Inland Norway University of Applied Sciences

Faculty of Applied Ecology, Agricultural Sciences and Biotechnology

Kauê de Sousa

Agrobiodiversity and climate adaptation:

insights for risk management in small-scale farming

PhD Dissertations in Applied Ecology and Biotechnology
2020

Ph.d-avhandlinger i anvendt økologi og bioteknologi
Phd Dissertations in Applied Ecology and Biotechnology

No. 1 **Kauê de Sousa:** Agrobiodiversity and climate adaptation: insights for risk management in smallscale farming

No. 2 **Gjermund Gomo:** The Vertebrate Scavenger Community Along a Boreal Forest-Alpine Gradient: The Importance of Ungulate Management, Small Rodent Cycles and Winter Climate

Kauê de Sousa

Agrobiodiversity and climate adaptation: insights for risk management in small-scale farming

PhD Thesis

2020

Faculty of Applied Ecology, Agricultural Sciences and Biotechnology

Inland Norway
University of
Applied Sciences

Printed by: Flisa Trykkeri A/S

Place of publication: Elverum

PhD Thesis in Applied Ecology and biotechnology no. 1

ISBN printed version: 978-82-8380-230-6

ISBN digital version: 978-82-8380-231-3

ISSN printed version: 2703-819X

ISSN digital version: 2703-8181

Abstract

Agriculture is a dynamic activity that sustains food and other goods for global human population. Aiming to ensure global food security, the sector has evolved dramatically, especially over the last century with the introduction of high-yielding crops, improved technology and pathogen resistant varieties, to name a few. According to the Food and Agriculture Organisation of the United Nations (FAO), the food security issue is still present. In 2019, around 26% of the world population experienced either a moderate or severe level of food insecurity. Climate change makes the challenge of food security even more pressing. It is argued that increased agrobiodiversity through farm diversification and varietal selection can help farmers to cope with the negative effects of climate change while ensuring food security. However, such approaches have been difficult to scale up. One could argue that we often lack information to understand the contexts that drive farmers' adaptation decisions and how to develop recommendations for adaptation. In this thesis, I developed methods and tools to support farmers and stakeholders in adapting to a changing climate. I present results from three continents to improve the understanding of the food systems at the farm level, and specifically in smallholder farming. I provide insights for the different biological levels: species level, focusing on trees as slow grower organisms for interspecific diversification; varieties level, looking for locally adapted phenotypes; and genotype level, focusing on $G \times E$ interactions to support crop breeding for intraspecific diversification. From the first part of the study, conducted in Central America, the results showed that farmers have a clear preference to a set of adaptation strategies, with reforestation (agroforestry) as the first choice (Paper 1). Crop variety management is the least preferred choice of the top-5. In the second part of the study, I assessed the impacts of climate change on the habitats of the 100 most common tree species used in coffee (*Coffea arabica* L.) and cocoa (*Theobroma cacao* L.) agroforestry in Central America (Paper 2). The results showed that the most preferred trees are the most vulnerable, but farmers could re-think the agroforestry designs using a portfolio of underutilised species already present in low densities at the current systems. In the third part of the study, I employed a citizen science approach that can scale variety testing and help farmers to select the right crop variety for their farms (Paper 3). I tested this approach with common bean (*Phaseolus vulgaris* L.) in Nicaragua, bread wheat (*Triticum aestivum* L.) in India and durum wheat (*Triticum durum* Desf.) in Ethiopia. The results showed that the approach reduces geographic sampling bias and could be scaled to provide tailored recommendations for crop variety management. I also show, with durum wheat genotypes in Ethiopia, that linking the farmer-generated data to scientist-generated data can support breeding programs targeting challenging crop production environments using a data-driven decentralised approach (Paper 4). The approach is fully replicable, and part of its workflow is presented in this thesis (Paper 5 and Paper 6). Overall, the results of this thesis should be seen as starting point to develop lines of research that support recommendations to adapt agricultural systems to a changing climate.

