growth not always with adequate technological backing, and to planting in areas not appropriate for this species. These problems have led to some investments ending in failure or low sustainability (Ramirez-Gil et al., 2014; Ramírez-Gil et al., 2017a). Avocado crops could thus benefit from a more detailed study of factors associated with success or failure, in search of a sustainable production system.

Ecological niche modeling (ENM) is an effort to estimate environmental requirements of species based on associations between known geographic occurrences of species and environmental characteristics of relevant landscapes, (ecological niches) identifying potential geographic distributions of species (Peterson et al., 2011). This approach

* Corresponding author at: Departamento de Ciencias Agronómicas, Facultad de Ciencias Agrarias, Universidad Nacional de Colombia, sede Medellín, Calle 59ª #63-20, Bloque 11, Núcleo El Volador, Colombia.
** Corresponding author.
E-mail addresses: j457r015@ku.edu, jgramireg@unal.edu.co (J.G. Ramírez-Gil), town@ku.edu (A.T. Peterson).

https://doi.org/10.1016/j.scienta.2018.04.021
Received 4 October 2017; Received in revised form 10 April 2018; Accepted 10 April 2018
0304-4238/ © 2018 Elsevier B.V. All rights reserved.

uses a heuristic device called the BAM diagram to understand causal factors, which considers geographic relationships between abiotic or uncoupled factors (**A**), biotic or coupled factors (**B**), and the movement or dispersal capacities of the species (**M**). This framework can be used to make explicit the possible arrangements of factors that determine distributions of species (Peterson et al., 2011), which have important implications for the success or failure of modeling endeavors (Saupe et al., 2012). Defining an appropriate area for calibrating the model (**M**) is crucial (Barve et al., 2011; Saupe et al., 2012). This area should be the area accessible to the species and that has been sampled, such that presence records can exist from those sites within the area that are suitable (Barve et al., 2011).

We applied ENM approaches to identify suitable areas for production of Hass avocadoes across Colombia, to test whether current production areas are the most appropriate ones, and to assess relationships between Hass avocado production and environmental dimensions. The result is a view of the geographic potential of this system across Colombia, and particularly of how avocado production systems could be optimized geographically across the country.

2. Materials and methods

2.1. Occurrence data and definition of calibration area (M)

We used 3950 records associated with Hass avocado production fields in Colombia. Data were gathered over a period of eight years (2009–2016) in Fitotecnia Tropical research group and were supplemented with information from state entities located in producing municipalities. We restricted analysis to avocado production areas of ≥ 1 ha. Each point was corroborated by visual inspection in Google Earth (Accessed 21 November 2016), eliminating sites that did not correspond to avocado fields (Fig. 1). Each data record consisted of coordinates (latitude, longitude), department, municipality, age (years), production (tons/ha), and technological level (low, medium, high, see below), although not all fields were populated with data in each case.

The area for calibrating the model (M) (Barve et al., 2011; Saupe et al., 2012) was focused in the Andean Region, at 1000–2500 m elevation, as those are the elevations at which Hass avocado can be established in Colombia. The region was further reduced to the limits of Colombia as sampling was only inside the country. Finally, we added a 50-km buffer zone, removing oceans and areas corresponding to other countries (Fig. 1).

2.2. Environmental variables

Because avocado production systems evaluated had an average age of 5 $\pm$ 3 yr, Colombian environmental characterization covered 8 yr (1 January 2009 – 31 December 2016). We used the Normalized Difference Vegetation Index (NDVI), in the form of 16-day composite images, with a spatial resolution of 250 m, from the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor (http://reverb.echo.nasa.gov). MODIS data were downloaded and reprojected using the MODIS Reprojection Tool (MRT) (https://lpdaac.usgs.gov/tools/modis_reprojection_tool_swath). Because MODIS data often include missing data caused by cloud cover, we filled these gaps via a spatial and temporal interpolation process implemented for this study by Huijie Qiao (Chinese Academy of Sciences). Other environmental variables were obtained from GMTED digital elevation model (DEM), with spatial resolution of 230 m (https://topotools.cr.usgs.gov/GMTED_viewer/). From the DEM, we obtained elevation (E), slope in degree (SD), slope in percent (SP), terrain ruggedness index (TRI) and topographic position index (TPI), using the ggplot, raster, and rgda packages in R (R Core Development Team, 2013).

