



Global mapping of *Jatropha curcas* yield based on response of fitness to present and future climate

ANTONIO TRABUCCO*[†], WOUTER M. J. ACHTEN*, COLM BOWE[‡], RAFAERTS*, JOS VAN ORSHOVEN*, LINDSEY NORGROVES[§] and BART MUYLS*

*Division Forest, Nature and Landscape, Katholieke Universiteit Leuven, BE-3001 Leuven, Belgium, [†]International Water Management Institute (IWMI), Colombo, Sri Lanka, [‡]Centre for Underutilised Crops, Environment Division, School of Civil Engineering and the Environment, University of Southampton, Southampton SO17 1BJ, UK, [§]CABI Europe Switzerland, CH-2800 Delémont, Switzerland

Abstract

Although acclaimed as a biofuel crop with high potential to sustainably replace fossil fuels, *Jatropha curcas* L. remains a poorly studied plant. Reliable yield assessments with conventional methods require agroclimatic and physiological knowledge, which is not yet available for *Jatropha*. To fill this gap, we tested a novel two-step approach integrating knowledge from biogeography and population biology with available *Jatropha* field data. In the first step, using MaxEnt, a widely implemented model in biogeography, we predicted *Jatropha* fitness in response to climate by relating natural occurrence recorded in herbaria with bioclimatic geodatasets. In the second step, we relied on population biology principles supported by seed mass addition experiments to relate fitness to reproductive potential, hence seed yield. *Jatropha* seed yield in response to climate was mapped worldwide for actual (1950–2000 average) and future (2020) climate conditions. The modelled *Jatropha* seed yield was validated against a set of on-field yield assessments ($R^2 = 0.67$, $P < 0.001$). The discrepancies between estimated and measured yields were partially explained by model uncertainties, as quantified by the sensitivity analysis of our modelling ($R^2 = 0.57$, $P = 0.001$). *Jatropha* has a pan-tropical distribution, plus specific adaptability to hot temperate areas. Climate variables most significantly affecting modelled yield response were annual average temperature, minimum temperature, annual precipitation and precipitation seasonality.

Keywords: climate change, environmental response, GIS, land suitability, MaxEnt, precipitation, productivity, temperature

Received 30 July 2009; revised version received 27 April 2010 and accepted 1 May 2010

Introduction

The financial sustainability of *Jatropha* projects requires reliable productivity estimates. Crop productivity depends primarily on the effect of environmental conditions, which can be mitigated or amplified by agronomic practices. However, *Jatropha* knowledge is currently not able to parameterize reliably traditional crop productivity models. We highlight the existence of data locating natural occurrence of *Jatropha*, which can be used, with biogeographic modeling and ecological

principles, to generate an alternative method to model *Jatropha* productivity.

Given food security and land scarcity concerns, *Jatropha curcas* L. (further referred as *Jatropha*) raised attention as a sustainable biofuel source for marginal and degraded semiarid areas (Fairless, 2007; Ndong *et al.*, 2009). *Jatropha*'s alleged ability to produce biofuel in areas unsuitable for intensive food crops production is expected to support food security, unlike most biofuel crops of first generation (Preston, 2009). Global demand for liquid biofuels together with optimistic claims about *Jatropha* has triggered a massive promotion and implementation of *Jatropha* plantations by private and public sectors (Planning Commission of India, 2003; GEXSI, 2008). Yet, given the scarcity of data on *Jatropha* genetics, its basic agronomy and biophysical requirements

Correspondence: B. Muys and A. Trabucco, tel. +32 16329726, fax +32 16329760, e-mail: bart.muys@ees.kuleuven.be, antonio.trabucco@ees.kuleuven.be

(Achten *et al.*, 2010), such promotion is economically risky (Openshaw, 2000). Financially sustainable biofuel policies should set regulatory frameworks based on a reliable understanding of those factors that influence biofuel production.

Crop performance is primarily constrained by the biophysical environment, which influences the availability of basic resources light, heat, water and mineral nutrients for plant growth (Mackey & Lindenmayer, 2001). Intensified agricultural systems take advantage of agronomic practices (e.g. plant breeding, irrigation, pesticides and fertilizers) to mitigate specific environmental constraints (Boyer, 1982). A clear understanding of the relationship between environment and crop performances can develop solid ground upon which to optimize agronomic practices. Unfortunately, scientific literature reports only limited knowledge describing how environmental factors influence *Jatropha* physiology. *Jatropha* grows on a wide range of climatic conditions, from semiarid to humid (annual rainfall varying from 300 to 3000 mm), with high tolerance to high temperatures, little frost tolerance and a preference for deep, well-drained soils (Achten *et al.*, 2008). Therefore, there is an urgent need for more specific knowledge, highlighting *Jatropha* productivity response to environmental factors. This knowledge can further be translated into spatial assessments of *Jatropha* yield across different agroclimatic zones.

Crop productivity in response to the environment is traditionally defined from either (1) land suitability models that empirically weigh combinations of crop responses to relevant agroclimatic factors (FAO, 1983, 1985; Sys *et al.*, 1991; Rossiter, 1996) or (2) physically based models such as STICS (Brisson *et al.*, 2003), EPIC (Tan & Shibasaki, 2003) and LPJML (Bondeau *et al.*, 2007), that simulate phenological processes, physiology, biogeochemical cycles, etc. Current *Jatropha* knowledge is too limited to reliably parameterize these methods. Furthermore, physically based models are based on specific plant traits not matching *Jatropha*, which is a stem succulent plant potentially shifting from C3 to CAM metabolism (Maes *et al.*, 2009a). These circumstances encourage the development of novel approaches, which can provide alternative estimates of *Jatropha* yield from available data.

Herbaria specimen can provide useful information on geographic locations of natural occurrences of plants (Elith *et al.*, 2006). In fact, a substantial collection of *Jatropha* specimen and their locations is available from online herbaria (WBIN, 2002). Availability of environmental geodatasets and advanced computational methods in statistics have catalyzed a rapid development of species distribution modeling (SDM) trained on herbaria specimen locations (Elith *et al.*, 2006). Thus, metrics of

species occurrence can be modeled and mapped as a function of environmental factors (Pearson & Dawson, 2003). MaxEnt (Phillips *et al.*, 2004) is a widely and successfully used SDM, which relates specimen locations and environmental geodatasets to derive the response of species occurrence probability as environmental conditions change (the 'environmental response'). These environmental functions are then applied to the same geodatasets to reconstruct the geographic distribution of species occurrence probability. This geographic distribution defines where environmental conditions are more suitable for species occurrence, including areas that have not been naturally 'colonized', due to biotic competition or geographic barriers (Hutchinson, 1957; Phillips *et al.*, 2004).

Compared with other SDM, MaxEnt has demonstrated excellent performance, even with a limited number of specimen locations (Elith *et al.*, 2006; Hernandez *et al.*, 2006). For example, species distribution classifications trained on as few as 15–20 specimen locations showed robust results for Gecko species in Madagascar (Pearson *et al.*, 2007) and for 76 plant species in Ecuador and Bolivia (Loiselle *et al.*, 2008). Loiselle *et al.* (2008) found that the effects of biases in specimen sampling were limited if a minimum of 100 specimen locations were available. A geographical sampling bias, commonly correlated to presence of towns and transportation routes (Reddy & Davalos, 2003), does not influence MaxEnt accuracy as long as infrastructures are distributed evenly along environmental gradients. MaxEnt provides flexible features, which reduce moderate levels of climate bias and sampling geo-local errors (Graham *et al.*, 2008; Phillips & Dudik, 2008). However, uncertainties in suitability prediction should be verified while looking at its transferability effectiveness, it is the ability to successfully transfer a model outside the training environmental conditions (Phillips, 2008).