Sammendrag

Landbruk er en dynamisk aktivitet som skal sikre mat og andre varer som verdens befolkning til enhver tid trenger. Med global matsikkerhet som mål har landbruket utviklet seg mye, særlig det siste århundre, blant annet ved å ta i bruk høytytende grøder, forbedret dyrkningsteknikk og sorter som er resistente mot sykdommer. I følge FNs organisasjon for ernæring og landbruk (FAO) er matsikkerhet fortsatt et aktuelt tema. I 2019 opplevde rundt 26% av verdens befolkning en moderat eller alvorlig grad av usikkerhet rundt tilgangen på mat. Klimaendringer gjør utfordringene rundt matsikkerhet enda mer krevende. Hele matvaresystemet må endres for å takle klimaendringer og samtidig sikre nok mat til alle. Det hevdes at økt biologisk mangfold i landbruket kan hjelpe bønder i å takle klimaendringene og samtidig sikre matproduksjonen, dette gjennom mer variasjon i hva som dyrkes på gårdsnivå og gjennom et bedre utvalg av sorter. Slike tilnærminger har imidlertid vist seg å være vanskelige å skalere opp. Man kan hevde at vi ofte mangler tilstrekkelig med informasjon for fullt ut kunne forstå hva som avgjør bønders valg knytta til klimatilpasning - og videre hvordan man så skal kunne utvikle rådgivingen for dette. I denne avhandlingen har jeg utviklet metoder og verktøy som kan brukes for å hjelpe bønder og andre i å tilpasse seg til endringer i klima. Jeg presenterer resultater fra tre ulike kontinent, dette for gi eksempel på hvordan en økt forståelse av matvaresystemene kan fungere på gårdsnivå, og særlig på små gårder. Jeg går inn på ulike biologiske nivå: på artsnivå, med fokus på trær som vokser langsomt og som gir stor diversitet mellom arter; på sortsnivå, ved å søke å finne lokalt tilpassa fenotyper; og på genotypenivå, ved å fokusere på samspillet mellom gener og miljø ($G \times E$), dette for å støtte foredlingsarbeid for økt diversitet innenfor arter. Resultater fra første delen av studiet som ble gjennomført i Mellom-Amerika viste at bønder har en klar preferanse for et sett av strategier i forhold til klimatilpasning, med agroskogbruk som førstevalg (Artikkel 1). Sortsvalg er det minst foretrukne valget av topp fem. I andre del av studiet undersøkte jeg hvor egnet dagens vokseplasser i Mellom-Amerika er for de 100 vanligste trærne som anvendes innenfor agroskogbruket med kaffe (*Coffea arabica* L.) og kakao (*Theobroma cacao* L.), dette med tanke på framtidige klimascenarier (Artikkel 2). Resultatene viste at de mest foretrukne trærne er de mest sårbare og at bøndene burde tenke nytt i forhold til utforming av agroskogbruket, dette ved å ta i bruk en rekke mindre anvendte arter som likevel finnes i dagens system. I tredje del av studiet anvendte jeg grasrotforskning som tilnærming for å skalere opp sortsforsøk og hjelpe bønder med å velge riktig sort for gårdene sine (Artikkel 3). Jeg undersøkte dette i bønner (*Phaseolus vulgaris* L.) i Nicaragua, vanlig brødhvete (*Triticum aestivum* L.) i India og durumhvete (*Triticum durum* Desf.) i Etiopia. Resultatene viste at en slik tilnærming reduserer feilkilder knyttet til geografisk representasjon og kan skaleres opp for å gi mer skreddersydde løsninger for bruk av sorter. I arbeidet med ulike genotyper av durumhvete i Etiopia viser jeg at foredlingsprogrammer kan styrkes at ved å koble grasrot-genererte data til forsker-genererte data gjennom en desentralisert tilnærming (Artikkel 4). Tilnærmingen er fullt mulig å gjenta, og en del av arbeidsflyten er presentert i avhandlingen (Artikkel 5 og 6). Samlet sett bør resultatene fra denne avhandlingen sees som en start på å utvikle en forskningen som kan bistå med anbefalinger slik at landbruket bedre kan tilpasse seg til et klima i endring.

Preface

This thesis summarises an amazing chapter of my personal and professional life. Growing up in the middle of the Brazilian Amazon, I always had an interest in plants, first to play with but later to actually study and research how they function and grow. My grandmother was my sponsor, supporting my interest to cultivate and harvest beans in our backyard. From my parents I learned that the way people interact with the environment is also an important aspect to look at. As I grew up, I met scientists that helped me to put these lessons into a more organised and scientific way. First at the Federal University of Amazonas (UFAM, Manaus, Brazil) and the National Institute of Amazonian Research (INPA, Manaus, Brazil) where I realised that, yes, I want to be a scientist. For this, thank you all. At the Tropical Agricultural Research and Higher Education Center (CATIE, Turrialba, Costa Rica) I learned how to integrate the complexity of landscapes into coherent approaches, including people, trees and crops. Lessons that were taken to a higher level when I joined the World Agroforestry (ICRAF, Costa Rica Office), Bioversity International (Costa Rica Office) and the Inland Norway University of Applied Sciences (INN, Hamar, Norway). During this time, I learned from an amazing group of scientists that gave me all the inputs that I used to develop this thesis. The thesis summarises my interests in plants and their interplay with people. That is why it assesses the different components of a landscape, first the people, then the trees and then the crops from where they take most of their food. I really enjoyed writing this thesis. All this thanks to the unbelievable opportunities that I had and all the nice and brilliant people that I met during this path. I wish that the children from the economically poor (but rich in biodiversity) regions in the tropics have access to the opportunities that I had.

Hamar, October 2, 2020
Kauê de Sousa

Contents

List of Figures

List of Tables

List of papers

Paper 1

K de Sousa, F Casanoves, J Sellare, A Ospina, JG Suchini, A Aguilar, L Mercado, L. (2018). "How climate awareness influences farmers' adaptation decisions in Central America?" *Journal of Rural Studies*, 64, 11–19. https://doi.org/10.1016/j.jrurstud.2018.09.018

Paper 2

K de Sousa, M van Zonneveld, M Holmgren, R Kindt, JC Ordoñez (2019). "The future of coffee and cocoa agroforestry in a warmer Mesoamerica". *Scientific Reports*, 9(1), 8828. https://doi.org/10.1038/s41598-019-45491-7

