

Modeling Land Suitability for *Coffea arabica* L. in Central America

Leonel Lara Estrada ^{1,2,*}, Livia Rasche ¹ and Uwe A. Schneider ¹

¹ Research Unit Suitability and Global Change, Center for Earth System Research and Sustainability, University of Hamburg. Grindelberg 5, 20144 Hamburg, Germany; livia.rasche@uni-hamburg.de, uwe.schneider@uni-hamburg.de

² School of Integrated Climate System Sciences, CLISAP, Grindelberg 5, 20144 Hamburg, Germany.

* Corresponding author: Leonel_larae@hotmail.com; Tel.: +49-40-42838-6593

Abstract

Agricultural planning processes at the farm to national level are essential for assessing and reacting to land conditions, opportunities, and threats for coffee production. However, the lack or uncertainty of information is common during these processes. Bayesian networks can be used to manage these uncertainties. We, therefore, developed the first Bayesian network model for an Agroecological Land Evaluation for *Coffea arabica* L. (ALECA). A newly developed set of suitability functions was used to populate the nodes in the network. ALECA was then adjusted and validated in Central America. The results show that even without the use of coffee maps as input, ALECA accurately scores the suitability of actual coffee areas for coffee production as higher than that non-coffee areas, and can accurately predict the known order of quality of coffee reference zones in Central America. The results also show that ALECA can be used as a decision-support tool even under data uncertainty.

Keywords:

Land evaluation, Ecological modeling, Suitability functions, Bayesian networks, Integrated modeling framework

Software Availability

Name: Agroecological Land Evaluation for *Coffea arabica* L. (ALECA)

Developer: Leonel Lara Estrada

Software requirements: Netica (www.Norsys.com)

Model availability: Upon request at Leonel_larae@hotmail.com

1. Introduction

The agricultural sector faces the challenge of producing enough goods for a growing population. This problem is exacerbated by changes in climate and the depletion of soil and water, factors that determine the suitability of land for agricultural production (Godfray et al., 2010). Information about the spatiotemporal variations of these factors is necessary for efficient agricultural planning processes at the farm, local and regional scales. Agricultural planning involves the inventory and classification of available resources to define the biophysical potential of land for crop production, called land suitability evaluation. Most land suitability evaluation systems assess climatic, soil and landform factors, while others also include anthropogenic factors such as local production systems, relevant cultural customs, policies, etc. (Littleboy et al., 1996; McRae and Burnham, 1981; Verheye, 1987).

Commonly, land suitability evaluation systems are used at national or regional scales to generate spatial representations of land suitability for different crops or animal production systems. However, they can also be used at local or farm level by farmers and other stakeholders themselves. The advantage, in this case, is that a more accurate evaluation can be performed due to the availability of more detailed farm data, but often the expertise required to operate the models is lacking (Jakeman et al., 2006). Another drawback is that the majority of land suitability evaluation systems provide limited options to deal with uncertainty, which is a common feature in land evaluation processes, and environmental modeling in general (Refsgaard et al., 2007). Bayesian networks offer a solution to this problem: They have the ability to manage uncertainty (of data and knowledge), integrate expert knowledge, and combine qualitative and quantitative information. Bayesian networks can also integrate complex systems from different domains and aggregate model dimensions to the level required, making them a suitable tool for ecological modeling and land suitability evaluations (Aguilera et al., 2011; Chen and Pollino, 2012; Poppenborg and Koellner, 2014).

Land suitability evaluation systems are especially relevant for managing risks in cropping systems with perennials, as the initial investments are usually higher, there is a waiting period of 3-4 years before the first harvest can be brought in, and resources are tied up for longer periods of time. One such example is coffee (*Coffea arabica* L.). Many smallholder farmers' livelihood in coffee-producing countries depends on its cultivation. Its production is reliant on a number of biophysical factors, such as climate, soil, landform, genetics and farming practices, whose relevance varies from site to site (Bertrand et al., 2011; Hagggar et al., 2011; Silva et al., 2013; Wang et al., 2015). In recent years, coffee producers have experienced a series of income losses due to market failures, pests and

diseases, the depletion of resources, extreme weather events, and a changing climate (Avelino et al., 2015; Tucker et al., 2010; Vega et al., 2003).

In light of the recent shocks and changing biophysical factors, it is desirable to develop mechanisms to describe and quantify the impacts of these changes on land suitability for coffee production. Some authors used species distribution models to explore the climatic suitability for coffee at regional and worldwide levels (Bunn et al., 2015, 2014; Chemura et al., 2015; Ovalle-Rivera et al., 2015). These models use coffee maps as presence data together with climate information to predict climatic suitability for coffee production. This approach assumes that climate is the only explanatory variable of coffee presence in a given location, and farmers' actions to improve climatic conditions or the influences of legal and socioeconomic factors are mostly excluded. Others, like D'haeze et al. (2005), explored the soil suitability (under the same climate conditions) for Robusta coffee using the Automated Land Evaluation System software based on the Framework for Land Evaluation of FAO (1976). Both studies focus on only one aspect of the environment, yet the response of species to environmental conditions is better explained by combining soil, topographical and climatic variables. While soil and topography are broadly aligned to climate, in practice, different soil types e.g. can be found under the same climatic conditions and vice versa (Coudun et al., 2006; Franklin and Miller, 2009). Mighty (2015) and Nzeyimana (2014) thus used multi-criteria analyses including climate, soil and landform variables in geographical information systems to evaluate land suitability for coffee in Jamaica and Rwanda, respectively.

All of these studies either look at a variety of factors but exclude uncertainty, or include uncertainty in their assessment but focus only on either soil or climate as the determining factor for coffee suitability. To date, no land suitability evaluation system includes all relevant factors for coffee production as well as uncertainty and has the potential to be used by scientists as well as stakeholders in coffee production. We, therefore, present in this paper the first Bayesian network model for the Agroecological Land Evaluation for *Coffea arabica* L. (ALECA) in Central America.

2. Materials and Methods

2.1. Study area

The study area encompasses the region of Central America spanning latitudes from 7.21° to 18.4°N and longitudes from 92.21 to 77.16°W, covering the countries of Guatemala, Belize, El Salvador, Honduras, Nicaragua, Costa Rica and Panama (Fig. 1). The area experiences tropical to subtropical conditions with a strong influence from the Atlantic Ocean. Dry winters and wet summers characterize the regional climate, whereby Atlantic

regions experience more rainfall and higher humidity (>80%) than Pacific ones (Taylor and Alfaro, 2005). Average variations in mean air temperature during the year are minimal (<4°C) and decrease from North to South, but maybe larger locally due to changes in topography (Magaña et al., 1999; Taylor and Alfaro, 2005). Due to the low variability in temperature, precipitation is the climatically more important determinant (Taylor and Alfaro, 2005). Even under future climate scenarios, where temperatures are expected to increase by 3 to 4°C, the projected decrease in precipitation and increase in precipitation variability is regarded as the major threat to the region (Giorgi, 2006; Karmalkar et al., 2011).

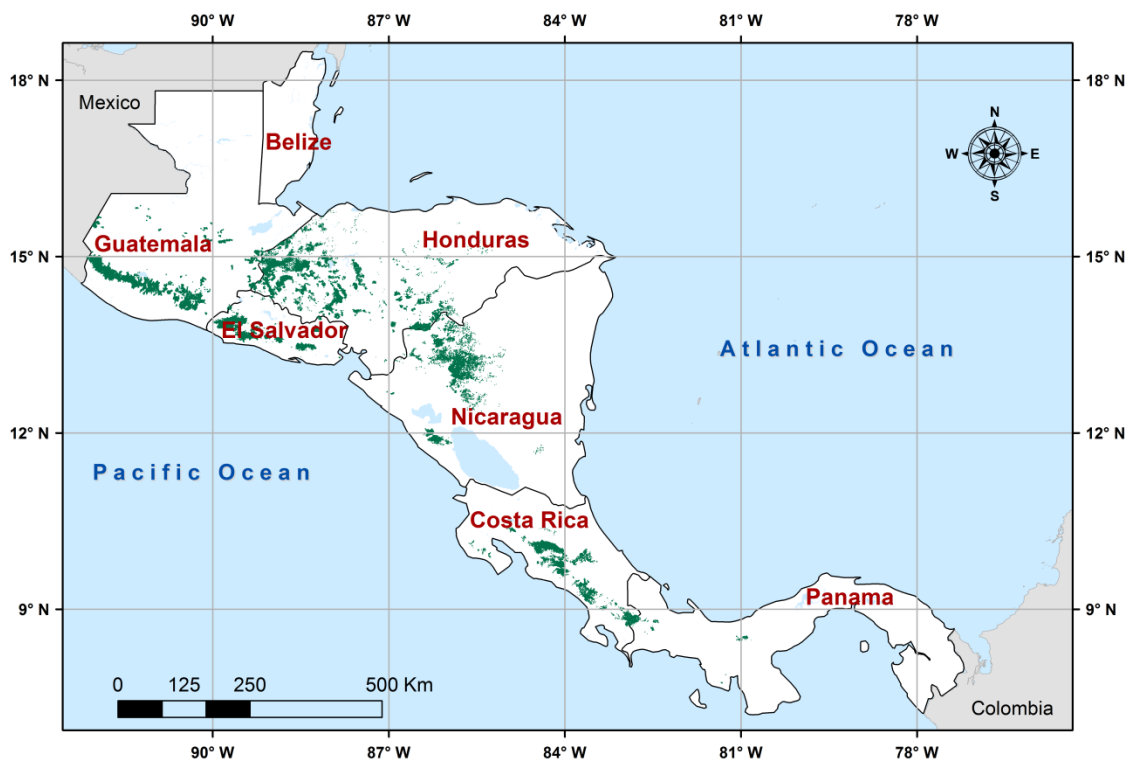


Fig. 1. Study area in Central America. Coffee areas are marked green¹.

Cultivation of *Coffea arabica* L. in Central America started in the 18th century in the Pacific region (Ukers, 1935). In some aspects, the cultivation systems differ little in the area (Samper, 1999); most of the coffee is cultivated in agroforestry systems, where trees are planted in coffee plantations to generate goods and services to the coffee plants and farmers (Somarriba et al., 2004), and 99% of coffee grown is *Coffea arabica*, with 1% *Coffea canephora* (Robusta) in Guatemala (USDA, 2015). Currently, nearly one million hectares

¹ Guatemala: 2010, Ministerio de Agricultura y Ganadería
 El Salvador: 2010, Ministerio de Agricultura y Ganadería
 Honduras: 2013, Instituto Nacional de Conservación y Desarrollo Forestal
 Nicaragua: 2012, Ministerio Agropecuario y Forestal
 Costa Rica: 2013, Instituto del Café de Costa Rica
 Panamá: 2012, Instituto Nacional de Estadísticas y Censo

are used for coffee production in Central America, generating 8 to 11% of the worldwide coffee supply (ICO, 2015).

2.2. Model development

2.2.1. Selection of variables

We started with a literature review to identify the key agroecological variables influencing coffee cultivation. Generally, coffee plants are sensitive to climate, soil conditions and farming practices (Camargo, 2010; Wang et al., 2015), in the literature described by the variables altitude, mean annual temperature, mean annual maximal temperature, annual mean precipitation, dry season length, relative humidity, wind speed, pH in H₂O, cation exchange capacity, soil organic carbon content, bulk density, soil texture, slope and aspect (Alégre, 1959; Descroix and Snoeck, 2004; Silva et al., 2013). From this list, we excluded altitude for our model, which is typically used as a proxy to define optimal and suboptimal climatic conditions for coffee (Avelino et al., 2005; Vaast et al., 2008), but which is only a substitute for temperature, which we use directly, wind speed and relative humidity due to gaps in data and information about its influence on coffee suitability, mean annual maximum temperature due its high correlation to mean annual temperature (Pearson coefficient $r=0.96$ and $p<0.00001$), which we calculated for the region based on WorldClim data (Hijmans et al., 2005), and soil organic matter due to its only indirect influence on coffee plants. With this variable selection, we are in line with other authors' recommendations for agricultural land evaluation (McRae and Burnham, 1981; Sys et al., 1991) and coffee land evaluation (Descroix and Snoeck, 2004; Mighty, 2015; Nzeyimana et al., 2014).

For the selected factors, we compiled their optimal, suboptimal and unsuitable levels for coffee production (Table 1). Mean annual temperatures of 18 to 21°C are thought to be ideal for coffee production (Alégre 1959). Temperatures above this range hasten the ripening of the coffee berry flesh before complete bean maturity is reached and consequently, the coffee quality declines (Vaast et al., 2006). Also, plant growth is reduced, and vegetative abnormalities start to occur at too low or high temperature (Camargo, 1985 & Franco, 1958 cited by DaMatta and Ramalho, 2006). For precipitation, Wallis (1963) estimated water requirements of 951 mm year⁻¹, in practice, as much as 1500 – 2000 mm are desirable. Precipitation higher than 2500 mm year⁻¹ can lead to waterlogging, a boost in fungal diseases, premature berry droppings, and ineffective fertilizer applications, amongst others (ANACAFE, 2006; Willson, 1985). Concerning the distribution of precipitation over the year, a dry season of 3 to 4 months is ideal to stimulate the main flowering and harvesting season in Central America. In regions closer to the equator, like Colombia, two short dry periods occur in a year facilitating two harvesting seasons. Longer dry seasons can lead to flowering

and fruit abortions, and quality and yield decline (Cannell, 1985; Willson, 1985), which is why coffee farmers identified precipitation (droughts and excessive rainfall) as the primary climatic concerns in the region, followed by temperature (Eakin et al., 2014; Tucker et al., 2010).