The fitness of a species is measured by its reproductive success (Hartl, 2000) and is often quantified with measures of relative and absolute abundance (Zielinski & Kucera, 1995). Probability of occurrence quantifies habitat suitability and, as a fitness measure, has positive relation with species density measures (He & Gaston, 2003; VanderWal *et al.*, 2009) and has been used to assess species absolute density distribution (Royle *et al.*, 2005; Gaston *et al.*, 2006). In evolutionary biology terms, natural selection directly rewards and selects those reproductive efforts that are most successful in recruiting new individuals, highlighting an optimal trade-off between resources allocated to seed mass and offspring recruitment probability (Vermeij, 2004). Indeed, reproductive performance of a species is simply an outcome of parental fecundity and offspring survival (Silvertown

& Charlesworth, 2001). Fecundity is the number of seeds produced by a flowering plant (Reekie & Bazzaz, 2005), while offspring survival is critically influenced by positive and vigorous seedling establishment (Rey & Alcantara, 2000). Seedlings from larger seeds of the same species, assuming low predation of a toxic seed such as *Jatropha*, have higher probabilities of germination and establishment as they have more embryonic resources to help rooting under adverse conditions (Kigel & Galili, 1995; Forget *et al.*, 2005). If fitness generates high seed mass, it increases the probability of both seedling establishment and survival, and therefore occurrence. Analyses of seed addition projects worldwide for 84 large-seeded tree species demonstrate significant positive linear relationships ($P = 0.0026$) between seed mass and seedling recruitment probability (Moles & Westoby, 2002; Moles *et al.*, 2004) and hence species occurrence probability. Thus, a linear correlation between reproductive potentials and density metrics can be assumed valid.

Crop productivity is traditionally assessed as the outcome of physiological processes or available agronomic knowledge. To estimate seed yield for an undomesticated and unknown plant, such as *Jatropha*, we used a backward approach, where we derive seed mass productivity from one of its outcomes, species occurrence probability. This approach is validated with available on-site assessments of *Jatropha* yield. In this study, we seek to: (1) test significance and response functions for different environmental variables to predict *Jatropha* fitness, (2) map *Jatropha* seed yield under 'present'

climatic conditions (1950–2000 average) and near-future climate scenarios (year 2020) and (3) validate results of *Jatropha* fitness and seed yield, and illustrate prediction uncertainties.

Materials and methods

Specimens/species natural occurrence data

Specimen observations describing the location of *Jatropha* natural occurrences were retrieved from the World Biodiversity Information Network (WBIN, 2002), an online network of herbaria databases. Since specimen locations were retrieved from multiple herbaria sources, different geographical or ecosystem focuses were integrated, reducing therefore environmental sampling bias. The geo-locations of these specimens were verified for consistency with available location labels, when reported by specimen collectors. Erroneous geo-locations were either corrected with coordinates from gazetteers (Global Gazetteer, 2006), if specimen labels provided sufficient location detail, or were discarded from further analyses. Only those specimen locations from the second half of the twentieth century were used to maintain temporal correspondence between occurrence sampling and climate described by geodataset. 325 validated specimen locations were used (Fig. 1) from Central America (259 locations), South America (43 locations) and East Africa (23 locations). The samples from East Africa and South America, outside *Jatropha* native range (Maes *et al.*, 2009b), represent

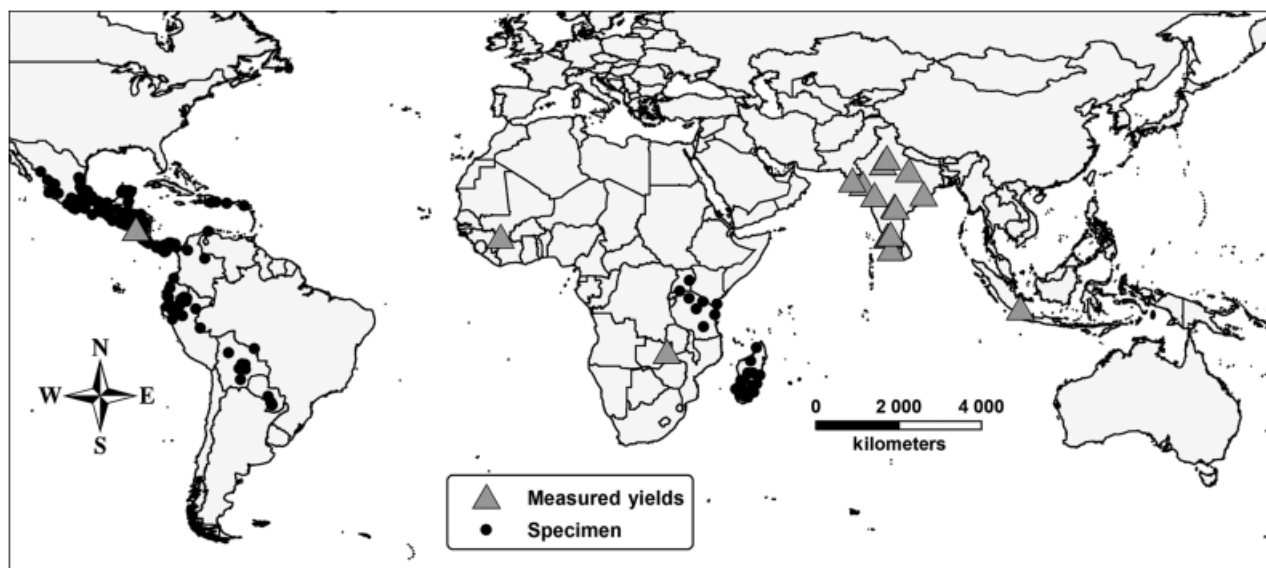


Fig. 1 Locations of *Jatropha* natural occurrence used for model training ('Herbarium Specimen') and locations of on-site yield assessments ('Measured Yields') used for model validation.

manly specimen of wild and vigorous individuals, reflecting favorable environmental conditions.

Environmental data

Several environmental geodatasets were acquired and evaluated for their usefulness to predict *Jatropha* ecological response. Bioclimatic factors are available from the WorldClim geodataset (Hijmans *et al.*, 2005) representing present (1950–2000) and future (2020, HADCM3 model with A2a emission scenario) conditions at high resolution (30 arc-seconds or ~ 1 km at equator). WorldClim is the best climate dataset available at high resolution for global studies, although in mountain and remote areas uncertainty is high as climate stations are sparse (Hijmans *et al.*, 2005). Slope was calculated from the CGIAR-CSI SRTM 90 m digital elevation model (Jarvis *et al.*, 2008), and was used as a potential explanatory variable, as it directly influences hydrological processes such as waterlogging and soil infiltration. Physical and chemical soil attributes were characterized using the 'Derived Soil Properties' dataset (5 arc-minute resolution, or ~ 10 km at equator) accompanying the FAO/UNESCO Digital Soil Map of the World (Nachtergaele *et al.*, 2007), developed for crop suitability studies on a global scale. All the global environmental geodatasets were aggregated to 1 min spatial resolution (~ 2 km at equator), which was the resolution used in our spatial modeling.