Paper 3

J van Etten, **K de Sousa**, A Aguilar, M Barrios, A Coto, M Dell'Acqua, C Fadda, Y Gebrehawaryat, J van de Gevel, A Gupta, AY Kiros, B Madriz, P Mathur, DK Mengistu, L Mercado, J Nurhisen Mohammed, A Paliwal, ME Pè, CF Quirós, JC Rosas, N Sharma, SS Singh, IS Solanki, J Steinke (2019). "Crop variety management for climate adaptation supported by citizen science". *Proceedings of the National Academy of Sciences*, 116(10), 4194–4199. https://doi.org/10.1073/pnas.1813720116

Paper 4

K de Sousa, J van Etten, J Poland, C Fadda, JL Jannink, YG Kidane, BF Lakew, DF Mengistu, ME Pè, SØ Solberg, M Dell'Acqua. (2020) "Data-driven decentralised breeding increases genetic gain in a challenging crop production environment". *Manuscript submitted.*

Paper 5

K de Sousa, AH Sparks, W Ashmall, J van Etten, SØ Solberg (2020). "chirps: API Client for the CHIRPS Precipitation Data in R". *The Journal of Open Source Software*, 5(51), 2419. https://doi.org/10.21105/joss.02419

Paper 6

K de Sousa, J van Etten, SØ Solberg (2020). "Climate variability indices for ecological and crop models in R: the `climatrends` package". *Preprint.*

Introduction

Climate change has increased the risks and uncertainties associated with agriculture (Campbell et al., 2016; Challinor et al., 2016; Steenwerth et al., 2014), which is one of the top sectors contributing to greenhouse gases accumulation (Poore & Nemecek, 2018), and one of the most vulnerable to its effects (Godfray et al., 2010; Tilman et al., 2011). Around the globe, changes in the frequency and intensity of extreme climatic events are expected (Aguilar et al., 2005; Imbach et al., 2018; Lin et al., 2017; Zohner et al., 2020), but its extent is hard to predict and will vary among regions. In the tropics, unpredictable precipitation and temperature oscillation associated with climate change has increased the concerns for farm adaptation, predominantly comprised by low-input smallholder systems that most likely will experience large yield gaps (Challinor et al., 2016, 2014; Peng et al., 2004; Zhao et al., 2017). On the other hand, an increase in temperatures, rainfall, and CO_2 levels may increase the productivity of agriculture in certain areas, most likely in some temperate zones (DaMatta et al., 2019; King et al., 2018; Tollenaar et al., 2017). This is however complicated, as increased pest pressure from weeds, insects and diseases, and highly seasonal climate variations are expected as well (Deutsch et al., 2018; King et al., 2018).

Agriculture, however, is a dynamic activity that has gone through a series of transformations over centuries. In its origins, human populations selected, domesticated and replicated the best plant genotypes and developed intercropping systems on which were the first steps of organised societies (Bellwood, 2004; Maezumi et al., 2018; Vavilov, 1992; Zeven, 1998). In the last century, in the attempt to provide food for a growing human population, agriculture evolved dramatically through the release of technological packages with high-yielding and high-tolerant crops that boosted agricultural production (Eshed & Lippman, 2019; Hickey et al., 2019). This revolution led to complex and diversified food systems that improved food access and economic growth throughout the world (Hickey et al., 2019; Pingali, 2012). However, this also produced trade-offs such as land degradation (Gibbs et al., 2010), the increase in the intake of high caloric and poorly nutritional foods (Pingali, 2012), and reduction in agricultural diversity through a greater focus on few crop species (Manners & van Etten, 2018). Also, the question on how to feed a growing human population still persists (FAO et al., 2019), despite the great contributions achieved in the agricultural sector (Pingali, 2012). Climate change magnifies the challenge. This calls for a new revolution and transformation of our food systems (Loboguerrero et al., 2020), at the farm level (Nelson, 2020; Sanchez, 2020), at the crop science level (Manners & van Etten, 2018), and also on how food is distributed and what we eat (Hirvonen et al., 2020; Willett et al., 2019).

It is argued that adoption of sustainable intensification and agroecological practices to meet the future challenges will help farmers to cope up with the effects of climate change (Andrade et al., 2020; Campbell et al., 2016; Lipper et al., 2014; Nelson, 2020; Steenwerth et al., 2014; Thornton et al., 2017; Zimmerer & De Haan, 2017). This transformation would involve a number of integrative approaches to help farmers to reorient their agricultural practices to ensure a rise in productivity, farm income and food security while adapting and mitigating to climate change. These practices include a range of actions, such as diversification, agroforestry, varietal selection, crop breeding, and landscape planning. Diversification, with the use of crop variety management or agroforestry, is one of the most promoted approaches (van Zonneveld et al., 2020). However, this approach needs to be explored to identify local solutions that embrace the complexity of the cropping system with its ecological interactions and abiotic factors.

For diversification with varietal selection, one proposed solution is to increase variety supply by accelerating crop breeding for more robust varieties and at the same time replace old varieties from the seed system (Atlin et al., 2017). Supply-driven variety replacement assumes that new varieties are well adapted to local environments and acceptable to farmers. For many climate-vulnerable crops and regions, this assumption is unwarranted. Varieties are often recommended without geographic analysis of climate adaptation to determine recommendation domains (Annicchiarico, 2002). A solution could come from a more scalable type of participatory research that *crowdsources* farmers' local knowledge through citizen science (Beza et al., 2017; Steinke et al., 2017; van Etten, Beza, et al., 2019).