Soil texture affects multiple soil properties that influence soil fertility and crop productivity. Since coffee productivity is sensitive to nutrient and water supply, sandy and heavy clay soils are avoided for their limitations in water and nutrient holding capacity and drainage. The pH and cation exchange capacity are critical indicators of nutrient availability in the soil. Coffee plants prefer slight to medium acidic soils (5.0 to 6.2) with a high cation exchange capacity. Unlike soil texture, pH and cation exchange capacity can be modified by farming practices, such as chemical fertilization, the addition of organic matter, burnings, etc. (Descroix and Snoeck, 2004; Osman, 2013). Slope defines the vulnerability of a site to erosion and determines the potential for mechanization. Thus, flat or low slopes are optimal, as steep slopes require major soil conservation practices and reduce the efficiency of farming practices (Descroix and Snoeck, 2004; Ramírez, 2009). The aspect influences temperature variations at the microclimate level; in northern latitudes, south-facing slopes receive more sunlight than north-facing ones (Adams, 2010; Bonan, 2008) and Avelino (2005) found that coffee cultivated on east-facing slopes produced a better coffee quality in two locations in Costa Rica. Philpott et al. (2008) discovered significantly higher landslides in southwestern-facing slopes during a hurricane in Mexico.

Table 1. Agroecological variables selected to describe coffee land suitability in Central America with unsuitable, suboptimal and optimal values as reported in the literature.

Component	Variables	Unit	Unsuitable	Suboptimal	Optimal
Climate	Mean annual temperature	°C	$\leq 10^{1,2}$, $\geq 30^{3,6}$, $> 32^2$	$< 15-16^3$, $< 17-18^4$ $> 23^{3,4}$, $> 26^5$	$18-21^3$, $18-23^4$
	Annual precipitation	mm	$< 1000^7$	$< 1300^8$, $> 3000^{7,9}$	$1550-2000^{10}$, $1600-1800^3$
	Dry season length	# months ≤ 60 mm	$> 6^{11}$	$< 2^{12}$, $5-6^{11}$	$2-4^{13}$, $3-4^{14}$
Soil	pH in H ₂ O	-	$< 4^8$, $> 8^8$	< 5 , > 6.5	$5.5-6.5^9$, $5.2-6.2^{14}$
	Cation exchange capacity	Meq 100g ⁻¹		$< 5^{15}$	$> 22^{15, 16}$
	Texture	Categorical	Sand ($> 30\%$), heavy clay (clay $> 70\%$) ^{12,13}		Loam, clay loam, clay
Landform	Slope	%	$> 50^9$, $> 70^{17}$	$> 40^7$	$0-40^7$
	Aspect	Cardinal directions			East ¹⁸

¹ Larcher (1981), ² Jaramillo and Guzmán (1984), ³ Alégre (1959), ⁴(1985) cited by DaMatta and Ramalho (2006), ⁵Nunes et al. (1973) cited by DaMatta and Ramalho (2006), ⁶ DaMatta and Ramalho (2006), ⁷ ANACAFE (2006), ⁸ Willson (1985), ⁹ Descroix and Wintgens (2004), ¹⁰Forestier (1969) cited by Willson (1985), ¹¹ Descroix and Snoeck (2004), ¹² Maestri and Barros (1977), ¹³ Haarer (1958), ¹⁴Robinson 1964 cited by Willson (1985), ¹⁵ Molina and Melendez (2002), ¹⁶ Verheye (2002), ¹⁷ Blanco and Aguilar (2015), ¹⁸Avelino et al. (2005).

2.2.2. Suitability functions

We created suitability functions, i.e. response curves, for the selected variables based on the agroecological requirements of coffee, and the literature listed in Table 1. We used two kinds of functions: tables for discrete variables (texture and aspect) and equations for continuous variables. To create the equations, we first defined the suitability scores for the variable values based on literature, secondly graphed the suitability scores for each variable and finally identified the functions that best fit each graph (Fig. 2). In the case of tabular suitability functions, for soil texture, we defined the suitability scores from literature. For aspect, we used a survey of 600 coffee farms in Nicaragua (Nitlapan, 2012) to define the suitability of each cardinal direction by using an analysis of covariance, a mean separation test, and a weighted mean.

In the case of wild species, these types of functions describe the relationship of a species' occurrence in relation to values of an environmental condition (Austin, 1980; Franklin and Miller, 2009). In our study, the suitability functions describe how suitable the value of a variable is for coffee production, ranging from 0 to 100% (Fig. 2 & Table 2).

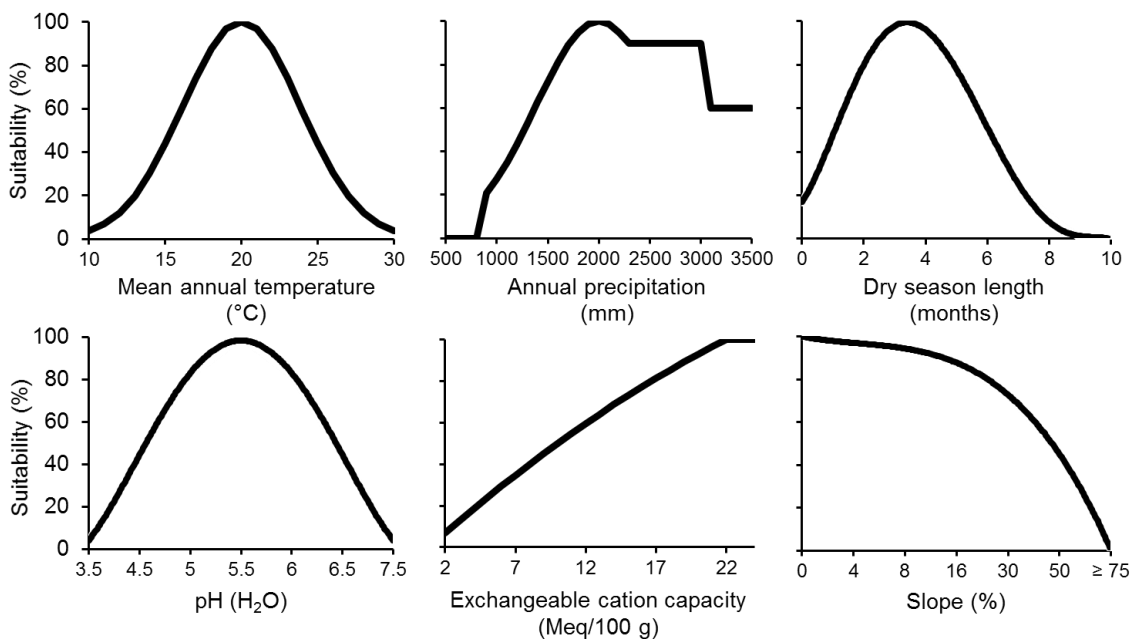


Fig. 2. Graphical display of the suitability functions for continuous variables.

Table 2. Suitability functions of the selected agroecological variables. Suitability scores range from 0 to 100%. The values in the texture table denote that there is a probability of 99% that ‘Sand’ e.g. only has suitability of 0-25% for coffee production and a 1% probability that it is in the range of 25-50%.

Variables	Equations																																																																																			
Mean annual temperature	$S_{ti} = T_i \sim N(\mu, \sigma^2) \div T_{\mu} \sim (\mu, \sigma^2) \cdot 100$ <p>Where S_{ti} is the suitability score for a given annual mean temperature in °C (T_i) and is distributed normally with mean $\mu = T_{\mu} = 20$ and variance $\sigma^2 = 3.89$</p>																																																																																			
Annual precipitation	$S_{pi} = \begin{cases} 0, & \text{if } P_i < 800; \\ P_i \sim N(\mu, \sigma^2) \div P_{\mu} \sim (\mu, \sigma^2) \cdot 100, & \text{if } P_i < 2300; \\ 90, & \text{if } P_i \leq 3000; \\ 60, & \text{Otherwise} \end{cases}$ <p>Where S_{pi} is the suitability score for given annual precipitation in mm (P_i) and is distributed normally with mean $\mu = P_{\mu} = 2000$ and $\sigma^2 = 620.48$</p>																																																																																			
Dry season length	$S_{di} = \begin{cases} 0, & \text{if } D_i > 8; \\ 0.252D_i^4 - 3.828D_i^3 + 14.149D_i^2 - 1.458D_i + 60, & \text{Otherwise} \end{cases}$ <p>Where S_{di} is the suitability score for a given dry season length in months (D_i)</p>																																																																																			
Slope	$S_{si} = 0.01 S_i^2 - 2S_i + 100$ <p>Where S_{si} is the suitability score for a given slope in percentage (S_i)</p>																																																																																			
Aspect	$S_{ai} = \begin{cases} 80, & \text{if } A = \text{Flat}; & 90, & \text{if } A = \text{West}; \\ 72, & \text{if } A = \text{Southeast}; & 93, & \text{if } A = \text{Southwest}; \\ 80, & \text{if } A = \text{Northeast}; & 97, & \text{if } A = \text{Northwest}; \\ 80, & \text{if } A = \text{East}; & 100, & \text{if } A = \text{South}; \\ 87, & \text{if } A = \text{North}; & & \end{cases}$ <p>Where S_{ai} is the suitability score for a given slope aspect (A)</p>																																																																																			
pH (H ₂ O)	$S_{phi} = pH_i \sim N(\mu, \sigma^2) \div pH_{\mu} \sim (\mu, \sigma^2) \cdot 100$ <p>Where S_{phi} is the suitability score for a given pH (pH_i) and is distributed normally with mean $\mu = pH_{\mu} = 5.5$ and $\sigma^2 = 0.79$</p>																																																																																			
Cation exchange capacity	$S_{ceci} = \begin{cases} 0, & \text{if } CEC_i < 1 \\ 100, & \text{if } CEC_i \geq 22 \\ -0.061CEC_i^2 + 6.114CEC_i - 5.053, & \text{Otherwise} \end{cases}$ <p>Where S_{ceci} is the suitability score for a given cation exchange capacity in meq/100g (CEC_i)</p>																																																																																			
Texture	<table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th rowspan="2">Texture</th> <th colspan="5">Suitability (%)</th> </tr> <tr> <th>0 to 25</th> <th>25 to 50</th> <th>50 to 75</th> <th>75 to 90</th> <th>90 to 100</th> </tr> </thead> <tbody> <tr><td>Sand</td><td>99</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>Sandy loam</td><td>99</td><td>1</td><td>0</td><td>0</td><td>0</td></tr> <tr><td>Loamy sand</td><td>0</td><td>99</td><td>1</td><td>0</td><td>0</td></tr> <tr><td>Loam</td><td>0</td><td>0</td><td>1</td><td>99</td><td>0</td></tr> <tr><td>Silt loam</td><td>0</td><td>0</td><td>1</td><td>99</td><td>0</td></tr> <tr><td>Silt</td><td>0</td><td>1</td><td>99</td><td>0</td><td>0</td></tr> <tr><td>Sandy clay loam</td><td>0</td><td>1</td><td>99</td><td>0</td><td>0</td></tr> <tr><td>Clay loam</td><td>0</td><td>0</td><td>1</td><td>99</td><td>0</td></tr> <tr><td>Silty clay loam</td><td>0</td><td>0</td><td>1</td><td>99</td><td>0</td></tr> <tr><td>Sandy clay</td><td>0</td><td>1</td><td>99</td><td>0</td><td>0</td></tr> <tr><td>Silty clay</td><td>0</td><td>0</td><td>0</td><td>1</td><td>99</td></tr> <tr><td>Clay</td><td>0</td><td>0</td><td>0</td><td>1</td><td>99</td></tr> </tbody> </table> <p>Where S_{ti} is the suitability score for given soil texture (T_i)</p>	Texture	Suitability (%)					0 to 25	25 to 50	50 to 75	75 to 90	90 to 100	Sand	99	1	0	0	0	Sandy loam	99	1	0	0	0	Loamy sand	0	99	1	0	0	Loam	0	0	1	99	0	Silt loam	0	0	1	99	0	Silt	0	1	99	0	0	Sandy clay loam	0	1	99	0	0	Clay loam	0	0	1	99	0	Silty clay loam	0	0	1	99	0	Sandy clay	0	1	99	0	0	Silty clay	0	0	0	1	99	Clay	0	0	0	1	99
Texture	Suitability (%)																																																																																			
	0 to 25	25 to 50	50 to 75	75 to 90	90 to 100																																																																															
Sand	99	1	0	0	0																																																																															
Sandy loam	99	1	0	0	0																																																																															
Loamy sand	0	99	1	0	0																																																																															
Loam	0	0	1	99	0																																																																															
Silt loam	0	0	1	99	0																																																																															
Silt	0	1	99	0	0																																																																															
Sandy clay loam	0	1	99	0	0																																																																															
Clay loam	0	0	1	99	0																																																																															
Silty clay loam	0	0	1	99	0																																																																															
Sandy clay	0	1	99	0	0																																																																															
Silty clay	0	0	0	1	99																																																																															
Clay	0	0	0	1	99																																																																															

2.2.3. Modeling in Bayesian Networks

Bayesian Networks are multivariate statistical models that comprise two main components: First, a directed acyclic graph composed of a set of random variables $X = (X_1, \dots, X_n)$ linked by arcs, where the arc direction defines direct dependencies between variables (parent and child nodes). Each variable has at least two mutually exclusive states. Second, a conditional probability distribution (conditional probability table) that quantifies the dependencies between variables. To illustrate, assume that X_1 is the child variable of parent variables X_2, \dots, X_n , written as $P(X_1 | X_2, \dots, X_n)$, which expresses the probability of X_1 occurring given the values of X_2, \dots, X_n (Jensen and Nielsen, 2007b; Pearl, 1988). In our case, the combination of agroecological variables determines the level of suitability of a given piece of land for coffee cultivation. In a Bayesian network framework, this implies direct causal links between each variable and the variable land suitability, i.e. $P(\text{Land Suitability} | \text{variable}_1, \dots, \text{variable}_n)$. These multiple links create an exponential increment of model complexity due to the need to estimate the probability distributions for all possible combinations of the states of the agroecological variables (Landuyt et al., 2013). We avoided this complexity and at the same time increased the explanatory ability of the model by adding two intermediate levels (the S-variables and components) between each agroecological variable and the land suitability. This technique is known as *divorce* in graphical models and is commonly used to keep conditional probability tables tractable (Chen and Pollino, 2012; Jensen and Nielsen, 2007a). We built our model using the software Netica v.5.17 (Norsys Software Corp.).