The Köppen climate classification, based on the observed interaction between climate and native vegetation (Köppen, 1923) is still widely used to define agronomic potentials across different climates (Gallup & Sachs, 2000). A recent Köppen classification map (Peel *et al.*, 2007) was used, and modified by masking out frost risk areas, to summarize *Jatropha* yield for climate zones.

Evaluation of environmental variables – data reduction

Responses of multivariate nonlinear models based on data with high collinearity can result in overfitting that can be highly exacerbated outside training conditions (Morlini, 2006). To increase transferability effectiveness, the number of explanatory factors was reduced to a limited set of more significant and less correlated variables. This allowed the creation of more generalized responses to the environment (and a more transferable model), with a balance between underfitted models with few parameters and overfitted models with too many correlated parameters (Burnham & Anderson, 2002). Furthermore, including highly correlated environmental variables adds limited information to explain *Jatropha* suitability, at the expense of additional error

associated with datasets (decreasing signal to noise ratio). The relative importance of each environmental variable was evaluated in MaxEnt by means of 'training gain', which is the improved predictability of MaxEnt based on the incorporation of a particular variable.

The use of slope as a predictor variable returned a high training gain with suitability decreasing as terrain flattens, as expected, due to increasing waterlogging risks. However, we noticed that the *Jatropha* suitability response to slope was not affected by changes in total precipitation. This suggests that the observed response to slope, rather than reflecting drainage suitability, may follow a likely topographic bias of natural occurrence, as trees have been removed preferentially in flatter areas for human infrastructures. Soil properties gave limited training gain to the *Jatropha* MaxEnt modeling, very likely because global soil datasets suffer of low accuracy (Gray *et al.*, 2009). Therefore, slope and soil variables were not used as explanatory variables.

The full set of 18 climatic predictor variables was reduced to a smaller subset by excluding those variables that (1) provided limited additional information (jackknife test of training gain with or without the variable) or (2) had the most complex bioclimatic meaning among highly correlated variable (Pearson's correlation > 0.85).

The final set of environmental dataset included eight bioclimatic variables: annual mean temperature (Mean T-Ann), mean diurnal temperature range (MeanTRng-Dly), maximum average temperature of the warmest month (MaxT-WrmMo), minimum average temperature of the coldest month (MinT-CldMo), annual precipitation (P-Ann), precipitation of wettest quarter (P-Wet Qrt), precipitation of driest quarter (P-DryQrt) and precipitation seasonality (standard deviation of monthly precipitation $\times 100$ over mean monthly precipitation, P-Seas).

Modeling Jatropha occurrence probability – MaxEnt parameterization

Except for those discussed below, default options were used for parameter selection in MaxEnt (v 3.2.19) as suggested by the model authors (Phillips & Dudik, 2008). Environmental responses were modeled from a training area including those countries where ≥ 5 *Jatropha* specimen locations were recorded (Fig. 1), and then spatially projected into global predictions. Pseudo-absence locations are sampled within the training area following the 'target-group absences' approach in order to minimize sampling biases and increase SDM predictive performances (Mateo *et al.*, 2010). From the available specimen locations, 80% (260 points) were randomly selected for training the MaxEnt model, while the remaining 20% (65 points) were used to validate the

predictions. The accuracy of MaxEnt predictions were tested within the training area using the AUC method (Fielding & Bell, 1997) on the subset of specimen locations reserved for validation.

Several features are implemented in MaxEnt to define the shape of the environmental response curves (Phillips *et al.*, 2006). The threshold feature was not used, thus the model was allowed to identify more generalized (and conservative for transferability) environmental constraints, more suitable for working at the broader global scale. Including the Threshold feature increases the risk of fitting too closely biased points in the specimen distribution or reflecting 'local' sharp impacts due to biotic competition (Pearson & Dawson, 2003).

The climatic responses for *Jatropha* suitability were derived and projected into two global maps of occurrence probability: one reflecting almost present climate conditions (1950–2000 average) and one for near-future climatic conditions (2020). Occurrence probability, characterizing *Jatropha* fitness, is presented as a logistic distribution varying from 0 to 1. Sensitivity of MaxEnt results was verified in a Monte Carlo analysis (50 iterations, due to computational limits) with different parameterization schemes: (1) with and without the Threshold feature and (2) climatic factors being randomly substituted by others excluded in the data reduction process. The sensitivity analysis produced local distributions of occurrence probability, from which the standard deviations are projected into a global sensitivity map.

Modeling Jatropha seed yield from occurrence probability

The linear relationship between seed yield and probability occurrence was adopted conservatively, in accordance with recruitment patterns observed globally in seed mass addition experiments (Moles & Westoby, 2002). Based on literature review and expert knowledge, the maximum seed yield for existing genotypes is assumed to be 5 ton dry seeds ha⁻¹yr⁻¹ for mature plantations under optimal environmental conditions (Achten *et al.*, 2008). A global *Jatropha* seed yield map has therefore been estimated by scaling the occurrence probability linearly into a yield range between 0 and 5 tons of dry ('air-dry') seeds (kernels) ha⁻¹yr⁻¹. The standard deviation of occurrence probability from Monte Carlo analysis is similarly scaled into the yield range to establish an uncertainty map of *Jatropha* yield.

Validation of Jatropha seed yield estimates

The estimated *Jatropha* seed yield was validated through linear regression fit with on-site yield assess-

ments (Fig. 1). Residuals, the absolute difference between estimated and measured yields, were used to validate the standard deviation of modeled sensitivity, which we used as predictor of yield estimate uncertainty. The modeled climate responses were tested by verifying whether the distribution of residuals was biased towards climatic factors.

Direct measurements of *Jatropha* yield are available from fifteen locations (Table 1), mainly commercial plantations, with plant age varying between 2.5 and 5 years. Mostly these locations had adequate agricultural inputs, in terms of pruning and fertilizer application. Measured yields were converted to yield at maturity, using a Chapman–Richards model (Chapman, 1961) developed from two measured chronosequences of *Jatropha* yield and plant age in Nicaragua (Foidl *et al.*, 1996) and Allahabad, India (own observation). The Chapman–Richards model simulates sigmoidal growth with an asymptotic peak in growth after a certain age:

$$Y = a(1 - e^{-b \times X})^c, \quad (1)$$

where Y is the organism growth quantity to evaluate (in our case, seed yield), X is age, and a , b and c are the asymptote or maximum achievable growth (seed yield at maturity), the growth rate, and the shape of the curve near the origin, respectively. The Chapman–Richards model regression for the two chronosequences derived similar growth coefficients for b (0.793 and 0.91) and c (3.344 and 3.588). This similarity suggests that a growth model with the average b and c (0.852 and 3.466) can be applied in both (and other) locations, where a is modified to match measured seed yields at specific ages and predict seed yield at maturity.

Results

Occurrence probability/fitness response to climate variables

MaxEnt uses trained responses from multiple variables altering simultaneously, a multivariate relationship being hard to represent in bi- or tri-dimensional graphics. Therefore the response in fitness for each climatic variable was graphically represented (Fig. 2) as the outcome of simplified MaxEnt simulations where only a single climatic variable was used as an explanatory factor.