In terms of agroforestry, adaptation is addressed by introducing trees to the agricultural production system to ameliorate abiotic stress and facilitate the performance of understory crops (Blaser et al., 2018; Holmgren et al., 1997). This approach is particularly challenging, as perennial crops take long before farmers fully benefit from their management decisions (Cerda et al., 2014; de Sousa et al., 2016; Ramírez et al., 2001). Two issues also increase the uncertainty of agroforestry to address climate adaptation and risk management. First, climate change may also affect the habitat for agroforestry tree species (Lyra et al., 2017). Second, interactions between trees and crops could drive negative results in productivity (Abdulai et al., 2017; Blaser et al., 2018), limiting the interest of farmers in adopting this approach. Understand the effects of climate change on the suitability of agroforestry species may be one approach to support farmers in selecting agroforestry designs at low risk.

The increasing intensity of extreme climatic events associated with climate change and the absence of knowledge to provide recommendation domains highlights the need for rapid, tailored and straightforward solutions to help farmers to adapt agricultural production. In this thesis, I analysed diversification approaches to bring new insights to generate accurate climate-friendly recommendations for farmers across different climatic spaces. I focussed on smallholder farming, a practice that is conducted in more than 570 million farms worldwide (Lowder et al., 2016), contribute to around 30-34% of global food supply (Ricciardi et al., 2018), is central for the conservation of agrobiodiversity (Altieri & Nicholls, 2017), but also highly vulnerable to climate change (Harvey et al., 2014; Lipper et al., 2014; Vermeulen et al., 2012). Here I bring evidence for different levels of biological levels: species, focusing on trees as slow grower organisms for interspecific diversification; varieties, looking for locally adapted phenotypes; and genotypes (genes) focusing on $G \times E$ interactions to support crop breeding for intraspecific diversification.

Objectives

This thesis aimed to provide insight and develop methods to support farmers and stakeholders in adapting agricultural systems to a changing climate, specifically for smallholders, by answering the following questions:

1. Can climate awareness influence sustainable adaptation decisions in smallholder farms? (Paper 1)

Climate change increases the risks and uncertainties associated with agriculture, particularly for smallholders (Altieri & Nicholls, 2017; Campbell et al., 2016). The evidence has shown that the adoption of agricultural innovations and climate-adapted practices can help vulnerable farmers to cope with the effects of climate variability and change (Lipper et al., 2014; Vermeulen et al., 2013, 2012). These practices include farm sustainable intensification, diversification of production, agroforestry, crop variety management and plant breeding. Farmers' awareness and perceptions of climate change are correlated with the adoption of such innovations (Elum et al., 2017; Niles & Mueller, 2016; Schattman et al., 2016; Singh et al., 2017), but no evidence is provided for smallholders in Central America. Paper 1 (de Sousa, Casanoves, et al., 2018) targets this knowledge gap.

2. What is the future of current agroforestry combinations in coffee and cocoa production systems? (Paper 2)

Agroforestry, the deliberate and simultaneous management of trees within crop or livestock systems (Nair, 1993), is considered an important climate-adapted innovation to increase the resilience of agricultural systems (Spurgeon, 1979). Trees can ameliorate the micro-climate and facilitate the performance of understory crops (Holmgren & Scheffer, 2010). Most perennial crop systems in Central America are managed following agroforestry practices (Beer et al., 1998; Somarriba et al., 2013), and have been increasingly encouraged as climate change is projected to affect future crop production (Ovalle-Rivera et al., 2015). Nevertheless, climate change can also affect the future ecological niches of several tree species (Holmgren et al., 2013; Lyra et al., 2017) and may hamper the prospects of agroforestry as a viable approach for climate adaptation. Paper 2 (de Sousa et al., 2019) assesses the future of the 100 most common tree species found in Arabica coffee (*Coffea arabica* L.) and cocoa (*Theobroma cacao* L.) production systems in Central America.

3. Can cocoa become a suitable alternative in vulnerable coffee production areas? (Paper 2)

As climate change projections points to a decline in coffee production due to the increasingly climate variability (Bunn et al., 2015; Läderach et al., 2017; Ovalle-Rivera et al., 2015), farmers have developed an interest for cocoa. The drivers of this shift are trends in recent years of increasing coffee production costs and large losses due to pests and diseases (*e.g.* the leaf rust crisis caused by climate oscillation that makes coffee susceptible to the fungus *Hemileia vastatrix* Berk. & Broome) (Avelino et al., 2015). Replacing coffee by cocoa has become one of the main strategies for climate change adaptation for producers in low elevation areas (Läderach et al., 2017). Nevertheless, there is no quantitative assessment of the feasibility of such strategy, starting from considering that cocoa is vulnerable to climate change itself (Schroth et al., 2016). Paper 2 (de Sousa et al., 2019) also explores this strategy assessing the potential areas where cocoa is a suitable alternative to coffee.