The final graphical model (Fig. 3) displays three evaluation levels: variables, components (groups of variables), and final land suitability (groups of components). The model scores and aggregates the suitability of the agroecological variables at each level to obtain the final land suitability, which we defined as the potential of a given unit of land to be used for viable coffee cultivation. In the first level, we created a node with discretized states for each variable and defined the maximum and minimum values of the variables (Pollino et al., 2007) and the priors from data available for the region (Table 3) (Hengl et al., 2014; Hijmans et al., 2005). Then, we added a new child variable "S" to each agroecological variable and used the suitability functions to populate the conditional probability tables of the S-variables. In the second level, we grouped and aggregated the S-variables into the components climate, soil, and landform by using the Linear Combination Method (Hopkins, 2014). This method is a simple weighted sum of factors. We assigned the same weight to all S-variables. In the third level, we again used the weighted sum method to aggregate the components into the final variable Land Suitability, but this time assigned different weights to each component (49% to climate, 36% to soil and 15% to landform). The weights were

calculated based on two coffee survey from Nicaragua (Lara Estrada, 2005; Nitlapan, 2012) using Pearson’s correlation coefficients between our selected variables and the coffee yields reported in the surveys.

After the structure of the model was determined, the priors of the agroecological variables were learned from data using the Counting-Learning Algorithm (Norsys, 2015) and the conditional probability tables of all child-nodes were populated using the described equations and weighted sums. Then the model was compiled and ready to use.

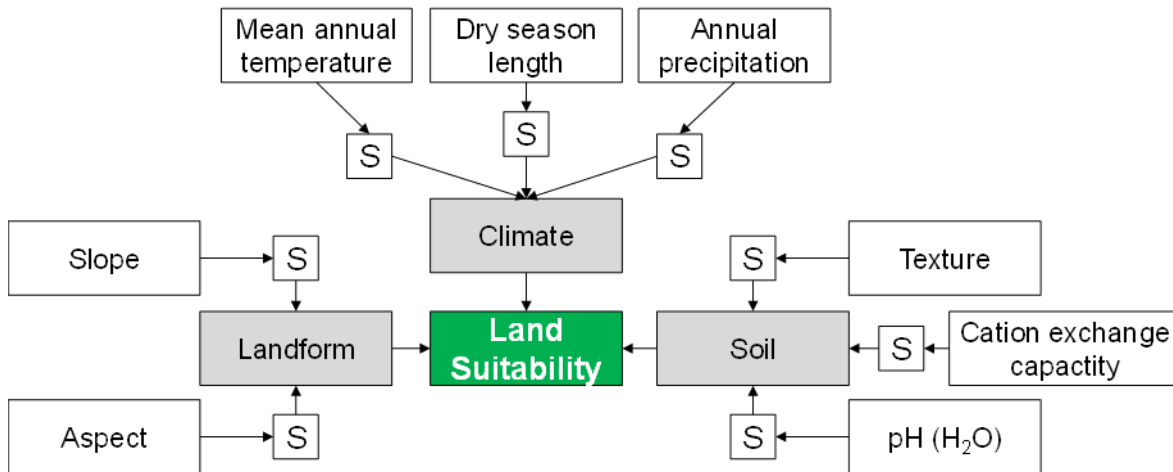


Fig. 3. Graphical structure of the Agroecological Land Evaluation model for *Coffea arabica* L. (ALECA). Arrows indicate causal relationships between variables (from parent to child nodes). Level 1 (white boxes): suitability functions [0-100%] are used in the “S” nodes to evaluate the suitability levels of the agroecological variables for coffee cultivation. Level 2 (gray boxes): the suitability scores of Level 1 are aggregated into components. Level 3 (green box): the components are aggregated and weighted to obtain the final land suitability score. The units of the S variables, the components, and the final land suitability score are percent.

Table 3. Description of the state values for the selected variables.

Variables	States*
Level 1	
Annual mean temperature [°C]	4-9.99, 10-11.99, [...], 26-27.99, 28-29.99
Annual precipitation [mm]	500-999.99, 1000-1249.99, [...], 2750-2999.99, 3000-6499.99
Dry season length [months]	0, 1, 2, 3, 4, 5, 6, 7, 8
Slope [%]	0-1.99, 2-3.99, [...], 16-29.99, 30-59.99, 60-99.99
Aspect [cardinal direction]	North, Northeast, East, Southeast, South, Southwest, West, Northwest, Flat
pH in H ₂ O [-]	2-2.49, 2.5-2.99, 3-3.49,[...], 7-7.49, 7.5-7.99
Cation exchange capacity [Meq 100 g ⁻¹]	2.5-4.99, 5-7.49, [...], 17.5-19.99, 20-22.49, ≥22.5
Texture [categorical]	Sand, Loamy sand, Sandy loam, loam, Silt loam, Silt, Sandy clay loam, Clay loam, Silty clay loam, Sandy clay, Silty clay, Clay
S variables [%]:	0-9.99, 10-19.99, [...], 80-89.99, 90-99.99

Annual mean temperature, Annual precipitation, Dry season length, Slope, pH in H ₂ O, Cation exchange capacity	
S variable Aspect [%]	70-79.99, 80-89.99, 90-99.99
S variable Texture [%]	0-24.99, 25-49.99, 50-74.99, 75-89.99, 90-99.99
Level 2	
Component suitability [%]: Climate, Landform and Soil	0-9.99, 10-19.99, [...], 70-79.99, 90-99.99
Level 3	
Land Suitability [%]	0-9.99, 10-19.99, [...], 70-79.99, 90-99.99

2.3. Data sources

For the simulations, we used the WorldClim dataset (Hijmans et al., 2005) to extract precipitation and temperature and calculate dry season length. WorldClim provides climate data in raster format from interpolated observations for the period 1950 to 2000 and also includes digital elevation [90 m aggregated at 1 km] from the Shuttle Radar Topography Mission (SRTM). We used this data to estimate the slopes and aspect [0 to 360°], which was converted to cardinal directions (ESRI, 2008). Sand, silt and clay content, cation exchange capacity and pH in H₂O were downloaded from the SoilGrids portal (www.soilgrids.org). SoilGrids is a worldwide 3D spatial dataset for chemical and soil physical properties at a 1 km resolution (Hengl et al., 2014). Soil data were available from 0 to 200 cm depth in layers of 5 cm; we used the average of the layers between 0 at 30 cm depth. Lastly, we merged all climate, landform and soil data into a single dataset, where each pixel [1 km resolution] corresponds to one “land case” to be evaluated in the model.

3. Sensitivity analysis

The influence of a parent variable on a child variable is defined by the prior distribution, the variable states (i.e. size and bounds) and the equations or conditional probability tables inside the child variable (Bennett et al., 2013). To explore the influence of the chosen agroecological variables on land suitability, we ran a sensitivity analysis using the variance reduction metric. The higher the variance reduction value of a variable is (scoring from 0 to 100%), the greater its influence on land suitability (Marcot et al., 2006; Norsys, 2015). We also examined the influence of the single components' scores on land suitability.

The analysis shows that the components climate and soil and their respective agroecological variables (except texture) influence land suitability most, and that results are not very sensitive to the landform component. Inside the components, mean temperature plays the most important role in the climate, pH in the soil and slope in the landform component (Fig. 4). The low variance reduction of landform results a) from the low weight

derived to the landform component, and b) from the narrow distribution of slope and aspect in our gridded input dataset, where, due to the resolution, extremes in slope, which would restrict or prohibit coffee cultivation on some plots, are filtered out. Since ALECA was also developed to be applied to specific plots on farms, the inclusion of landform is nevertheless important. A flat or low slope is more desirable on a piece of land with excellent or good conditions of soil and climate (land suitability > 85%) than a steep slope and will thus be better ranked. For example, steep slopes facilitate soil erosion and the efficiency of agronomic practices (Okoth et al., 2007; Tilman et al., 2002).

For an explanation of how the components' scores interact to define the final land suitability, see appendix.

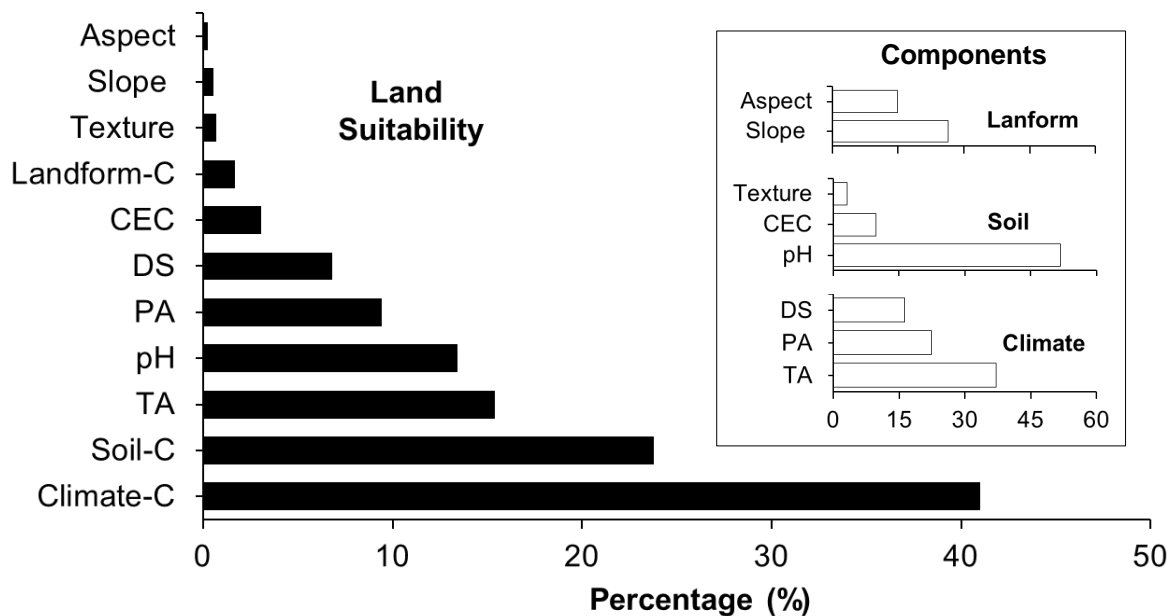


Fig. 4. Sensitivity analysis of model results using variance reduction for land suitability (left) and components (right). CEC: cation exchange capacity, DS: dry season length, PA: mean annual precipitation, TA: mean annual temperature.

4. Model validation

Validating a model like ALECA is difficult due to the nature of its results, which do not give data on the presence/absence of coffee, but provide a score of 0-100% indicating the suitability of a specific land unit for coffee cultivation. We nevertheless decided to compare our model results with regional data of the spatial distribution of coffee areas (Fig. 1), assuming that the distribution of current coffee areas is the result of a historical and technical selection for the best possible and available areas (Rueda and Lambin, 2013; Samper, 1999) and that consequently, ALECA should score their land units high on the suitability scale. Some plantations may be located in areas of lower suitability since farming practices

like the planting of shade trees improve land suitability by lowering temperatures (Hagggar et al., 2011; Siles et al., 2010), but no coffee plantations should be located on completely unsuitable areas.

The results of this exercise show that there is indeed a close fit between patterns of coffee areas reported in national coffee maps and areas scored as suitable for coffee production by our model (Fig. 5; maps of the entire region can be found in the supplementary material). Areas with coffee plantations have a mean land suitability score of 85% (SD=5.34), nearly 98% of the plantations are located in areas with a suitability higher 70%, only 2% of the areas are ranked 60-70%, and no coffee plantations have land suitability scores below 60% (Table 4). In comparison, non-coffee areas have a mean land suitability score of 76% (SD=7.7), with 80% located in areas with a suitability higher 70%, 17% in areas ranked 60-70%, and 3% in areas with suitability scores below 60%. We assume that the occurrence of coffee plantations in areas ranked 60-70% is the result of social and agronomical factors, such as planting coffee for land reclamation purposes (Charlip, 2003). Additionally, the use of farming practices like agroforestry systems and better adapted genetic material have permitted farmers to extend coffee production to sites with a lower suitability ranking (Lopez-Rodriguez et al., 2015; Muschler, 2001; Philippe et al., 2009). Guatemala, for example, the country with the highest proportion (54%) of coffee areas rated between 60 and 80% suitability, is also the only country in Central America with a significant Robusta production. This coffee species is better adapted to warmer conditions than *Coffea arabica* L. (USDA, 2015; Willson, 1985), indicating that the simulated lower suitability scores may be realistic. The relatively high proportion of non-coffee areas with scores higher 70% shows that there is great expansion potential in Central America for coffee production. It should be kept in mind, however, that non-coffee areas include urban and protected areas, coasts, roads, and other land uses that are not immediately available for plantations.

We conclude from this validation study that, taking local social and agronomical factors into account, ALECA performs well in terms of land suitability scoring of potential coffee areas.