Jatropha showed significant fitness (occurrence probability above 0.25) for MeanT-Ann between 18 and 29 °C, with optimal values of 26–27 °C (Fig. 2a). MinT-CldMo defined a strong increase in *Jatropha* fitness above 8–9 °C (Fig. 2b). Unlike MeanT-Ann

Table 1 On-site assessments of *Jatropha* seed yields for model validation. The measured yields are projected to predicted yields at maturity using a Chapman-Richards model we parameterized for *Jatropha* (see text)

Plantation location	Country	Longitude	Latitude	Plantation age (years)	Measured yield (kg dry seeds ha ⁻¹ yr ⁻¹)	Source	Predicted yield at maturity (kg dry seeds ha ⁻¹ yr ⁻¹)
Digini	Mali	-7.6320	10.9568	2.0	550	Wijgerse (2008)	1500
Bhavnagar	India	71.2494	22.4035	2.0	1270	Daey Ouwens (2007)	2400
León	Nicaragua	-86.7113	12.5894	4.0	2500	Foidl <i>et al.</i> (1996)	2750
Jaipur	India	75.8177	26.8871	2.5	313	Achten <i>et al.</i> (2008)	500
Andhra Pradesh	India	77.9420	17.4851	2.5	1000	Gexsi (2008)	1680
Bawal	India	76.8747	28.0605	3.0	208	NOVOD (2008)	300
Hyderabad	India	78.6298	17.2193	3.0	911	Rao (2006)	1400
Sadivayal	India	76.6173	10.8723	5.0	4000	Gunaseelan (2009)	4000
Allahabad	India	82.5104	25.5020	4.0	2000	Achten <i>et al.</i> (2008)	2300
Nashik	India	73.9256	20.0196	5.0	1200	Wani (2006)	1200
Tamil Nadu	India	76.8758	11.0912	3.0	1573	D1 Oils (personal communication)	2000
Bandar Surubaya	Indonesia	105.6509	-4.7042	2.0	1000	D1 Oils (personal communication)	1950
Kapiri Mposhi	Zambia	28.4570	14.3603	2.5	500	D1 Oils (personal communication)	780
Tirunelveli	India	77.5326	8.2741	3.0	2000	D1 Oils (personal communication)	2600
Erode	India	77.4345	11.3801	2.5	350	D1 Oils (personal communication)	550

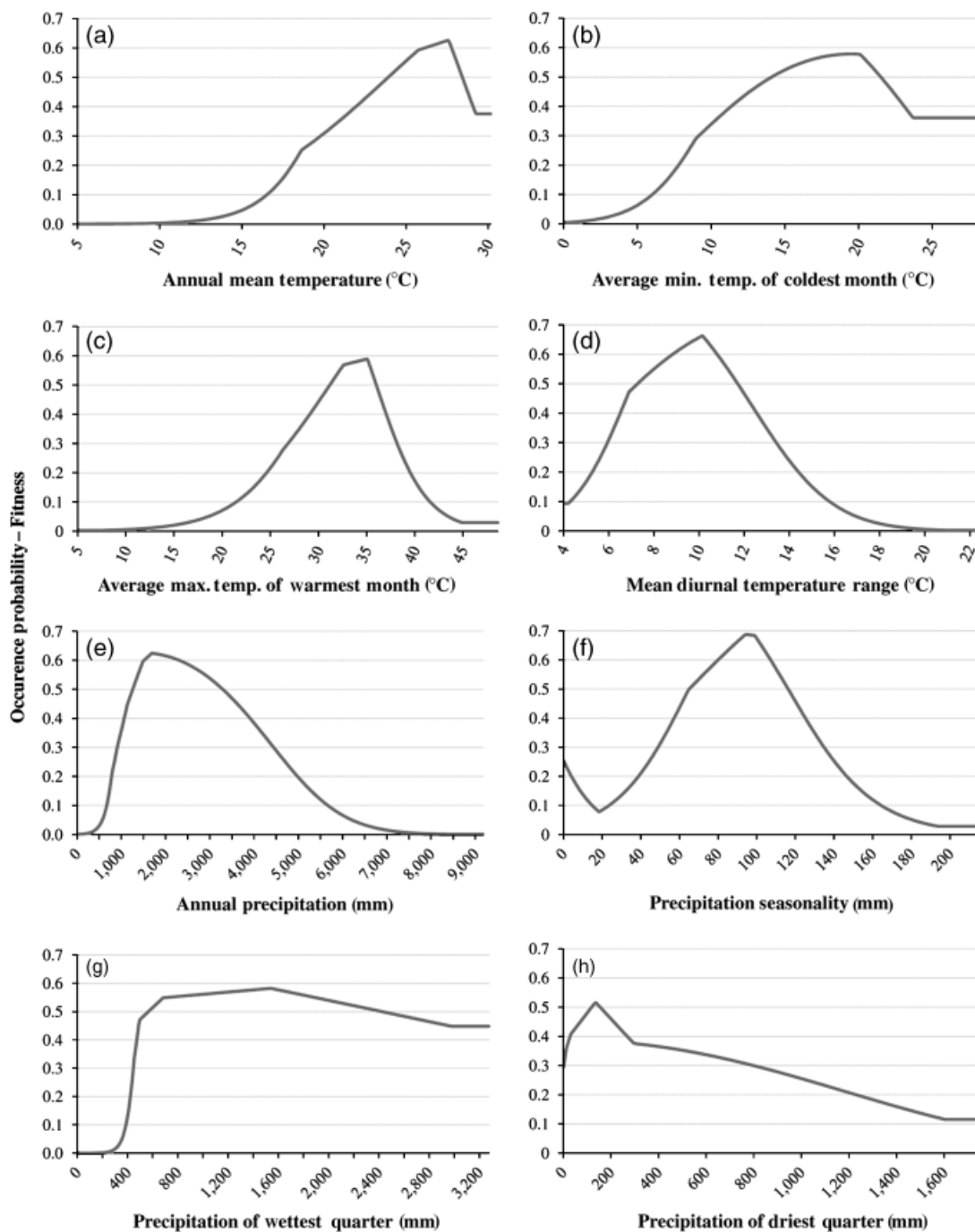


Fig. 2 Response of Jatropha fitness (logistic distribution of occurrence probability) to bioclimatic variables.

and MinT-CldMo, MaxT-WrmMo was less relevant for overall Jatropha fitness, as it did not decrease much the overall training gain when omitted. However, the simulation graph did show fitness increased where MaxT-WrmMo falls between 35 and 45 °C (Fig. 2c). Jatropha fitness increased when MeanTRng-Dly exceeded 4 °C but was <12–14 °C (Fig. 2d).