4. Can on-farm participatory crop trials generate insights into variety management for climate adaptation? (Paper 3)

Crop improvement increases production and contributes to food and nutrition security (Godfray et al., 2010; Hickey et al., 2019). Previous studies have shown that it is critical that farmers keep a continuous turnover of improved and locally adapted varieties for climate adaptation (Atlin et al., 2017; Challinor et al., 2016). One constraint, however, in adopting this practice is the cost of the seeds. Farmers often rely on their local varieties and changing to new ones can increase the risks when the performance of these varieties under local conditions is unknown (Dawson et al., 2008). Existing experimental agricultural approaches lack the ability to provide such recommendations across space and time, particularly to marginal production environments. Paper 3 (van Etten, de Sousa, et al., 2019) address this issue by exploring a participatory approach to characterise varietal climatic responses allowing for seasonal and geographical extrapolation.

5. Can a data-driven decentralised approach improve the selection of genotypes in challenging crop production environments? (Paper 4)

To adapt to climate change farmers require accelerated selection of genotypes and production of locally adapted varieties (Eshed & Lippman, 2019; Godfray et al., 2010). Conventional breeding programs have proven high success in maximizing genetic diversity in the early stages of selection and then identifying superior germplasm (Hickey et al., 2019). At present, plant breeders use genomic-driven approaches to increase selection intensity while reducing the time of the breeding cycle and deriving greater genetic gain. However, the same approach may not translate well in marginal environments, often in the periphery of research stations and characterised by a diversity of environments and management practices (Annicchiarico et al., 2019). Decentralised participatory approaches could help breeders in accelerate the selection of genotypes while addressing the $G \times E \times M$ interactions (genotype by environment by management) required for challenging crop production environments (Annicchiarico et al., 2019; Ceccarelli & Grando, 2019; Tester & Langridge, 2010; van Eeuwijk et al., 2001; van Etten, Beza, et al., 2019). Paper 5 (de Sousa, van Etten, Poland, et al., 2020a) addresses this issue by proposing a decentralised data-driven approach scaled by citizen science.

6. Can we develop tools to allow reproducible and replicable workflows for crop recommendation domains? (Paper 5 and Paper 6)

Reproducibility, the ability to repeat the analysis, and replicability, the ability to repeat an experiment (Stevens, 2017), are key to perform collaborative scientific research (Munafò et al., 2017; Powers & Hampton, 2019). It allows scientists to re-perform analysis after a long hiatus and the peers to validate analysis and get new insights using original or new data. In the walk of climate change it is key to ensure that recommendation domains are made based on replicable and reproducible workflows that can be updated as new data becomes available. To address this issue we developed a series of tools to create a workflow for the analysis on crop variety management using the R language (R Core Team, 2020). Paper 5 and 6 illustrates the applicability of some of these tools.

Methods

Research sites

The studies took place in three different regions, Central America, East Africa and South Asia (Fig. 1). The regions are characterised by its rich plant diversity being *Centre of Origin* and domestication (Vavilov, 1992) of important staples and crops such as common bean (*Phaseolus vulgaris* L.), maize (*Zea mays* L.) and cocoa in Central America; durum wheat (*Triticum durum* Desf.) and coffee in East Africa; and rice (*Oryza sativa* L.), and bread wheat (*Triticum aestivum* L.) in South Asia. Smallholder agriculture and livestock production are the main livelihood for the majority of the population in these regions. Poverty and food insecurity levels are still among the higher in the world. According to the 2019's report on the State of Food Security and Nutrition (FAO et al., 2019), in 2018 Central America, East Africa and South Asia had a prevalence of severe or moderate food insecurity for 31.5%, 62.7% and 34.3%, respectively, of their total population.

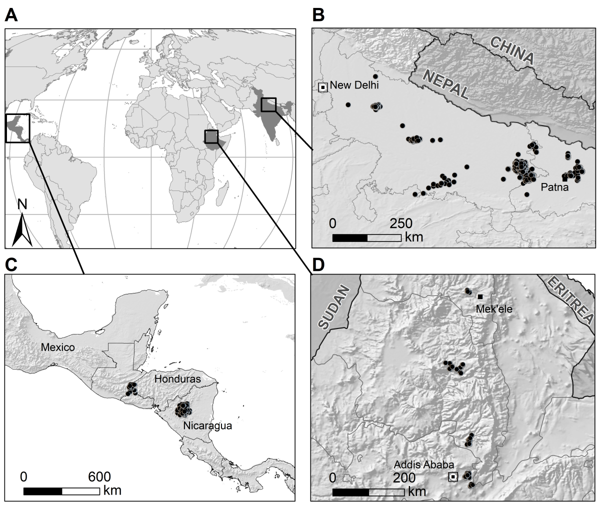

Figure 1: Research sites. (A) Overview. (B) India. (C) Central America. (D) Ethiopia. Farms included in the trials or interviews are indicated as dots. Research in Paper 2 was performed across the framed area in panel C.