To minimize the dimensionally of the NDVI data, we applied principal components analysis (PCA) using Spatial Analysis Tools in ArcGIS (Version 10.3). To reduce dimensionality further, correlations among environmental variables (including PCA variables) were obtained using R (R Development Core Team, 2013). We eliminated one among each pair of variables that presented Pearson correlation coefficients ≥ 0.6.

2.3. Model selection

To select models that were optimal in terms of both information content and predictive ability, we used a multi-step process. First, we used jackknife routines in the algorithm Maxent (version 3.3.3k; Phillips et al., 2006) to explore the relative contribution of each variable; we eliminated layers that showed low contribution to model predictions in an iterative series of assessments. We eventually reduced the set of variables to as few as 7 environmental variables, but we retained the final series of 9, 11, 13, etc., variables to allow exploration of more complex environmental spaces later in the model selection process.

We used model selection approaches to optimizing parameter settings in Maxent, with Akaike Information Criterion (AIC_c) values calculated in ENMTools (version 1.3) (Warren et al., 2010) from models calibrated with 100% of input points, 10 cross-validated replicates, and raw model outputs. We evaluated changes in the regularization multiplier (β), exploring values of 0.01, 0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9, 3, 4, 5, 10, 15, and 20; for each β value, we assessed the following combinations of response features L, LQ, LQP, LQPT, and LQPTH (L = linear, Q = quadratic, P = product, T = threshold, and H = hinge). As such, we generated 90 models, each based on 7 environmental variables: PCA1-5, slope in degrees (SD), and elevation (E). We inspected the results for parameter combinations that yielded minimum or near minimum values of AIC_c.

This initial exploration resulted in multiple minimum or near-minimum AIC_c values, but we had not yet considered inclusion of more environmental variables in the models. We decided, however, that it was not feasible to do a full examination of all combinations of β, response features, and environmental variables. Hence, we assessed a second group of 54 models, as the combination of 2 β values (1.5 and 4.0), 3 response types or feature combinations (LQPTH, LQPT, and LQP) and 9 combinations of environmental variables (7, 9, 11, 13, 15, 17, 19, 21, and 23), in the reverse order in which they were eliminated in the jackknife process described above. Performance of models in this second round of testing was evaluated using AIC_c as described above.

Finally, to extend our results from model selection to actual performance testing, we selected 6 ENMs that presented low AIC_c values, but with different parameter settings and environmental data layers. We created three subareas based on dividing the calibration area into eastern, western, and southern areas, with approximately equal numbers of occurrences available from each subarea (Fig. 1). We created combinations of 6 models and 3 subareas, and each pair of subareas was used to create a model that predicted the spatial distribution of occurrences in the third subarea. We used Maxent, with 100% of calibration points for training, 10 bootstrapped replicates, and logistic output format. Performance of predictions was evaluated using partial receiver operating characteristic (partialROC) approaches (Peterson et al., 2008), via functions available in Niche Toolbox (http://shiny.conabio.gob.mx:3838/nichetoolb2/). We used an acceptable omission rate of $E = 5\%$, and ran 1000 replicate analyses based on random subsamples of 50% of each testing data for each of the models generated. We determined the probability associated with the test as the proportion of replicates in which the partial ROC statistic was ≤ 1.

Finally, we calculated the omission rate (OR) for each of the predictions in independent subareas. We assessed OR via thresholding models on the calibration subareas, based on the highest model output value that would be associated with 5% omission of calibration data; once this threshold was imposed on the "other" subarea, and OR could be calculated based on the evaluated data. Hence, overall, from among the models with lower AIC_c values, we chose best models based on both

Download English Version:

https://daneshyari.com/en/article/8892593

Download Persian Version:

https://daneshyari.com/article/8892593

Daneshyari.com