The environment for Jatropha was more suitable when P -Ann exceeded 600–900 mm, with an optimum at 1500 mm followed by a slow decrease as precipitation further increases (Fig. 2e). We observed high Jatropha fitness when P -WetQrt is between 500 and 1400 mm with slowly decreasing fitness for higher precipitations (Fig. 2f). P -DryQrt response denoted the highest

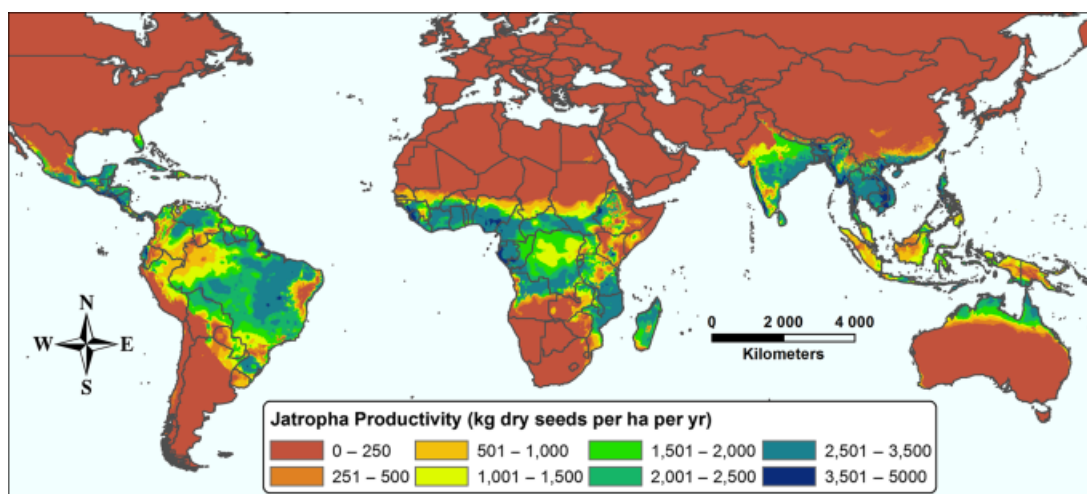


Fig. 3 Estimated *Jatropha* productivity ($\text{kg dry seeds ha}^{-1} \text{ yr}^{-1}$) for present climatic conditions (1950–2000 average).

preference for small levels of precipitation (200 mm) followed by a strong decreasing in fitness as precipitation in the dry period increased (Fig. 2g). Fitness for *P*-Seas peaked when the standard deviation of monthly precipitation is analogous to the monthly average precipitation, $P\text{-Seas} \approx 100$ (Fig. 2h).

The environmental variables that caused the greatest decrease in training gain when omitted were *P*-Seas and MeanTRng-Dly. Therefore, these variables had very specific information to explain *Jatropha* distribution that was not present in other variables.

Jatropha seed yield map

Jatropha has mainly a pan-tropical distribution, favoring areas with a dry season (Fig. 3). Additional suitable areas are located in warm temperate climate zones with sufficient rainfall. Such areas are mainly in the south of Brazil, in the north of Argentina, and in a strip dominated by monsoons extending from south of the Himalayas to the South China Sea.

Results aggregated into relevant Köppen climate classes (Table 2) denote high yields in those areas characterized by a tropical climate (mean temperature of the coldest month $> 18^\circ\text{C}$), either monsoonal (Am) or with a distinct dry season (Aw). However, outside the tropics, suitable growing opportunities for *Jatropha* are found in warm temperate climates with no frost risk, characterized by either having dry seasons (Cw) or being fully humid (Cf). Other climates, such as tropical climates with no dry seasons (Af) and subtropical deserts (Bsh) have moderate yield potential.

The standard deviation of yield results (Fig. 4) from Monte Carlo analysis describes modeling sensitivity and potential uncertainties. Areas with highest sensi-

Table 2 Average predicted *Jatropha* productivity within Köppen climate zones*

Köppen climate zone * (Peel <i>et al.</i> 2007)	<i>Jatropha</i> average productivity ($\text{kg dry seeds ha}^{-1} \text{ yr}^{-1}$)
Tropical humid climates	
Af (tropical wet – no dry season)	1150
Am (tropical monsoonal – short dry season; heavy monsoonal rains in other months)	2200
Aw (tropical savanna – winter dry season)	2300
Dry climates	
BSh (subtropical steppe – low-latitude)	750
Subtropical temperate climates	
Cw (humid subtropical – dry winter)	1950
Cf (humid subtropical/ marine without dry season – hot or warm summers)	1550

*Original classification has been modified to exclude frost risk conditions (minimum average temperature of the coldest month $< 8^\circ\text{C}$).

tivity are those where climate regimes differ the most from the ones encountered in the training area. Among areas deemed suitable, the highest sensitivity is observed in subtropical areas (i.e. North India and South Brazil), and in tropical areas of the eastern coast of India, South-East Asia and on the coasts of Guinea, Guinea-Bissau and Sierra Leone.

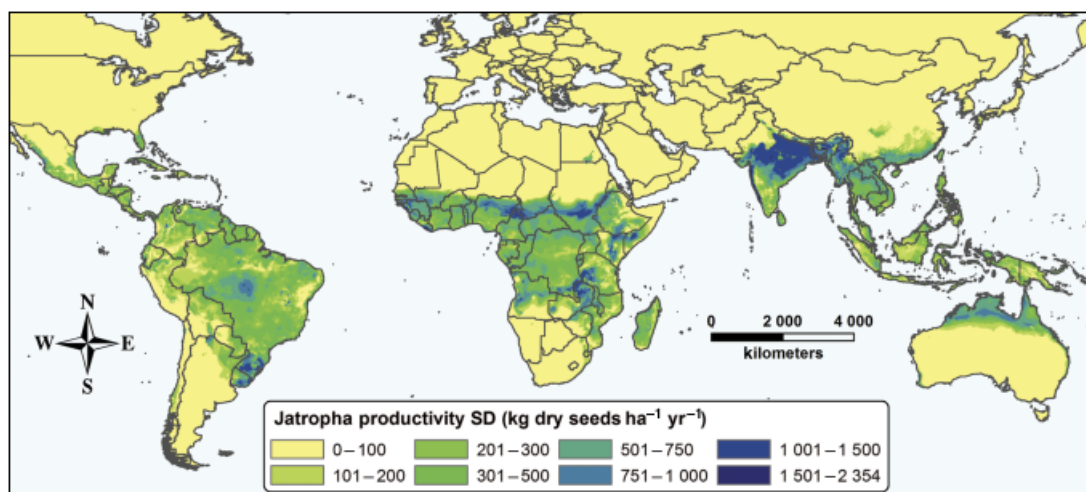


Fig. 4 Standard deviation of *Jatropha* productivity (kg dry seeds ha⁻¹ yr⁻¹) from Monte Carlo sensitivity analysis. Areas with highest standard deviation show the highest prediction uncertainty.