Field data was collected between 2010 and 2016 as part of several Research for Development programs performed across the regions. In Central America, farmer's surveys were conducted in El Salvador, Guatemala, Honduras and Nicaragua (Paper 1). A subset of participatory crop trials was conducted in Nicaragua (Paper 3). Sampled farmers extended across three ecological regions, the Central American Atlantic Moist Forests, the Central American Dry Forests and the Central American Pine-Oak Forests. The research in Paper 2 was conducted across the entire area shown in Fig. 1C. Important crops grown by smallholders in the region are maize, common bean, sorghum (*Sorghum bicolor* (L.) Moench), banana (*Musa* spp.), coffee and cocoa, the last two grown for the international markets, while the others for local markets and household consumption.

In East Africa, the research (Paper 3 and 4) was conducted in the regions of Amhara, Oromia and Tigray in Ethiopia, which encompasses one main ecological region, the Ethiopian Montane Grasslands and Woodlands. The main crop grown by smallholder farmers in this region are durum wheat, teff (*Eragrostis tef* (Zucc.) Trotter), barley (*Hordeum vulgare* L.) and sorghum, mostly for household consumption and local markets. In South Asia, the research (Paper 3) was conducted across the States of Bihar, Madhya Pradesh and Uttar Pradesh in India encompassing three ecological regions, the Upper Gangetic Plains Moist Deciduous Forests, the Lower Gangetic Plains Moist Deciduous Forests, and the Narmada Valley Dry Deciduous Forests. The main crop grown by smallholders are rice, bread wheat, maize, and several pulses and vegetables. Table 1 presents a description on the environmental characteristics of each region extracted from the sampled locations used in this research (Hijmans et al., 2005; Jarvis et al., 2008; Olson et al., 2001).

Table 1: Environmental characteristics of sampled locations across the three research sites. Average values with minimum and maximum shown between parenthesis.

Research site	Elevation (m)	Temperature (°C)	Precipitation (mm)	Ecoregions
Central America	597 (190–1,900)	23 (17–29)	1,717 (905–2,122)	3
East Africa	2,598 (1,960–3,200)	15 (6–25)	976 (671–1,078)	1
South Asia	85 (42–571)	25 (9–39)	1,024 (808–1,280)	3

Farmers' climate awareness and adaptation strategies

This part of the research was performed only in Central America (Paper 1). In 2014, we performed a survey to 283 households participating in the Mesoamerican Environmental Program (MAP) (Gutierrez-Montes et al., 2020). Farmers were questioned about their perceptions regarding changes in precipitation and temperature over the 10 years before the interviews (2005–2014). Farmers who reported to have felt changes in climatic patterns were asked to list the farm management practices they have adopted in their crop systems to cope with such changes. These practices were ranked by the order they were mentioned. We wanted to answer two main questions, how accurate are the farmers' perceptions to climate change with observed time series data, and whether socio-economic factors can influence farmers' adaptation decisions.

To address the first question, we linked the farmers' responses as categorical variables (*e.g.* more precipitation, less precipitation, uncertain rain season) to a gridded time series precipitation database from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) dataset (Funk et al., 2015). This dataset incorporates global daily rainfall from 1983 to near-present with a resolution of 2.5 arc-min, which is obtained by weather stations and combined with remote sensing. Changes in precipitation were assessed by calculating three extreme precipitation indices using an alpha version of the R (R Core Team, 2020) package `climatrends` (de Sousa, van Etten, & Solberg, 2020). For this analysis we used the simple daily intensity index (SDII, total precipitation/rainy days), the maximum 5-day precipitation (Rx5day), and the maximum length of consecutive dry days (MLDS). We linked the observed changes in precipitation and farmers' perceptions with a multiple correspondence analysis using the R package `FactoMineR` (Lê et al., 2008). The analysis takes multiple categorical variables and seeks to identify associations between levels of those variables. The associations were visualised with a biplot.

The second question was answered by linking the farmers prioritised adaptation management strategies with their socio-economic data. Household socio-economic data was obtained by the baseline survey performed with all farmers for the Mesoamerican Environmental Program. From the adaptation strategies derived from all responses, we compiled a list of 10 options: (*i*) Change in Agricultural Calendar, (*ii*) Change in Varieties, (*iii*) Production Diversification, (*iv*) Introduction of New Crops, (*v*) Less Fertilizers and Pesticides, (*vi*) Reforestation and Restoration, (*vii*) Sustainable Soil Management, (*viii*) Sustainable Water Management, (*ix*) Leave Farming System, and (*x*) More Fertilizers and Pesticides.

The relative importance of the different strategies was measured following the Luce's Choice Axiom (Luce, 1959), which states that the probability that one item beats another is independent from the presence or absence of any other items in the set.

Equation [1]

$$P(i \succ j) = \frac{p_i}{p_i + p_j}$$

where p_i is a positive real-valued score assigned to individual i. The comparison $i \succ j$ can be read as "i is preferred over j"

We estimated the *worth* parameters (relative importance) using the maximum log-likelihood estimation with the Bradley-Terry (BT) model (Bradley & Terry, 1952). In the BT model, each m individuals choice is compared with one another in pairs to compute the number of times i is preferred over j. The outcomes of different pairings are assumed to be independent, and the log-likelihood based on the BT model is estimated using the equation (Hunter, 2004):

Equation [2]

$$\ell(p) = \sum_{i=1}^{m} \sum_{j=1}^{m} [wij \;\ln\; p_i - wij \;\ln(p_i + p_j)]$$

where w_{ij} denotes the number of times individual i has beaten individual j and we assumes $w_{ii} = 0$ by convention.