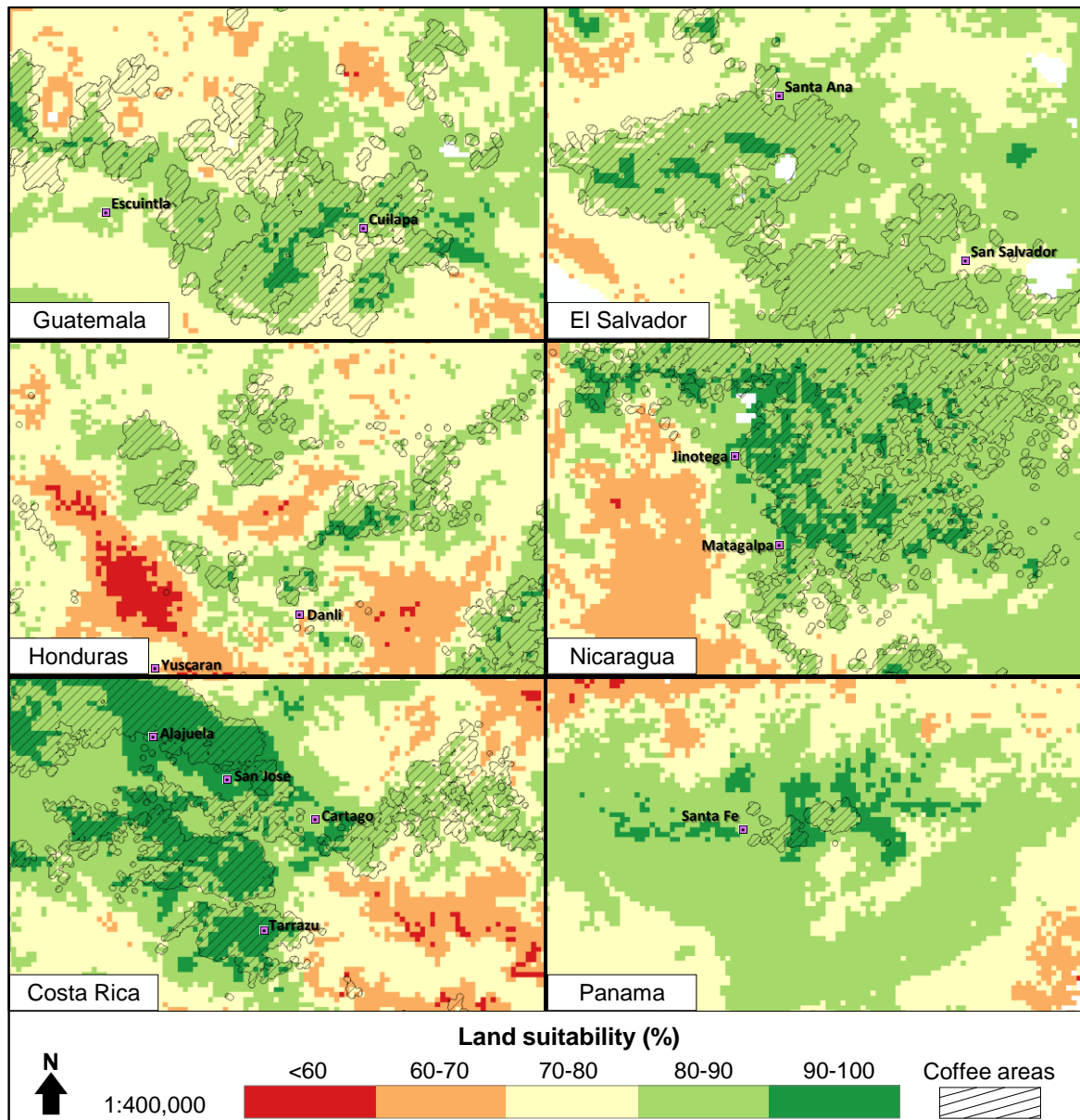


Fig. 5. Map of reported coffee areas and simulated land suitability scores in selected areas of Central America. Coffee is usually grown in areas with high estimated suitability, indicating a good model performance. Pixel size is 1 km.

Table 4. Simulated land suitability scores of current coffee and non-coffee areas for *Coffea arabica* L. in Central America with GUA = Guatemala, HON = Honduras, SV = El Salvador, NIC = Nicaragua, CR = Costa Rica, PAN = Panamá. A graphical presentation of the data can be found in the supplementary material.

Land suitability score (%)	Coffee areas								Non-coffee areas	
	GUA	HON	ESV	NIC	CR	PAN	∑ %	∑ km ²	%	km ²
<60	<0.01	<0.01	-	-	<0.01	-	<0.01	<1	3.20	15,776
60 - 69	1.90	0.06	<0.01	0.09	0.06	-	2.11	224	17.34	85,469
70 - 79	18.25	2.62	0.96	1.56	0.99	0.30	24.68	2,614	44.75	220,678
80 - 89	14.99	17.33	11.88	11.15	3.66	0.46	59.47	6,302	32.75	161,489
90 - 100	2.08	2.99	1.82	2.49	4.26	0.10	13.74	1,456	1.96	9,677
Total	37.22	23.00	14.66	15.29	8.97	0.87	100	10,598	100	493,089

In a second validation study, we compared the simulated land suitability scores of eight coffee reference zones with their conventionally reported suitability for coffee cultivation, hoping to show that ALECA simulates higher scores for prime coffee areas than for areas known to be of lesser suitability. The selected zones are located in Honduras, Nicaragua and Costa Rica and include the zones Marcala and El Paraíso in Honduras (Teuber, 2009), Jinotega, Masatepe and Nueva Guinea in Nicaragua (Haggar et al., 2011; Vaast et al., 2005), and Tarrazú, Turrialba and San Carlos in Costa Rica (Muschler, 2001; Siles et al., 2010). According to the literature and commonly held views, the region Tarrazú is considered as optimal, Jinotega and Marcala as very good, and Turrialba, Masatepe, El Paraíso, Nueva Guinea and San Carlos as suboptimal for coffee production (Haggar et al., 2011; Muschler, 2001; Rojas, 1989; Vaast et al., 2005).

We found that Tarrazú, the only region ranked optimal, was also the region simulated to have the highest mean land suitability-scores among the reference zones. The determining factors for the high score were mainly climate and soil conditions (Fig. 6), which is line with the assessments of other authors for Tarrazú (Bornemisza and Segura, 1999; Chinchilla et al., 2011; Rojas, 1989). Next in the land suitability ranking were Jinotega and Marcala, both also known in reality for their overall good production levels and coffee quality (IICA, 2003; Teuber, 2009). Jinotega scored slightly worse in the soil compartment and Marcala in the climate compartment than Tarrazú. Turrialba, Masatepe, and El Paraíso received suitability scores of 80-89% due to several limitations in climate, soil, and landform, with Turrialba having a better climate, and Masatepe better soil and landform conditions (cf. Haggar et al., 2011). Finally, even though Nueva Guinea showed the best landform conditions, non-optimal values in mean temperature and soil chemical properties resulted in a land suitability score of 73%, only above of San Carlos.

This analysis shows that ALECA is not only able to assign high suitability scores to actual coffee areas in Central America as shown in the previous validation study but also to reproduce reported quality patterns between single coffee areas. Fig. 6 also shows that looking at the different component suitability scores instead of only the final land suitability score can help to understand the reasons behind the final ranking and facilitate better decision support.

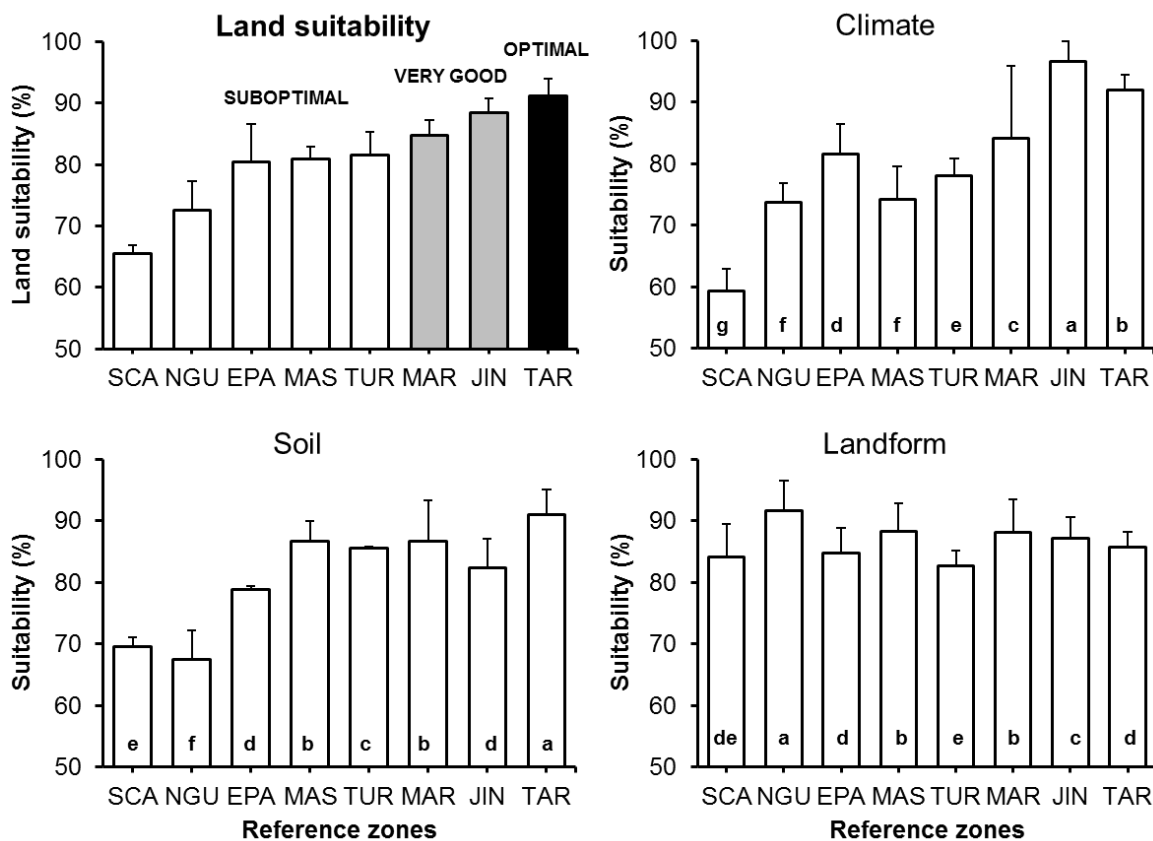


Fig. 6. Overall land suitability and single component suitability scores of the coffee reference zones in Central America. In the land suitability figure, the colors black (optimal), grey (very good) and white (suboptimal) indicate the commonly accepted classifications of the reference zones. Honduras: EPA=El Paraíso, MAR=Marcala; Nicaragua: NGU=Nueva Guinea, MAS=Masatepe, JIN=Jinotega; Costa Rica: SCA=San Carlos, TUR=Turrialba, TAR=Tarrazú. Errors bars represent the standard deviation of the mean. Different letters inside bars indicate a significant difference between reference zones (ANOVA by Fischer test; $p < 0.01$).

Based on the results of this validation study, we propose that land suitability scores above 90% can be categorized as optimal (cf. Tarrazú), 90-85% as very good (cf. Marcala and Jinotega), 84-75% as moderate (cf. El Paraíso, Masatepe, Turrialba), 74-60% as suboptimal (cf. San Carlos and Nueva Guinea) and values below 60% as unsuitable for coffee production.

5. Application example using uncertain information

One of the main motivations for developing ALECA was that the Bayesian network approach allows users to consider data uncertainty. Users may have precise information for some variables and incomplete information for others, or may simply want to consider the inherent uncertainty of some input data, like e.g. soil sampling deviations or errors in rainfall measurements. To demonstrate ALECA's ability to deal with this issue and still deliver reliable results, we conducted a simulation using input data with added uncertainty and compared the results to simulations without uncertainty in the input data. We assumed that farmers can easily measure slope, soil texture, and dry season length and thus have precise information (hard evidence) for these factors. For the remaining variables, we used Netica's Uncertain Value Format, which allows users to represent different types of data uncertainty by using Gaussian distributions, intervals, a set of (im)possibilities and others (Norsys, 2015). In this case, we used the Gaussian uncertain value format to calculate a mean and standard deviation for each case value based on the input data for each land case and the state values of each agroecological variable (Table 3). We then ran the model with both input datasets, thus estimating land suitability scores with and without uncertainty in the input data, and finally calculated for both result datasets in order to evaluate model performance the Bayesian metrics quadratic loss and spherical payoff, and the conventional metrics bias, RMSE, and Index of Agreement (Marcot, 2012; Willmott, 1981).

Results show that the mean land suitability-scores were approximately the same, with 76.13% and an SD=4.88 calculated from the input dataset with added uncertainty and 76.29% with an SD=7.7 from the original dataset (Fig. 8). The use of data with uncertainty generated a land suitability score distribution with shorter tails, however. This finding indicates that in areas with land suitability scores ranged 60-90%, ALECA's performance is quite accurate even under uncertainty, but that in regions with low (<60%) or high suitability scores ($\geq 90\%$) results differ by ca. 5% between data with and without uncertainty. This difference is negligible in our case, as the qualitative ranking of areas does not change with a difference of 5%: areas with suitability values below 60% are unsuitable for coffee production in any case, and areas with values above 90% will remain optimal or very good.

Overall, the study shows that in the suitability range most coffee areas are in Central America are found in, ALECA can deliver a reliable analysis even under uncertainty and can thus be used as a decision support tool even in situations where data is missing or uncertain.