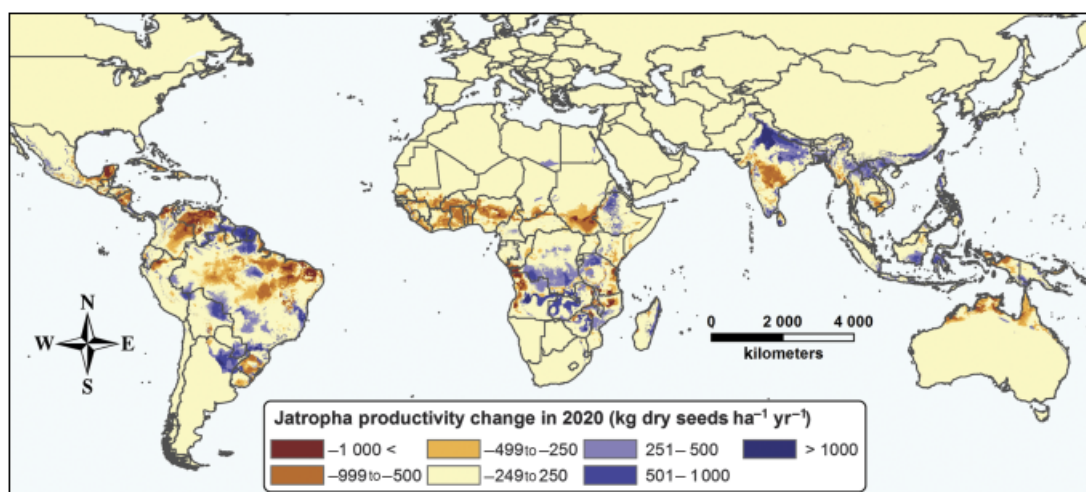


Fig. 5 Change in estimated *Jatropha* yield (kg dry seed ha⁻¹ yr⁻¹) between future (2020) and present climatic conditions (1950–2000 average).

Changes in seed yield due to climate changes in the near future (Fig. 5) follow clear subcontinental patterns with yield decreasing in zones with lower predicted precipitations (Sahel, Eastern Brazil, Northern Australia), and yield increasing in regions with decreasing frost risk (Northern part of India and East Asia, and Southern Africa). Average global yield in areas suitable for *Jatropha* will decrease slightly due to climate change (~11% or 300 kg seeds ha⁻¹ yr⁻¹).

Validation

Validation of estimated MaxEnt probability occurrence reported AUC values of 0.92 when evaluated on testing

data (65 specimen locations). AUC values theoretically equals 1 for perfect models, while for AUC values equal to 0.5 or less, models are considered random or worse than random. In general, when the AUC values is > 0.9, a probabilistic classification is considered reliable (Araújo *et al.* 2005). Linear regression between measured vs. estimated yields (Fig. 6a) showed a good fit ($R^2 = 0.674$, $P < 0.001$). When plotting the residuals of this regression against the standard deviation of productivity obtained from the sensitivity analysis (Fig. 6b), we obtained a significant linear regression fit ($R^2 = 0.569$, $P = 0.001$). Thus, approximately half of the residual values can be explained by modelling uncertainties and the remaining half by unaccountable explanatory variables (e.g. nutrients availability). Residuals do not show a significant

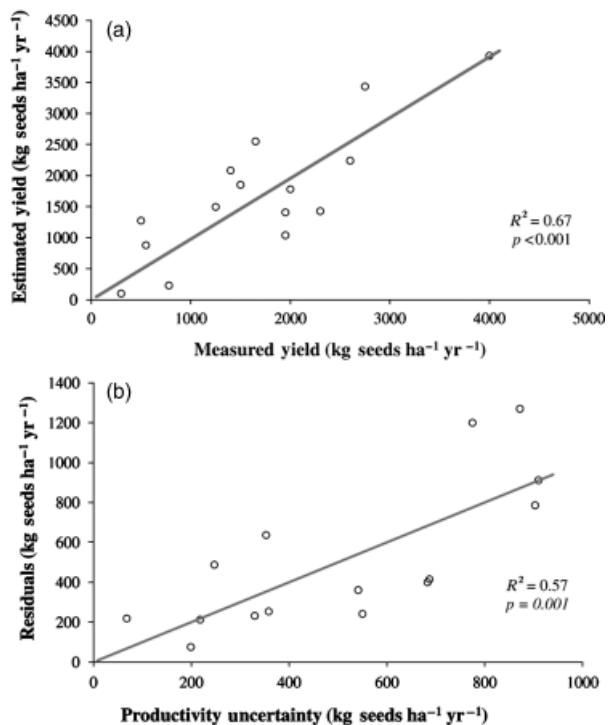


Fig. 6 (a) Linear regression fit of estimated versus measured Jatropha seed yields, (b) linear regression fit of residuals (|estimated–measured yields|) versus productivity uncertainty (standard deviation of Jatropha yield from sensitivity analysis).

linear trend across any of the climatic variables used to simulate occurrence probability ($R^2 = 0.1–0.3$). Therefore, it can be assumed that our model reliability is unbiased towards climate factors and thus appropriate for studying climatic influences.

Discussion

Several environmental variables describing soil, topographic and climatic conditions were tested as explanatory factors of Jatropha fitness. However, we could infer significant and reliable responses only from climate datasets. Lack of explanatory significance for edaphic factors is very likely due to the inaccuracy typical of available global soil datasets (Gray *et al.*, 2009). In reality, soil properties have an important effect on Jatropha productivity (Openshaw, 2000). We also have excluded slope as explanatory variable, although significant response was reported by its use. The modeled response to slope shows increasing suitability towards terrains with medium slopes ($20–30^\circ$), but likely reflecting a topographic bias of natural vegetation occurrence and its sampling.

In accordance with previous observations (Achten *et al.*, 2008), we found that Jatropha establishment requires mean annual temperatures between 18 and

28°C (with optimal values around $26–27^\circ\text{C}$) and average minimum temperatures of the coldest month above $8–9^\circ\text{C}$, indicating a clear lack of tolerance to frost (Sys *et al.*, 1991). Furthermore, our results show preferences for average maximum temperatures in the warmest period varying between 35 and 45°C . However, maximum temperatures had a marginal significance to predict suitability. This implies that Jatropha is slightly constrained by maximum temperatures, which is a trait typical of succulent plants (Ackerly *et al.*, 2000). Suitable conditions were found with annual precipitation above $600–900$ mm, which is higher than early expectations (Foidl *et al.*, 1996), with an optimum at 1500 mm. While Jatropha has a drought avoidance strategy and a relatively high water-use efficiency (Maes *et al.*, 2009a), no scientific study has confirmed previous high yields expectations for arid and semiarid areas. Fitness decreased when annual precipitation exceeds 1500 mm, probably because of increased risk of waterlogging (Biswas *et al.*, 2006; Singh *et al.*, 2006) or a lower likelihood of drought. It is a common belief among Jatropha practitioners that seasonal droughts are somehow functional for higher yields. In fact, our results show specific preferences for limited precipitation in the driest quarter (optimum at $50–200$ mm) and seasonal precipitation regimes (optimum precipitation seasonality at ~ 100). When precipitation seasonality approaches 100 , the standard deviation of the monthly precipitation distribution approximates its average. Jatropha showed a preference for a mild mean daily temperature range ($4–12^\circ\text{C}$), generally found in hot but not extremely dry zones (e.g. subject to oceanic influences).

Jatropha seed yield estimates (Fig. 3 and Table 1) showed that suitable zones for high yields are located in both tropical and hot temperate areas (with sufficient precipitation and absence of frost). Results show that average yield of Jatropha in tropical climates was halved when planted in fully humid zones rather than humid tropics with a strong or partial dry season. Vegetation on coastal ranges benefits from horizontal precipitations (e.g. sea-fog) from nearby water bodies (Bruijnzeel & Proctor, 1995). Owing to the lack of climate geodatasets describing these events, results exclude the positive impacts of horizontal precipitations on Jatropha suitability.