We used the top five strategies mentioned by each interviewed farmer and converted it into pairwise comparisons using an alpha version of the R package `gosset` (de Sousa, van Etten, Dumble, et al., 2020). Socio-economic variables were linked to the pairwise rankings using the

Model-Based Recursive Partitioning approach (Zeileis et al., 2008) a non-parametric modelling approach implemented in R by the package `psychotree` (Strobl et al., 2011). The algorithm starts by fitting a BT model to the full data, then it assesses the stability of the *worth* parameters, if there is a significant instability, the full data is split by the covariate with strong instability. The process is repeated until there is no more significant instability (Zeileis et al., 2008). We linked six covariates: (*i*) the ecoregion (Dry or Rainforest), (*ii*) the Progress Out of Poverty Index (PPI), (*iii*) the literacy level of the head of household, (*iv*) the area of the main crop system (ha), (*v*) the age of the head of household, and (*vi*) the number of practices adopted by the farmers after participating in the Farmers Field Schools led by MAP.

Mapping future suitability of coffee and cocoa agroforestry

This assessment focussed on Central America within the coordinates 101° to 77° E and 7° to 22° N (Paper 2). We wanted to answer two main questions, how ecologically resilient are the trees commonly used by farmers in coffee and cocoa agroforestry systems, and whether cocoa can be a suitable alternative for coffee growers. These questions came out as a result from the assessment on farmers' adaptation decisions described in the earlier section. Among the 10 adaptation decisions mentioned by the interviewed farmers, restoration and reforestation (using agroforestry) were the most preferred (de Sousa, Casanoves, et al., 2018).

To answer the first question, we assessed the current and future potential distribution of the top-100 commonly used tree species in cocoa and coffee plantations across Central America. The selection of the 100 species was based on three criteria. The first was the abundance reported in relevant datasets of agroforestry inventories conducted in smallholder farms across the region (Bonilla Zunhiga et al., 2014; Orozco et al., 2014; Sepulveda & Barrios, 2016), selecting from the most abundant to the least. We then filtered the list of species based on ecological and economic services identified by farmers and reported in the literature (CATIE & OFI, 2003), taxa were classified by their main use; N-fixing (for soil amelioration), timber or fruit. The last criterion was the availability of at least 60 geographical records to ensure accurate modelling results. To answer the second, question we compared the current potential areas for coffee and cocoa production with their projected distribution under climate change scenarios.

We compiled presence location points of selected tree species (including coffee and cocoa) from the Global Biodiversity Information Facility (GBIF) (GBIF.org, 2020), MAPFORGEN (http://www.mapforgen.org) and from the database of farm inventories used to select the tree species. No distinction was made between locations from natural forests or farms because this information was not always available in the original sources.

Bioclimatic predictors from WorldClim v1.4 (Hijmans et al., 2005) were used to model the current distribution of the 100 species, and of coffee and cocoa. These bioclimatic variables are widely used in ecology to model the distribution of species based on their interaction with the variation in precipitation and temperature (Booth, 2018). To avoid model overfitting, we selected the least correlated variables based the variance inflation factors, retaining those with VIF $<$ 10 (Ranjitkar et al., 2014). This resulted in nine bioclimatic variables. Future projections were based on two Representative Concentration Pathways (RCP) scenarios of climate change (van Vuuren et al., 2011), RCP 4.5 as an intermediate scenario which predicts an average temperature increase of 1.4 °C (0.9–2.0 °C), and RCP 8.5 as a high emissions scenario, which predicts an average temperature increase of 2.0 °C (1.4–2.6 °C) by 2050. These scenarios are defined by the value of radiative forcing (ability to absorb or release heat) from the atmosphere to 2050, ranging from 2.6 to 8.5 Watts $\cdot$ m^{-2}. The scenario RCP 2.6 was not chosen because it represents the most effective mitigation scenario, aiming to keep temperature below 2 °C. Currently this scenario is unlikely with projections of current policies (expected temperature increase of from 3.3 °C to 3.9 °C). To cover the uncertainty of the General Circulation Models, we used future bioclimatic variables downscaled from 17 General Circulation Models that were available for both RCP scenarios.

The distribution of the 100 species and coffee and cocoa was modelled using an ensemble suitability method implemented by the R package `BiodiversityR` (Kindt, 2018). The procedure consists of four steps that, first, calibrate the model by assessing the performance of 18 algorithms of species distribution models (SDM) measured with the area under the curve (AUC). In this step, the AUC values obtained by each algorithm are weighed using the following equation:

Equation [3]

$$S_e = \frac{\sum_i w_i S_i}{\sum_i w_i}$$

where the ensemble suitability (S_e) is obtained as a weighted (w) average of suitabilities predicted by the contributing algorithm (S_i).

The second step consists in retaining only the algorithms that contributed at least in 5% to the ensemble suitability (S_e). The third step generates the suitability maps using the predictions from the algorithms that were selected in the second step. Finally, to generate the presence–absence layers, we convert the consensus suitability from the third step using the threshold of maximum specificity + maximum sensitivity (Liu et al., 2013). Replication data and code used in this analysis are available through Dataverse (de Sousa, van Zonneveld, et al., 2018).