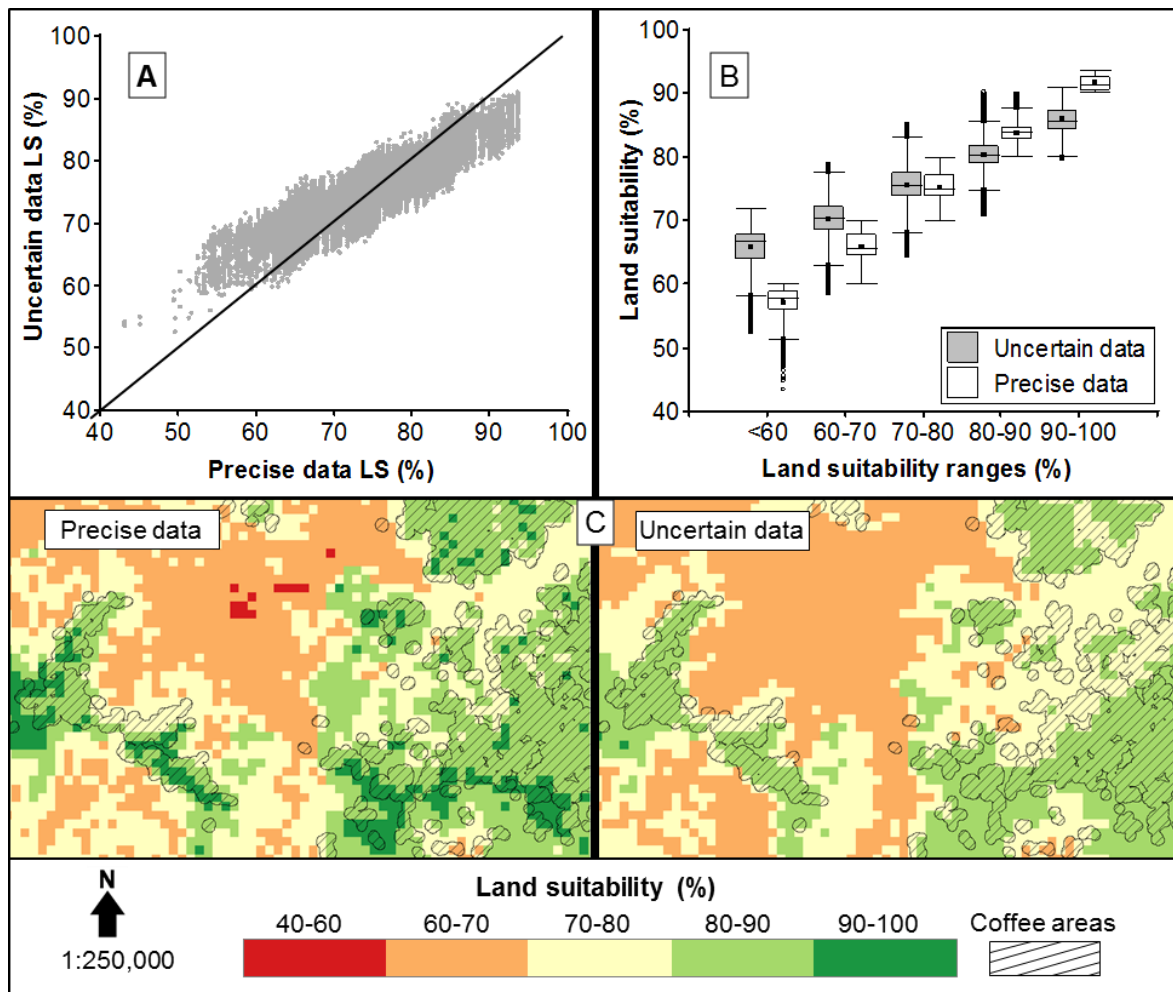


Fig. 7. Land suitability (LS) scores simulated from input data with and without added uncertainty. A) Dispersion graph of LS scores. A bias (ca. 5%) is visible between datasets at low and high LS values. B) Box plot of the same values as shown in A). The classes on the x-axis only apply to the white boxes (input data without uncertainty). The grey boxes show the LS scores of the same land units, but with added uncertainty in the input data. C) Two maps of a randomly chosen coffee zone in Nicaragua.

6. Discussion

Bayesian networks are versatile tools for the creation of land suitability evaluation systems for coffee production. Unlike other land suitability evaluation systems (Bunn et al., 2014), models like ALECA have a transparent and informative graphical interface that allows users to evaluate the suitability of a given land unit for coffee cultivation at variable, component and aggregate level. Decision makers can see, explore and learn from the evaluation process (Jakeman et al., 2006; Ranatunga et al., 2008). Like Mighty (2015) and Nzeyimana et al. (2014), we used climate, soil and topography variables to define land suitability. The main difference between their variable selection and ours is that they used proxy variables like soil orders and geological information to represent the effect of soil properties in some cases, while we focused on variables with a direct influence on land

suitability (Austin, 2002). Finally, the combined deterministic and probabilistic updating in ALECA allows the model to produce quick and accurate results by using i) deterministic updating when there is no uncertainty in the evidence for the parent variable(s) and no uncertainty in the equation or tables relating the child variable to its parent(s), and ii) probabilistic updating when uncertainty is present from any of those sources (B. Boerlage, Norsys Software Corp.; personal communication).

Some caveats should be kept in mind, however. First, we had to use data from different types of studies (surveys and trials), regions and conditions to develop the suitability functions, but assumed that the information appropriately represents the process under study (Rodríguez et al., 2011; van Oijen et al., 2010). Second, the literature largely refers to only *Coffea arabica* L. and ignores varietal differences (Bertrand et al., 2011; Philippe et al., 2009). Thus, we also excluded varietal differences from our model. Third, there is uncertainty associated with the datasets we used that we did not consider in the validation procedure for the sake of clarity. We opted to demonstrate how this factor can be considered in the uncertainty exercise instead. Finally, like other land evaluation systems, ALECA uses a linear combination method to estimate land suitability, which is a weighted sum of individual variables' suitability that excludes any interactions and assumes variable independence (Hopkins, 2014). However, the influence of some variables can depend on the state of other variables. Medium to low amounts of annual precipitation, for example, have less of an impact on coffee when soil texture is clayey (Willson, 1985) and the slope aspect modifies the temperature at the site level (Barry, 2008). Including these interactions would have required defining all conditional dependencies between the states of the different variables, which was not feasible due to a lack of information. The model validation exercise showed that ALECA performs well even without considering these interactions, but we will nevertheless keep refining the model with any new data set that becomes available.

Plans for ALECA also include the addition of variables describing the impact of farming practices like mulching on soil properties and erosion reduction, irrigation on water supply and the planting of trees on microclimate and soil conditions. The shade of trees, for example, reduces the temperature under the canopy in agroforestry coffee plantations (Siles et al., 2010). By defining the potential of temperature reduction based on the level of shade, it is easily possible to explore changes in temperature suitability due to trees in ALECA. ALECA can also be adapted for use in coffee regions outside Central America by updating the bounds and priors of the variables to the new conditions. For the regions that experience lower annual or seasonal temperatures than Central America, the introduction of a variable like 'temperature of the coldest month' may be necessary (Woodward and Williams, 1987), but no further changes to the structure would likely be needed.

7. Conclusions

In this study, we introduce ALECA, the first Bayesian network model to evaluate land suitability for coffee production in Central America under uncertainty. The validation showed that even without the use of coffee maps as input, ALECA reliably scored the suitability of actual coffee areas for coffee production as higher than that non-coffee areas, and was able to accurately predict the known order of quality of several coffee reference zones in Central America. We further showed that the model can also be used as a reliable decision support tool for coffee stakeholders in situations where some input data is uncertain. The graphical structure of the model permits users to easily assess the main factors determining land suitability for coffee production, to explore how changes in these factors impact suitability, and to plan adaptation measures accordingly.

Acknowledgments

This study was funded and supported by the School of Integrated Climate System Sciences (SICSS), the Kompetenzzentrum Nachhaltige Universität (KNU) and the cluster of excellence “Integrated Climate System Analysis and Prediction” (CliSAP) at Universität Hamburg. We would further like to thank Nitlapan-UCA and FUNICA in Nicaragua for providing the data of Nitlapan (2012) and Dr. Christophe Mongtanon, Dr. Enrique Succar and two anonymous reviewers for their comments on the manuscript.

References

- Adams, J.M., 2010. *Vegetation-climate interaction: how plants make the global environment*, 2nd ed, Springer-Praxis books in environmental sciences. Springer ; In Association with Praxis, Berlin ; New York : Chichester, UK.
- Aguilera, P.A., Fernández, A., Fernández, R., Rumí, R., Salmerón, A., 2011. Bayesian networks in environmental modelling. *Environmental Modelling & Software* 26, 1376–1388. doi:10.1016/j.envsoft.2011.06.004
- Alégre, C., 1959. *Climates et caféiers d’Arabie*. *Agronomie Tropicale* 14, 23–58.
- ANACAFE, 2006. *Guía Técnica de Caficultura*. ANACAFE, Guatemala, Guatemala.
- Austin, M.P., 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling* 157, 101–118. doi:10.1016/S0304-3800(02)00205-3
- Austin, M.P., 1980. Searching for a model for use in vegetation analysis. *Vegetatio* 42, 11–21. doi:10.1007/BF00048865
- Avelino, J., Barboza, B., Araya, J.C., Fonseca, C., Davrieux, F., Guyot, B., Cilas, C., 2005. Effects of slope exposure, altitude and yield on coffee quality in two altitude terroirs of Costa Rica, Orosi and Santa María de Dota. *Journal of the Science of Food and Agriculture* 85, 1869–1876. doi:10.1002/jsfa.2188

- Avelino, J., Cristancho, M., Georgiou, S., Imbach, P., Aguilar, L., Bornemann, G., Läderach, P., Anzueto, F., Hruska, A.J., Morales, C., 2015. The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. *Food Security* 7, 303–321. doi:10.1007/s12571-015-0446-9
- Barry, R.G., 2008. *Mountain Weather and Climate*. Cambridge University Press, Leiden.
- Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. *Environmental Modelling & Software* 40, 1–20. doi:10.1016/j.envsoft.2012.09.011
- Bertrand, B., Alpizar, E., Lara, L., SantaCreo, R., Hidalgo, M., Quijano, J., Montagnon, C., Georget, F., Etienne, H., 2011. Performance of *Coffea arabica* F1 hybrids in agroforestry and full-sun cropping systems in comparison with American pure line cultivars. *Euphytica* 181, 147–158. doi:10.1007/s10681-011-0372-7
- Blanco, R., Aguilar, A., 2015. Soil erosion and erosion thresholds in an agroforestry system of coffee (*Coffea arabica*) and mixed shade trees (*Inga* spp and *Musa* spp) in Northern Nicaragua. *Agriculture, Ecosystems & Environment* 210, 25–35. doi:10.1016/j.agee.2015.04.032
- Bonan, G.B., 2008. *Ecological climatology: concepts and applications*, 2nd ed. Cambridge University Press, Cambridge ; New York.
- Bornemisza, E., Segura, A., 1999. Los suelos cafetaleros de América Central y su fertilización, in: *Desafíos de la caficultura en Centroamérica*. IICA, PROMECAFE, San José, C. R, pp. 97–137.
- Bunn, C., Läderach, P., Pérez Jiménez, J.G., Montagnon, C., Schilling, T., 2015. Multiclass Classification of Agro-Ecological Zones for Arabica Coffee: An Improved Understanding of the Impacts of Climate Change. *PLOS ONE* 10, e0140490. doi:10.1371/journal.pone.0140490
- Bunn, C., Läderach, P., Rivera, O.O., Kirschke, D., 2014. A bitter cup: climate change profile of global production of Arabica and Robusta coffee. *Climatic Change* 129, 89–101. doi:10.1007/s10584-014-1306-x
- Camargo, A., 1985. O clima e a cafeicultura no Brasil. *Informe Agropecuario* 11, 13–26.
- Camargo, M.B.P. de, 2010. The impact of climatic variability and climate change on arabic coffee crop in Brazil. *Bragantia* 69, 239–247. doi:https://dx.doi.org/10.1590/S0006-87052010000100030
- Cannell, M.G.R., 1985. Physiology of the coffee crop, in: *Coffee*. Springer, pp. 108–134.
- Charlip, J.A., 2003. *Cultivating coffee: the farmers of Carazo, Nicaragua, 1880-1930*, Ohio University research in international studies. Latin America series. Ohio University Press, Athens.
- Chemura, A., Kutuywayo, D., Chidoko, P., Mahoya, C., 2015. Bioclimatic modelling of current and projected climatic suitability of coffee (*Coffea arabica*) production in Zimbabwe. *Reg Environ Change* 1–13. doi:10.1007/s10113-015-0762-9

- Chen, S.H., Pollino, C.A., 2012. Good practice in Bayesian network modelling. *Environmental Modelling & Software* 37, 134–145. doi:10.1016/j.envsoft.2012.03.012
- Chinchilla, M., Mata, R., Alvarado, A., 2011. Andisoles, inceptisoles y entisoles de la subcuenca del río Pirrís, región de Los Santos, Talamanca, Costa Rica. *Agronomía Costarricense* 35, 83–107.
- Coudun, C., Gégout, J.-C., Piedallu, C., Rameau, J.-C., 2006. Soil nutritional factors improve models of plant species distribution: an illustration with *Acer campestre* (L.) in France. *Journal of Biogeography* 33, 1750–1763. doi:10.1111/j.1365-2699.2005.01443.x
- DaMatta, F.M., Ramalho, J.D.C., 2006. Impacts of drought and temperature stress on coffee physiology and production: a review. *Brazilian Journal of Plant Physiology* 18, 55–81. doi:10.1590/S1677-04202006000100006
- Descroix, F., Snoeck, J., 2004. Environmental Factors Suitable for Coffee Cultivation, in: Wintgens, J. (Ed.), *Coffee: Growing, Processing, Sustainable Production*. Wiley-VCH, Alemania, pp. 164–177.
- Descroix, F., Wintgens, J.N., 2004. Establishing a Coffee Plantation, in: Wintgens, J.N. (Ed.), *Coffee: Growing, Processing, Sustainable Production*. Wiley-VCH Verlag GmbH, pp. 178–245.
- D'haeze, D., Deckers, J., Raes, D., Phong, T.A., Loi, H.V., 2005. Environmental and socio-economic impacts of institutional reforms on the agricultural sector of Vietnam: Land suitability assessment for Robusta coffee in the Dak Gan region. *Agriculture, Ecosystems & Environment* 105, 59–76. doi:10.1016/j.agee.2004.05.009
- Eakin, H., Tucker, C.M., Castellanos, E., Diaz-Porrás, R., Barrera, J.F., Morales, H., 2014. Adaptation in a multi-stressor environment: perceptions and responses to climatic and economic risks by coffee growers in Mesoamerica. *Environ Dev Sustain* 16, 123–139. doi:10.1007/s10668-013-9466-9
- ESRI, 2008. How Aspect works [WWW Document]. URL <http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=How%20Aspect%20works> (accessed 9.23.15).
- FAO, 1976. A framework for land evaluation, *Soils bulletin*. Food and Agriculture Organization of the United Nations, Rome, Italy.
- Forestier, J., 1969. *Culture du caféier Robusta en Afrique centrale*: par J. Forestier, Institut français du café et du cacao, Paris.
- Franco, C.M., 1958. Influence of temperatures on growth of coffee plant, *Bulletin*. IBEC Research Institute, New York (EUA), New York, NY (EUA).
- Franklin, J., Miller, J.A., 2009. *Mapping species distributions spatial inference and prediction*. Cambridge University Press, Cambridge; New York.
- Giorgi, F., 2006. Climate change hot-spots. *Geophys. Res. Lett.* 33, L08707. doi:10.1029/2006GL025734
- Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C., 2010. Food Security: The Challenge of Feeding 9 Billion People. *Science* 327, 812–818. doi:10.1126/science.1185383
- Haarer, A.E., 1958. *Modern Coffee Production*. Leonard Hill Limited, London, England.