The average global yield of Jatropha is estimated to decrease slightly ($\sim 11\%$) under current climate change (i.e. difference between average 1950–2000 and 2020 conditions), although greater effects can be seen within specific regions. Cold days, cold nights, and frost events will become on average less frequent as climate changes (IPCC 2007). This will make areas of Southern Africa (e.g. Zambia), South America (e.g. Argentina, Paraguay), the northern part of South and East Asia (e.g. Northern India,

Nepal and China) more suitable for *Jatropha*. Similarly, projected precipitation decreases in already dry conditions (e.g. Sahel) may reduce *Jatropha* suitability.

Our MaxEnt modeling effort was based on a set of primary data (259 specimen locations) considered quite sufficient to model correctly suitability distribution over environmental conditions encountered in the training area (Pearson & Dawson, 2003; Loiselle *et al.*, 2008). In fact, validation of MaxEnt results over training environmental conditions was rated very positively (AUC = 0.92). As we move away from these environmental conditions, higher model sensitivity to MaxEnt parameterization (and uncertainty) is observed. The global MaxEnt estimate of occurrence probability is converted linearly into seed yield, following observed worldwide ecological patterns for 64 large-seeded tree species (Moles & Westoby, 2002). A limited number of on-site yield assessment ($n = 15$) is available. Luckily, the locations of these assessments follow an ideal distribution (i.e. outside training area, across a gradient of model uncertainty) that maximizes validation effectiveness. Measured and estimated yields show a strong linear fit ($R^2 = 0.67$). The yield data were found in locations where the annual precipitation ranged from 400 to 1500 mm. As such, they provide a valid benchmark for semiarid and subhumid areas, where is the major interest in *Jatropha*.

Since the residuals analysis showed no significant trend ($R^2 = 0.1-0.3$) in error across the climate gradients, our model provides unbiased and conservative *Jatropha* responses to climate factors. The spatial assessment of sensitivity provides a tool to detect uncertainty within predicted results, inherent to any modeling, allowing risk assessment of their use by land managers and policy makers.

It should be noted that our results describe *Jatropha* suitability to climate factors from an ecological perspective. Therefore, they do not include the potential of agronomic practices to mitigate climatic stresses and widen the area suitable to *Jatropha*. Overall the specimen collection corresponds to individuals naturally occurring, and therefore responses are calibrated for a range of genotypes adapted to local climatic conditions in the training area. Although, agronomic and genetic improvements can increase *Jatropha*'s global potential, local performances will always be primarily constrained by the biophysical environment. As such this study can be used to identify locations where the basic requirements for *Jatropha* yield are met, or to indicate, for a certain location, which essential factor has to be artificially adapted (e.g. irrigation for precipitation). With this exploration of the occurrence and yield response to climatic conditions, further research can focus on the response of management inputs (e.g. agronomic practices, genetic improvement).

This study has successfully produced a spatial high resolution (~2km) global estimate of *Jatropha* seed yield under both current and future climate conditions by applying principles of biogeography and ecology to the limited data available for *Jatropha*. We have acquired further information on the species environmental requirements based on analysis of a near global distribution of the species. A measure of uncertainty has been provided to allow appropriate and informed use of results.

Acknowledgements

This research was funded by the K. U. Leuven Research Fund, International Water Management Institute (IWMI), Flemish Interuniversity Council – University Development Co-operation (VLIR-UOS) and the Research Foundation – Flanders (FWO). The work is a contribution of the R&D Platform on Climate and Development Cooperation (KLIMOS) and the JATROPHABILITY project. We treasure the thoughtful suggestions from the anonymous reviewers and the valuable field assessments of D1 Oils Plant Science NV.

References

- Ackerly DD, Dudley SA, Sultan SE *et al.* (2000) The evolution of plant ecophysiological traits: recent advances and future directions. *BioScience*, **50**, 979–995.
- Achten WMJ, Maes WH, Aerts R *et al.* (2010) *Jatropha*: from global hype to local opportunity. *Journal of Arid Environments*, **1**, 164–165.
- Achten WMJ, Verchot L, Franken YJ *et al.* (2008) *Jatropha* biodiesel production and use. *Biomass and Bioenergy*, **32**, 1063–1084.
- Araujo MB, Whittaker RJ, Ladle RJ, Erhard M. (2005) Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, **14**, 529–538.
- Biswas S, Kaushik N, Srikanth G (2006) Biodiesel: technology and business opportunities – an insight. In: *Proceedings of the Biodiesel Conference Toward Energy Independence—Focus of Jatropha* (eds Singh B, Swaminathan R, Ponraj V), pp. 303–330. Rashtrapati Bhawan, New Delhi, India.
- Bondeau A, Smith PC, Zaehle S *et al.* (2007) Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology*, **13**, 679–706.
- Boyer JS (1982) Plant productivity and environment. *Science*, **218**, 443–448.
- Brisson N, Gary C, Justes E (2003) An overview of the crop model STICS. *European Journal of Agronomy*, **18**, 309–332.
- Bruijnzeel LA, Proctor J (1995) Hydrology and biogeochemistry of tropical montane cloud forests: what do we really know? In: *Tropical Montane Cloud Forests* (eds Hamilton LS, Juvik JO, Scatena FN), pp. 38–78. Springer-Verlag, New York.
- Burnham KP, Anderson DR (2002) *Model Selection and Multi-Model Inference. A Practical Information—Theoretic Approach*, 2 edn. Springer-Verlag, Heidelberg.

- Chapman DG (1961) In: *Proceedings of 4th Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 4 (ed. Neyman J), pp. 153–168. Berkeley, CA.
- Daey Ouwens K, Francis G, Franken YJ *et al.* (2007) Position paper on *Jatropha curcas*. State of the art, small and large scale project development. Fact Foundation.
- Elith J, Graham CH, Anderson RP *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Fairless D (2007) Biofuel: the little shrub that could – maybe. *Nature*, **449**, 652–655.
- FAO (1983) *Guidelines: Land Evaluation for Rainfed Agriculture*. FAO, Rome.
- FAO (1985) *Guidelines: Land Evaluation for Irrigated Agriculture*. FAO, Rome.
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.
- Foidl N, Foidl G, Sanchez M, Mittelbach M, Hackel S (1996) *Jatropha curcas* L. as a source for the production of biofuel in Nicaragua. *Bioresource Technology*, **58**, 77–82.
- Forget PM, Lambert JE, Hulme PE, Vander Wall SB (2005) *Seed Fate Predation, Dispersal, and Seedling Establishment*. CABI Pub., Wallingford, Oxfordshire, UK.
- Gallup JK, Sachs JD. (2000) Agriculture, climate, and technology: why are the Tropics Falling Behind? *American Journal of Agricultural Economics*, **82**, 731–737.
- Gaston KJ, Borges Pa, He F, Gaspar C (2006) Abundance, spatial variance and occupancy: arthropod species distribution in the Azores. *The Journal of Animal Ecology*, **75**, 646–656.
- GEXSI (2008) *Global market study on Jatropha – final report*. GEXSI LLP, Berlin, Germany:
- Global Gazetteer (2006) *Global gazetteer, version 2.2*. Falling Rain Genomics Inc. Database. Available at: <http://www.fallinrain.com/world/index.html> (accessed 1 December 2008).
- Graham CH, Elith J, Hijmans RJ, Guisan A, Peterson AT, Loiselle BA (2008) The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, **45**, 239–247.
- Gray JM, Humphreys GS, Deckers JA (2009) Relationships in soil distribution as revealed by a global soil database. *Geoderma*, **150**, 309–323.
- Gunaseelan VN (2009) Biomass estimates, characteristics, biochemical methane potential, kinetics and energy flow from *Jatropha curcas* on dry lands. *Biomass Bioenergy*, **33**, 589–596.
- Hartl DL (2000) *A Primer of Population Genetics*. Sinauer Associates, Sunderland, MA.
- He F, Gaston KJ (2003) Occupancy, spatial variance, and the abundance of species. *American Naturalist*, **162**, 366–375.
- Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773–785.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Hutchinson GE (1957) Concluding remarks. *Cold Spring Harbor Symposium on Quantitative Biology*.
- IPCC (2007) *The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge.
- Jarvis A, Reuter HI, Nelson A, Guevara E (2008) Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m. Database. Available at: <http://www.srtm.csi.cgiar.org> (accessed 1 September 2008).
- Kigel J, Galili G (1995) *Seed Development and Germination*. Marcel Dekker, New York.
- Köppen W (1923) *Die Klimate der Erde*. Walter de Gruyter. Walter de Gruyter, Berlin.
- Loiselle BA, Jørgensen PM, Consiglio T *et al.* (2008) Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes? *Journal of Biogeography*, **35**, 105–116.
- Mackey BG, Lindenmayer DB (2001) Towards a hierarchical framework for modelling the spatial distribution of animals. *Journal of Biogeography*, **28**, 1147–1166.
- Maes WH, Achten WMJ, Reubens B, Raes D, Samson R, Muys B (2009a) Plant–water relationships and growth strategies of *Jatropha curcas* L. seedlings under different levels of drought stress. *Journal of Arid Environments*, **73**, 877–884.
- Maes WH, Trabucco A, Achten WMJ, Muys B (2009b) Climatic growing conditions of *Jatropha curcas* L. *Biomass and Bioenergy*, **10**, 1481–1485.
- Mateo RG, Croat TB, Felicísimo ÁM, Muñoz J (2010) Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. *Diversity and Distributions*, **1**, 84–94.
- Moles AT, Falster DS, Leishman MR, Westoby M (2004) Small-seeded species produce more seed per square meter of canopy per year, but not per individual lifetime. *Journal of Ecology*, **92**, 384–396.
- Moles AT, Westoby M (2002) Seed addition experiments are more likely to increase recruitment in larger-seeded species. *Oikos*, **99**, 241–248.
- Morlini I (2006) On multicollinearity and concurrency in some nonlinear multivariate models. *Statistical Methods and Applications*, **15**, 3–26.
- Nachtergaele FO, Zanetti M, Bloise M, Ataman E (2007) Global Land Resources GIS Models and Databases for Poverty and Food Insecurity Mapping (accessed 1 November 2008).
- National Oilseeds and Vegetable Oils Development (NOVOD) Board (2008) *Third Research and Development Report on Tree Borne Oilseeds 2007–2008*. NOVOD Board, Ministry of Agriculture, Government of India, New Delhi, India.
- Ndong R, Montrejeaud-Vignoles M, Girons OS *et al.* (2009) Life cycle assessment of biofuels from *Jatropha curcas* in West Africa: a field study. *Global Change Biology and Bioenergy*, **1**, 197–210.
- Openshaw K (2000) A review of *Jatropha curcas*: an oil plant of unfulfilled promise. *Biomass and Bioenergy*, **19**, 1–15.
- Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.

- Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.
- Peel MC, Finlayson BL, McMahon TA (2007) Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences*, **11**, 1633–1644.
- Phillips SJ (2008) Transferability, sample selection bias and background data in presence-only modelling: a response to Peterson et al (2007). *Ecography*, **31**, 272–278.
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Phillips SJ, Dudik KM (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.
- Phillips SJ, Dudik M, Schapire RE (2004) A maximum entropy approach to species distribution modeling. In: *Proceedings of the Twenty-First International Conference on Machine Learning*. ACM Press, New York, USA.
- Planning Commission of India (2003) *Report of the committee on development of bio-fuel*. Planning Commission, Government of India, New Delhi, India.
- Preston TR (2009) Environmentally sustainable production of food, feed and fuel from natural resources in the tropics. *Tropical Animal Health and Production*, **41**, 1071–1080.
- Rao VR (2006) The *Jatropha* hype: promise and performance. In: *Proceedings of the Biodiesel Conference Toward Energy Independence—Focus of Jatropha, Hyderabad, India, June 9–10* (eds Singh B, Swaminathan R, Ponraj V), pp. 16–20. Rashtrapati Bhawan, New Delhi.
- Reddy S, Davalos LM (2003) Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, **30**, 1719–1727.
- Reekie EG, Bazzaz FA (2005) *Reproductive Allocation in Plants*. Elsevier Academic Press, Burlington, MA.
- Rey PJ, Alcantara JM (2000) Recruitment dynamics of a fleshy-fruited plant (*Olea europaea*): connecting patterns of seed dispersal to seedling establishment. *Journal of Ecology*, **88**, 622–633.
- Rossiter DG (1996) A theoretical framework for land evaluation. *Geoderma*, **72**, 165–190.
- Royle JA, Nichols JD, Kéry M (2005) Modelling occurrence and abundance of species when detection is imperfect. *Oikos*, **110**, 353–359.
- Silvertown JW, Charlesworth D (2001) *Introduction to Plant Population Biology*. Blackwell Science, Oxford.
- Singh L, Bargali SS, Swamy SL (2006) Production Practices and Post-Harvest Management in *Jatropha*. In: *Proceedings of the Biodiesel Conference Toward Energy Independence – Focus of Jatropha* (eds Singh B, Swaminathan R, Ponraj V), pp. 252–267. Rashtrapati Bhawan, New Delhi, India.
- Sys C, Van Ranst E, Debaveye J (1991) Land evaluation Part I: principles in land evaluation and crop production calculations. In: *Agricultural Publication No. 7*. General Administration for Development Cooperation, Brussels, Belgium.
- Tan G, Shibasaki R (2003) Global estimation of crop productivity and the impacts of global warming by GIS and EPIC integration. *Ecological Modelling*, **168**, 357–370.
- VanDerWal J, Shoo LP, Johnson CN, Williams SE (2009) Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist*, **174**, 282–291.
- Vermeij GJ (2004) *Nature: An Economic History*. Princeton University Press, Princeton, NJ.
- Wani SP, Osman M, Silva ED, Sreedevi TK (2006) Improved livelihoods and environmental protection through biodiesel plantations in Asia. *Asian Biotechnology and Development Review*, **8**, 11–29.
- Wijgerse I (2008) The electricity system for a rural village in Mali. Msc Thesis, Delft University of Technology, Delft, 104 pp.
- World Biodiversity Information Network (WBIN) (2002). Online network of herbarium databases, supported by CONABIO (National Commission for the Knowledge and Use of Biodiversity), Mexico. Available at: http://www.conabio.gob.mx/remib_ingles/doctos/remib_ing.html; 2008 (accessed 19 December 2008).
- Zielinski WJ, Kucera TE (1995) *American Marten, Fisher, Lynx, and Wolverine: Survey Methods for their Detection*. Pacific Southwest Research Station, Albany, CA.