- Haggar, J., Barrios, M., Bolaños, M., Merlo, M., Moraga, P., Munguia, R., Ponce, A., Romero, S., Soto, G., Staver, C., de M. F. Virginio, E., 2011. Coffee agroecosystem performance under full sun, shade, conventional and organic management regimes in Central America. *Agroforestry Systems* 82, 285–301. doi:10.1007/s10457-011-9392-5
- Hengl, T., de Jesus, J.M., MacMillan, R.A., Batjes, N.H., Heuvelink, G.B.M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Gonzalez, M.R., 2014. SoilGrids1km — Global Soil Information Based on Automated Mapping. *PLoS ONE* 9, e105992. doi:10.1371/journal.pone.0105992
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25, 1965–1978. doi:10.1002/joc.1276
- Hopkins, L.D., 2014. Methods for Generating Land suitability Maps: A Comparative Evaluation, in: Ndubisi, F.O. (Ed.), *The Ecological Design and Planning Reader*. Island Press/Center for Resource Economics, pp. 348–367.
- ICO, 2015. Historical Data on the Global Coffee Trade [WWW Document]. URL http://www.ico.org/new_historical.asp?section=Statistics (accessed 9.11.15).
- IICA, 2003. *Cadena de Comercialización del Café*. Managua, Nicaragua.
- Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software* 21, 602–614. doi:10.1016/j.envsoft.2006.01.004
- Jaramillo, A., Guzmán, M., 1984. Relación entre la temperatura y el crecimiento en *Coffea arabica* L. variedad Caturra. *CENICAFE* 35, 57–65.
- Jensen, F.V., Nielsen, T., 2007a. Building Models, in: *Bayesian Networks and Decision Graphs, Information Science and Statistics*. Springer New York, pp. 51–108.
- Jensen, F.V., Nielsen, T.D., 2007b. *Bayesian networks and decision graphs.*, 2nd ed, Information science and statistics. Springer, New York.
- Karmalkar, A.V., Bradley, R.S., Diaz, H.F., 2011. Climate change in Central America and Mexico: regional climate model validation and climate change projections. *Climate Dynamics* 37, 605–629. doi:10.1007/s00382-011-1099-9
- Landuyt, D., Broekx, S., D'hondt, R., Engelen, G., Aertsens, J., Goethals, P.L.M., 2013. A review of Bayesian belief networks in ecosystem service modelling. *Environmental Modelling & Software* 46, 1–11. doi:10.1016/j.envsoft.2013.03.011
- Lara Estrada, L., 2005. Effects of the altitude, shade, production and fertilization on coffee quality (*Coffea arabica* L. var. Caturra) produced in agroforestry systems of coffee zone Northern Central of Nicaragua (M.Sc.). Spanish. Postgraduate School - CATIE, Turrialba, Costa Rica. <http://hdl.handle.net/11554/4936>.
- Larcher, W., 1981. Effects of low temperature stress and frost injury on plant productivity, in: *Physiological Processes Limiting Plant Productivity*. Butterworth-Heinemann, pp. 253–269.
- Littleboy, M., Smith, D.M., Bryant, M.J., 1996. Simulation modelling to determine suitability of agricultural land. *Ecological Modelling, Environmental and Ecological Models for Simulation and Management* 86, 219–225. doi:10.1016/0304-3800(95)00055-0

- Lopez-Rodriguez, G., Sotomayor-Ramírez, D., Amador, J., Schroder, E., 2015. Contribution of nitrogen from litter and soil mineralization to shade and sun coffee (*Coffea arabica* L.) agroecosystems. *Tropical Ecology* 56, 155–156. <http://tropecol.com/volumes/toc/en/toc56-2.htm>.
- Maestri, M., Barros, R.S., 1977. Coffee, in: Alvin, P., Kozlowski, T. (Eds.), *Ecophysiology of Tropical Crops*. Academic Press, Inc, New York, USA.
- Magaña, V., Amador, J.A., Medina, S., 1999. The midsummer drought over Mexico and Central America. *Journal of Climate* 12, 1577–1588. doi:[http://dx.doi.org/10.1175/1520-0442\(1999\)012<1577:TMDOMA>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(1999)012<1577:TMDOMA>2.0.CO;2)
- Marcot, B.G., 2012. Metrics for evaluating performance and uncertainty of Bayesian network models. *Ecological Modelling* 230, 50–62. doi:10.1016/j.ecolmodel.2012.01.013
- Marcot, B.G., Steventon, J.D., Sutherland, G.D., McCann, R.K., 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. *Can. J. For. Res.* 36, 3063–3074. doi:10.1139/x06-135
- McRae, S.G. (Stuart G., Burnham, C.P. (Christopher P., 1981. *Land evaluation, Monographs on soil survey*. Clarendon Press ; Oxford University Press, Oxford : New York.
- Mighty, M.A., 2015. Site suitability and the analytic hierarchy process: How GIS analysis can improve the competitive advantage of the Jamaican coffee industry. *Applied Geography* 58, 84–93. doi:10.1016/j.apgeog.2015.01.010
- Molina, E., Melendez, G., 2002. *Tabla de interpretación de análisis de suelos*. Centro de Investigaciones Agronomicas, Universidad de Costa Rica, Costa Rica.
- Muschler, R.G., 2001. Shade improves coffee quality in a sub-optimal coffee-zone of Costa Rica. *Agroforestry Systems* 131–139. doi:10.1023/A:1010603320653
- Nitlapan, 2012. Proyecto “Acceso a mercados de café diferenciados”. Resultados de monitoreo y auditoría ambiental. (Project report). NITLAPAN-UCA, Managua, Nicaragua.
- Norsys, 2015. *Netica Help [WWW Document]*. URL <http://www.norsys.com/WebHelp/NETICA.htm> (accessed 9.15.15).
- Nunes, M., Brumby, D., D Davies, 1973. Estudo comparativo do metabolismo fotossintético em folhas de cafeeiro, beterraba e cana-de-açúcar. *Garcia de Orta, Série Estudos Agronómicos* 1, 1–14.
- Nzeyimana, I., Hartemink, A.E., Geissen, V., 2014. GIS-Based Multi-Criteria Analysis for Arabica Coffee Expansion in Rwanda. *PLoS ONE* 9, e107449. doi:10.1371/journal.pone.0107449
- Okoth, P.F., Ng’ang’a, J.K., Kimani, P.K., 2007. Consequences of Field Management and Soil Erosion on the Sustainability of Large Scale Coffee Farming in Kiambu, in: Bationo, A., Waswa, B., Kihara, J., Kimetu, J. (Eds.), *Advances in Integrated Soil Fertility Management in Sub-Saharan Africa: Challenges and Opportunities*. Springer Netherlands, pp. 299–309.
- Osman, K.T., 2013. *Soils: Principles, Properties and Management*. Springer Netherlands, Dordrecht.

- Ovalle-Rivera, O., Läderach, P., Bunn, C., Obersteiner, M., Schroth, G., 2015. Projected Shifts in *Coffea arabica* Suitability among Major Global Producing Regions Due to Climate Change. *PLoS ONE* 10, e0124155. doi:10.1371/journal.pone.0124155
- Pearl, J., 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, 1st ed. Morgan Kaufmann, San Francisco, USA.
- Phillippe, L., Benoît, B., Hervé, E., 2009. Breeding Coffee (*Coffea arabica*) for Sustainable Production, in: Jain, S.M., Priyadarshan, P.M. (Eds.), *Breeding Plantation Tree Crops: Tropical Species*. Springer New York, pp. 525–543.
- Philpott, S.M., Lin, B.B., Jha, S., Brines, S.J., 2008. A multi-scale assessment of hurricane impacts on agricultural landscapes based on land use and topographic features. *Agriculture, Ecosystems & Environment* 128, 12–20. doi:10.1016/j.agee.2008.04.016
- Pollino, C.A., Woodberry, O., Nicholson, A., Korb, K., Hart, B.T., 2007. Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. *Environmental Modelling & Software, Bayesian networks in water resource modelling and management* 22, 1140–1152. doi:10.1016/j.envsoft.2006.03.006
- Poppenborg, P., Koellner, T., 2014. A Bayesian network approach to model farmers' crop choice using socio-psychological measurements of expected benefits of ecosystem services. *Environmental Modelling & Software* 57, 227–234. doi:10.1016/j.envsoft.2014.03.006
- Ramírez, J.E., 2009. *Hacia la caficultura sostenible*. San José (Costa Rica).
- Ranatunga, K., Nation, E.R., Barratt, D.G., 2008. Review of soil water models and their applications in Australia. *Environmental Modelling & Software* 23, 1182–1206. doi:10.1016/j.envsoft.2008.02.003
- Refsgaard, J.C., van der Sluijs, J.P., Højberg, A.L., Vanrolleghem, P.A., 2007. Uncertainty in the environmental modelling process – A framework and guidance. *Environmental Modelling & Software* 22, 1543–1556. doi:10.1016/j.envsoft.2007.02.004
- Rodríguez, D., Cure, J.R., Cotes, J.M., Gutierrez, A.P., Cantor, F., 2011. A coffee agroecosystem model: I. Growth and development of the coffee plant. *Ecological Modelling* 222, 3626–3639. doi:10.1016/j.ecolmodel.2011.08.003
- Rojas, O., 1989. Determinación del potencial agroecológico para el cultivo del café (*Coffea arabica*) en Costa Rica. *Turrialba* 39, 279–287.
- Rueda, X., Lambin, E.F., 2013. Linking Globalization to Local Land Uses: How Eco-Consumers and Gourmands are Changing the Colombian Coffee Landscapes. *World Development* 41, 286–301. doi:10.1016/j.worlddev.2012.05.018
- Samper, M., 1999. Trayectoria y viabilidad de las Caficultoras Centroamericanas, in: *Desafíos de La Caficultura En Centroamérica*. IICA/PROMECAFE/CIRAD/IRD/CCCR, San José, Costa Rica, pp. 1–68.
- Siles, P., Harmand, J.-M., Vaast, P., 2010. Effects of *Inga densiflora* on the microclimate of coffee (*Coffea arabica* L.) and overall biomass under optimal growing conditions in Costa Rica. *Agroforestry Systems* 78, 269–286. doi:10.1007/s10457-009-9241-y

- Silva, S.A., Lima, J.S.S., Bottega, E.L., 2013. Yield mapping of arabic coffee and their relationship with plant nutritional status. *Journal of soil science and plant nutrition* 13, 556–564. <http://ref.scielo.org/s5dspw>.
- Somarriba, E., Harvey, C.A., Samper, M., Anthony, F., González, J., Staver, C., Rice, R.A., 2004. Biodiversity Conservation in Neotropical Coffee (*Coffea arabica* L.) Plantations, in: Schroth, G., da Fonseca, G., Harvey, C., Gascon, C., Vasconcelos, H., Izac, A. (Eds.), *Agroforestry and Biodiversity Conservation in Tropical Landscapes*. Washington, DC, USA.
- Sys, C., Ranst, E. van, Debaveye, J., 1991. *Land Evaluation. Part I. Principles in Land Evaluation and Crop Production Calculations*, Agricultural Publications. Brussels, Bergium.
- Taylor, M.A., Alfaro, E.J., 2005. Central America and the Caribbean, Climate of, in: *Encyclopedia of World Climatology*. Springer, pp. 183–189.
- Teuber, R., 2009. Café de Marcala - Honduras' GI Approach to Achieving Reputation in the Coffee Market. *Estey Centre Journal of International Law and Trade Policy* 10, 131–148. <http://purl.umh.edu/48798>.
- Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R., Polasky, S., 2002. Agricultural sustainability and intensive production practices. *Nature* 418, 671–677. doi:10.1038/nature01014
- Tucker, C.M., Eakin, H., Castellanos, E.J., 2010. Perceptions of risk and adaptation: Coffee producers, market shocks, and extreme weather in Central America and Mexico. *Global Environmental Change* 20, 23–32. doi:10.1016/j.gloenvcha.2009.07.006
- Ukers, W., 1935. *All about coffee*, Second. ed. New York, USA.
- USDA, 2015. *Coffee Annual Report: Guatemala* (GAIN Report No. 15003). Foreign Agricultural Service, USDA. <http://www.fas.usda.gov/data/guatemala-coffee-annual-0>.
- Vaast, P., Bertrand, B., Perriot, J.-J., Guyot, B., Génard, M., 2006. Fruit thinning and shade improve bean characteristics and beverage quality of coffee (*Coffea arabica* L.) under optimal conditions. *Journal of the Science of Food and Agriculture* 86, 197–204. doi:10.1002/jsfa.2338
- Vaast, P., Cilas, C., Perriot, J.J., Davrieux, J.J., Guyot, B., Bolaños, M., 2005. Mapping of Coffee Quality in Nicaragua According to Regions, Ecological Conditions and Farm Management, in: *ASIC Conference*. Bangalore, India.
- Vaast, P., Kanten, R., Siles, P., Angrand, J., Aguilar, A., 2008. Biophysical Interactions Between Timber Trees and Arabica Coffee in Suboptimal Conditions of Central America, in: Jose, S., Gordon, A.M., Nair, P.K.R. (Eds.), *Toward Agroforestry Design, Advances in Agroforestry*. Springer Netherlands, pp. 133–146.
- van Oijen, M., Dauzat, J., Harmand, J.-M., Lawson, G., Vaast, P., 2010. Coffee agroforestry systems in Central America: II. Development of a simple process-based model and preliminary results. *Agroforestry Systems* 80, 361–378. doi:10.1007/s10457-010-9291-1
- Vega, F.E., Rosenquist, E., Collins, W., 2003. Global project needed to tackle coffee crisis. *Nature* 425, 343–343. doi:10.1038/425343a

- Verheye, W., 2002. Management of Agricultural Land: Chemical and Fertility Aspects, in: Encyclopedia of Life Support Systems (EOLSS).
- Verheye, W.H., 1987. Land Suitability Evaluation in Major Agro-ecological Zones and its Application in Land Use Planning and Nature Protection, in: Barth, H., L'Hermite, P. (Eds.), Scientific Basis for Soil Protection in the European Community. Springer Netherlands, pp. 377–388.
- Wallis, J.A.N., 1963. Water use by irrigated Arabica coffee in Kenya. *The Journal of Agricultural Science* 60, 381–388. doi:<http://dx.doi.org/10.1017/S0021859600011977>
- Wang, N., Jassogne, L., van Asten, P.J.A., Mukasa, D., Wanyama, I., Kagezi, G., Giller, K.E., 2015. Evaluating coffee yield gaps and important biotic, abiotic, and management factors limiting coffee production in Uganda. *European Journal of Agronomy* 63, 1–11. doi:10.1016/j.eja.2014.11.003
- Willmott, C.J., 1981. On the validation of models. *Physical geography* 2, 184–194. doi:10.1080/02723646.1981.10642213
- Willson, K.C., 1985. Climate and Soil, in: *Coffee. Botany, Biochemistry and Production of Beans and Beverage*. New York, USA.
- Woodward, F.I., Williams, B.G., 1987. Climate and plant distribution at global and local scales. *Vegetatio* 69, 189–197. doi:10.1007/BF00038700

Appendix

Influence of single components on final land suitability scoring

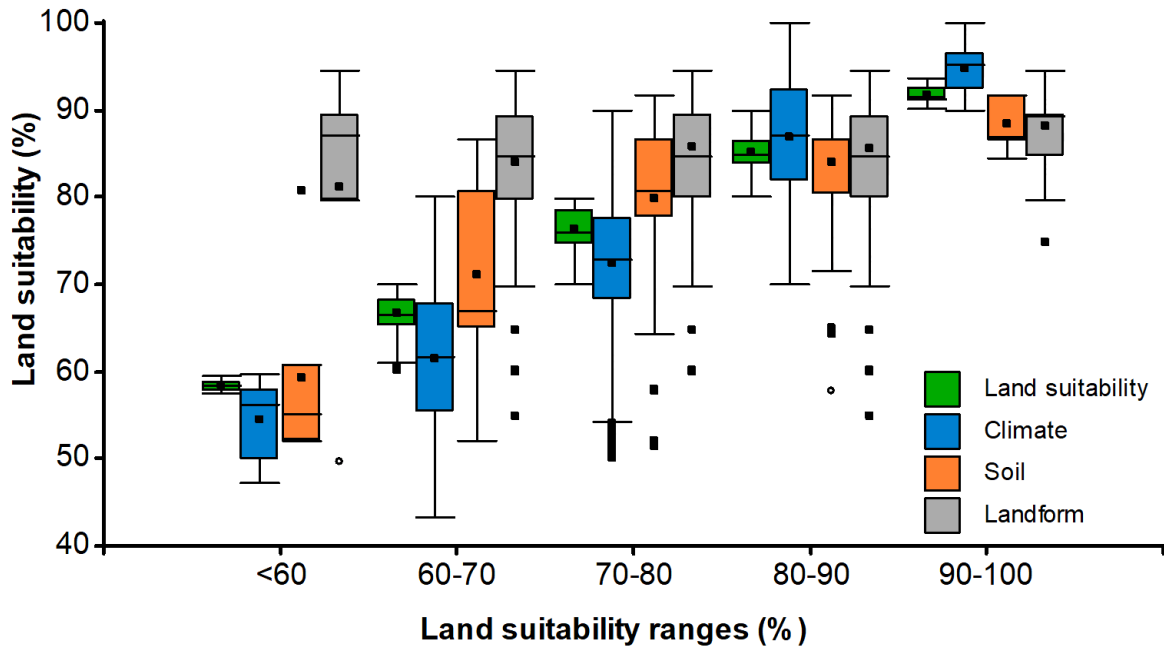


Fig. A1. Suitability scores of single components and final land suitability score of coffee areas in Central America. The climate and soil components have a larger influence on the final land suitability score than landform. To achieve very high or low land suitability scores, both climate and soil components need to have high or low values, indicating that no further analysis of causes is required. Intermediate land suitability values (60 to 80%) can arise from diverse component values, in which case the single component and variable values need to analyze in order to identify the limiting factors for coffee production.

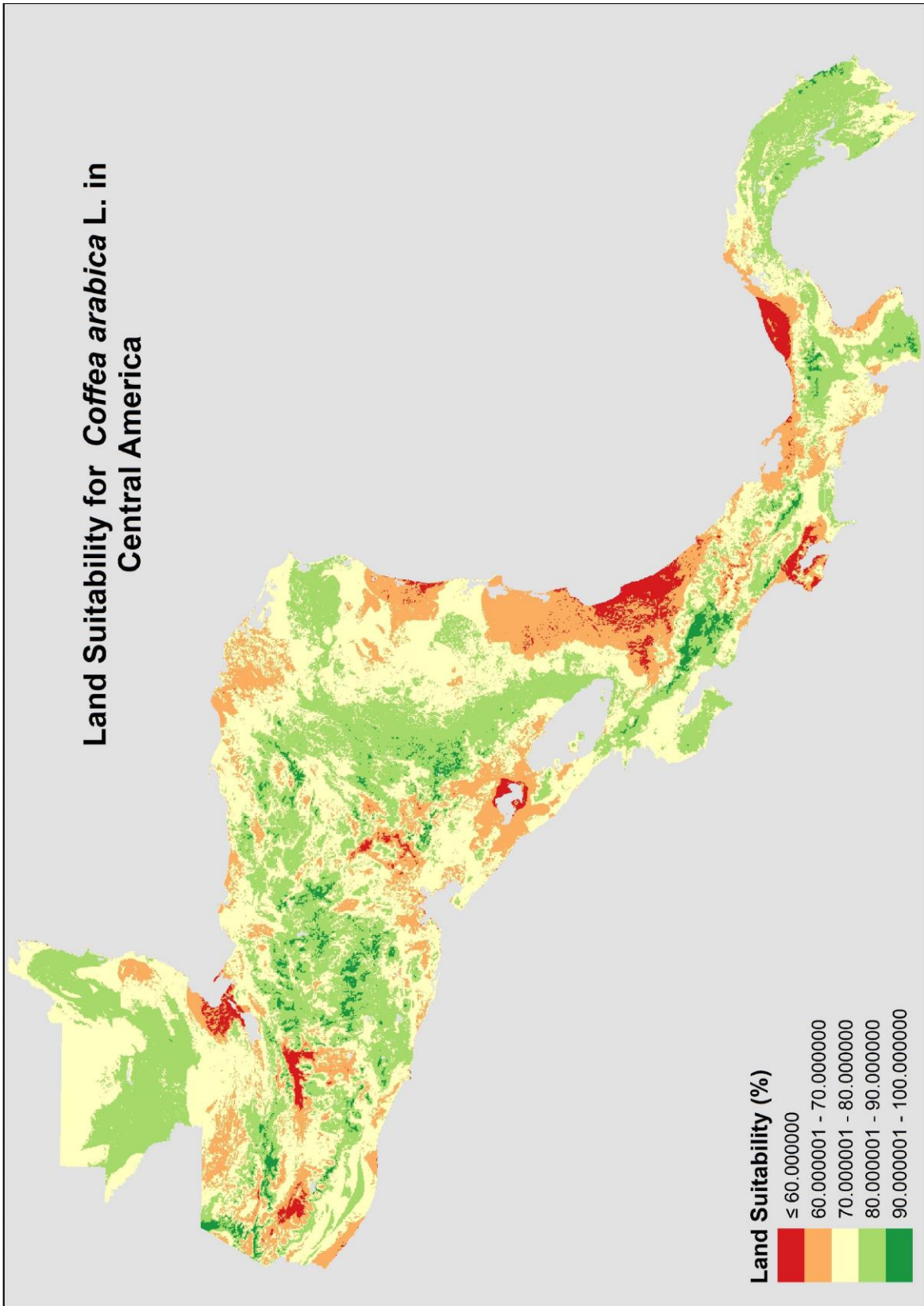


Fig. A2. Map of current land suitability for *Coffea arabica* L. in Central America

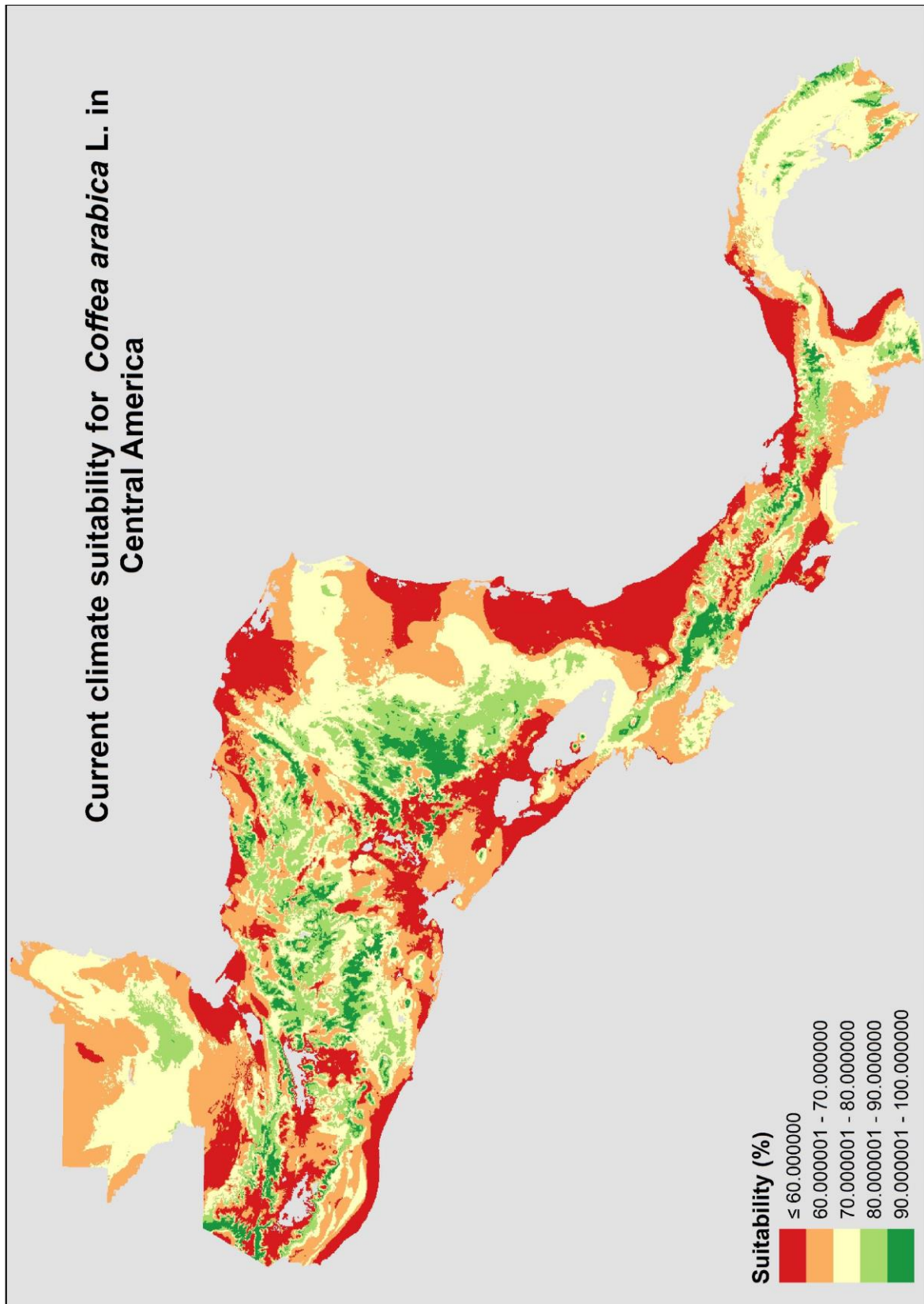


Fig. A3. Map of current climate suitability for *Coffea arabica* L. in Central America

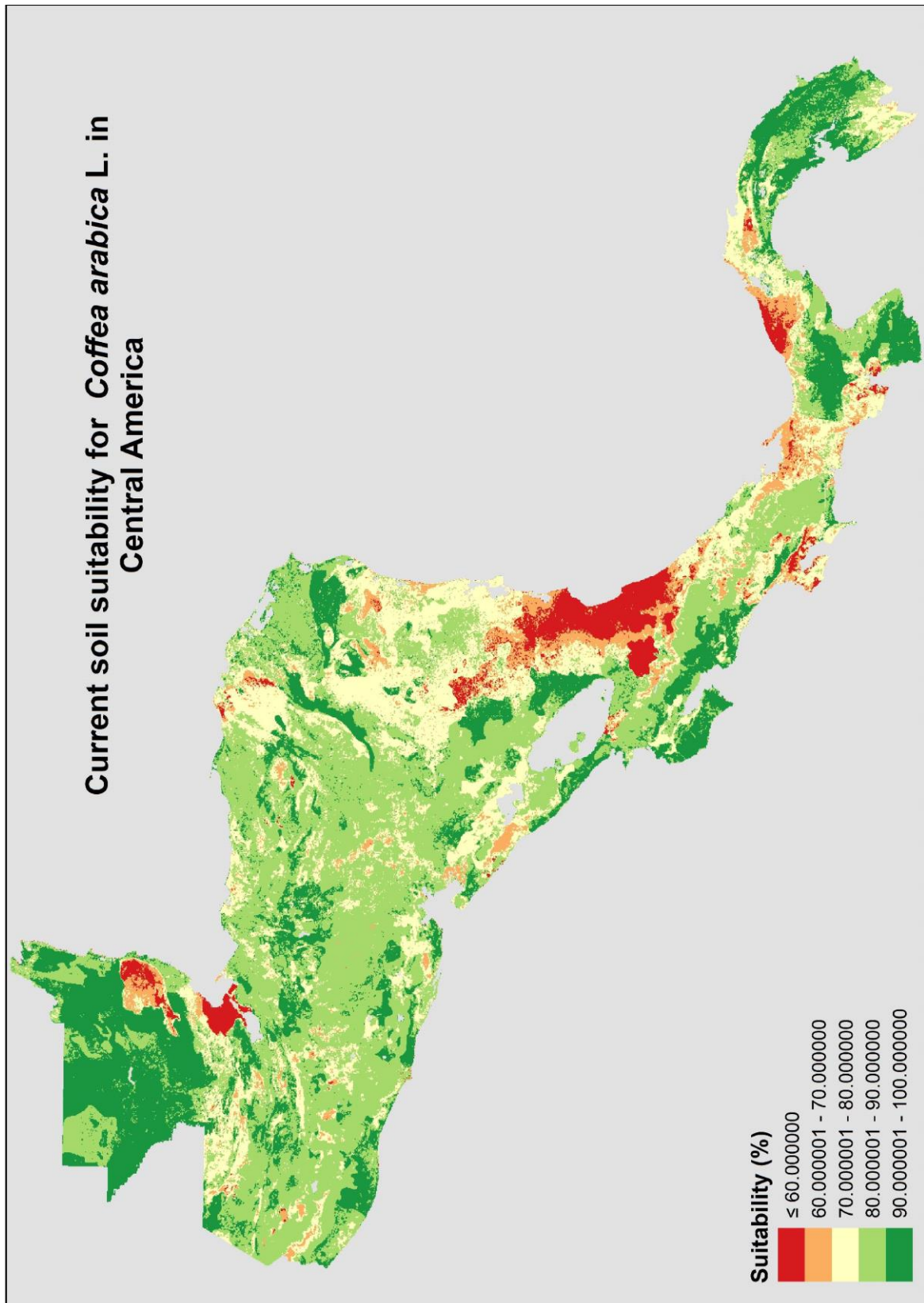


Fig. A4. Map of current soil suitability for *Coffea arabica* L. in Central America

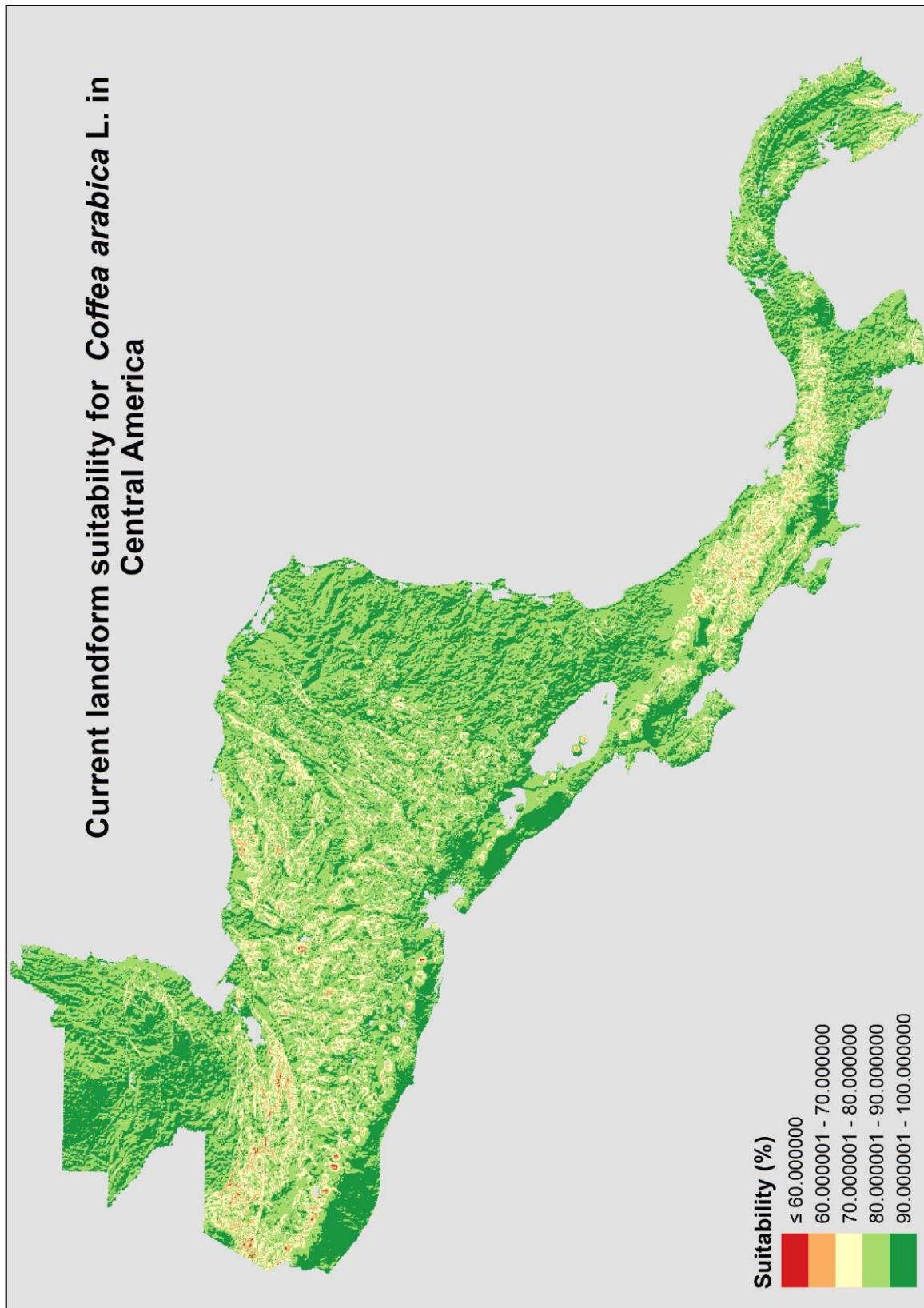


Fig. A5. Map of current landform suitability for *Coffea arabica* L. in Central America

Supplementary Material

Links to download four high resolution images in format PDF:

- i. [Map of current land suitability for Coffea arabica L. in Central America](#)
- ii. [Map of current climate suitability for Coffea arabica L. in Central America](#)
- iii. [Map of current soil suitability for Coffea arabica L. in Central America](#)
- iv. [Map of current landform suitability for Coffea arabica L. in Central America](#)

Accompanying figures to Table 5

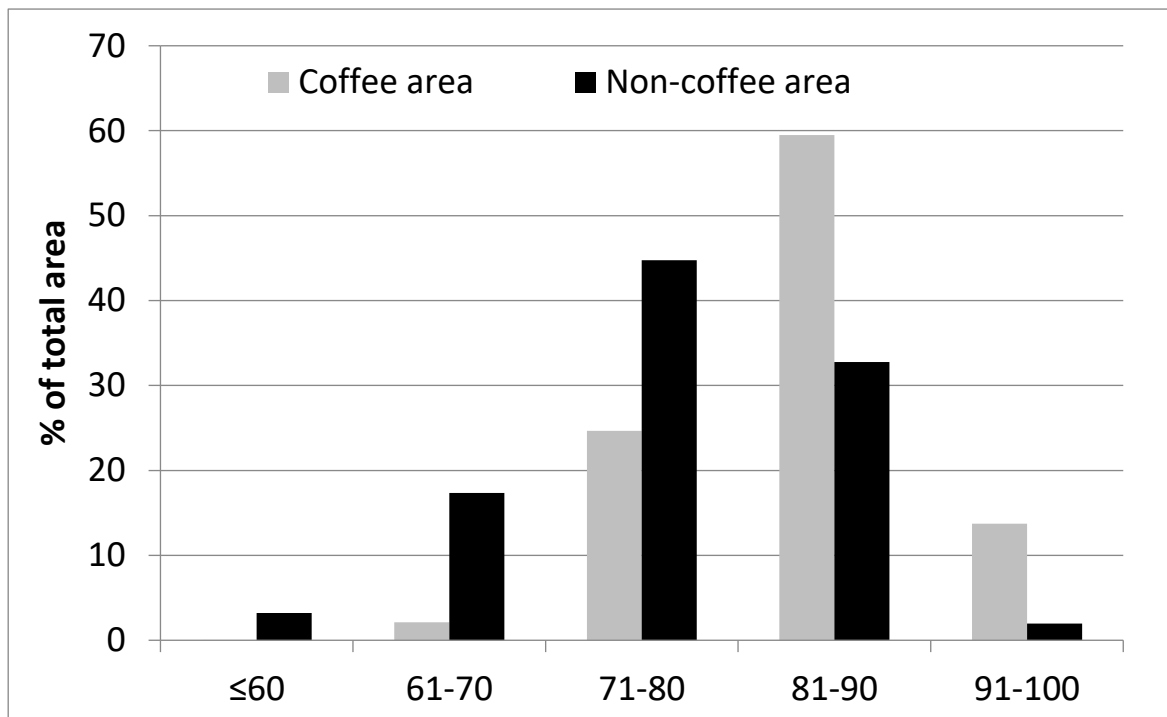


Fig. S1. Simulated share of total coffee and non-coffee area in Central America in the different land suitability classes for coffee production.

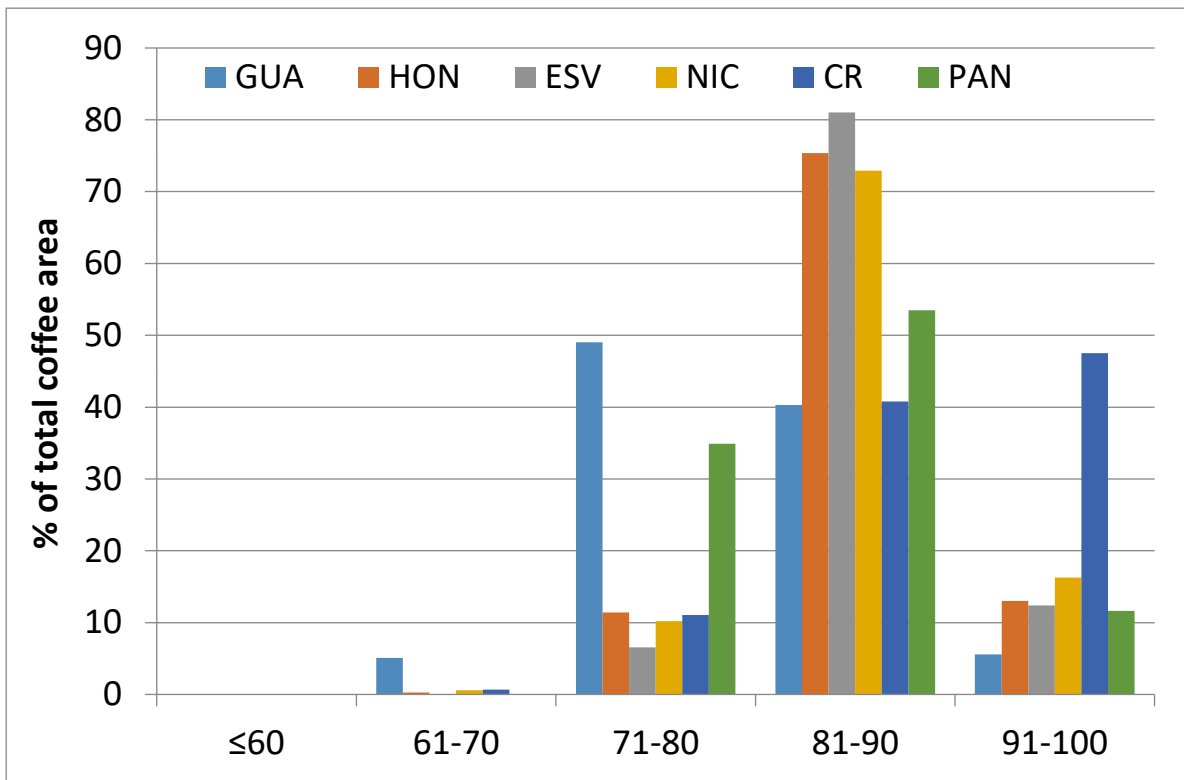


Fig. S2. Simulated share of total coffee areas in the countries Guatemala, Honduras, El Salvador, Nicaragua, Costa Rica, and Panama in the different land suitability classes for coffee production.