Evaluation of crop varieties

This part of the research (Paper 3) was performed between 2012 and 2016 during three growing seasons in Ethiopia, five growing seasons in Nicaragua, and four growing seasons in India (Table 2). Three crops were evaluated, common bean in Nicaragua, durum wheat in Ethiopia and bread wheat in India. The question that we addressed was whether on-farm participatory crop trials, scaled through a citizen science approach, can provide robust, actionable information on varietal climate adaptation. This aimed to respond to one open question in the assessment of farmers' adaptation decisions where change in crop varieties (or crop variety management) showed to be one of the least choices in adaptation decisions among the farmers.

We compiled data from 12,409 farmer-managed plots across the research sites. Trial design followed the *tricot* approach, standing for triadic comparison of technologies (van Etten, Beza, et al., 2019). The approach follows five principles: (*i*) anonymous subsets of three varieties (out of a larger set) are allocated randomly as incomplete blocks (Atlin et al., 2001); (*ii*) participants receive one subset to grow in their farms under their own management practices (Figure 2); (*iii*) plots are set up within the crop system, plots are small to facilitate participation but large enough to avoid strong edge effects; (*iv*) participants indicate the relative performance of varieties through ranking answering to two short statements for each targeted characteristic (*e.g.* which variety had the best leaf development? which variety had worst leaf development?); (*v*) data from each farmer-managed plot is collated into a single dataset. Across the research sites the pool of varieties comprised a list of 11 varieties in Nicaragua, 62 varieties and genotypes in Ethiopia, and 21 varieties in India.

Table 2: Number of tricot trials per cropping season of durum wheat (Ethiopia, Meher season), bread wheat (India, Rabi season), and common bean (Nicaragua).

			Nicaragua		
Year	Ethiopia	India	Primera	Apante	Postrera
2012	–	562	–	–	–
2013	176	4,134	–	–	–
2014	578	4,947	–	–	–
2015	336	834	–	481	177
2016	–	–	64	87	33

For the analysis of the ranking data generated by farmers, we used the Plackett–Luce (PL) model (Luce, 1959; Plackett, 1975), implemented in R with the package `PlackettLuce` (Turner et al., 2020). The PL model is similar to the BT model above and follows the Luce's Choice Axiom (Eq. 1). Whereas the BT model is used for pairwise comparisons, the PL model is used for rankings of three or more items. This makes possible to compare items across the entire rank permutation whereas BT model breaks the comparison into pairs. The PL model determines the values of positive-valued parameters α_i (*worth*) associated with each item i. These parameters α are related to the probability (P) that item i wins against all other n items. We report *worth* values that sum to one. This makes each *worth* value α_i equal to the probability of item i outperforming all other items:

Equation [4]

$$P(i \succ \{j, ..., n\}) = \frac{a_i}{a_1 + ... + a_n} = \frac{a_i}{1} = a_i$$

In the trials, we used rankings of three varieties ($i \succ j \succ k$), which have the following probability of occurring according to the PL model:

Equation [5]

$$P(i \succ j \succ k) = P(i \succ j, k) \cdot P(j \succ k)$$

The likelihood for a ranking ($i \succ j \succ k$) follows from Eqs. 1, 4, and 5 and takes the following form:

Equation [6]

$$
\begin{aligned}
\ell(\boldsymbol{\alpha}) &= \ln(P(i \succ \{j,k\})) + \ln(P(j \succ k)) \\
&= \ln(\alpha_i) - \ln(\alpha_i + \alpha_j + \alpha_k) + \ln(\alpha_j) - \ln(\alpha_j + \alpha_k)
\end{aligned}
$$

The likelihood is then the sum of the log-likelihood $\ell(\boldsymbol{\alpha})$ values across all rankings. Using an iterative algorithm, the log-likelihood is maximised to identify the α values that make the observed rankings most probable.

Climatic variables were linked to the rankings using the Model-Based Recursive Partitioning approach (Zeileis et al., 2008) which builds the partitioning trees. This process is explained in the previous section on farmers' adaptation decisions. For the climatic variables, we used free publicly available datasets with coverage across all the research sites to make comparable studies. We derived rainfall and temperature indices using an alpha version of the R package `climatrends` (de Sousa, van Etten, & Solberg, 2020). Rainfall was obtained using the CHIRPS dataset (Funk et al., 2015), while temperature was obtained from MODIS MYD11A2 (Wan et al., 2015).

Fourteen climatic variables were extracted for the vegetative, reproductive and grain filling period and the whole growth period (from planting date to harvesting) in each observation point. This resulted in 110 variables. To create models that provide generalizable predictions across seasons, we used blocked cross-validation (with seasons as blocks) combined with a forward variable selection procedure (Meyer et al., 2018). We used the deviance values of each validation season to calculate an Akaike weight, which is the probability that a given variable combination represents the best model (Wagenmakers & Farrell, 2004). We performed forward variable selection, using this combined Akaike weight as our selection criterion. From each study case (country) this procedure retained one variable, which were the maximum night temperature (°C) during the vegetative and reproductive periods for common bean in Nicaragua, the minimum night temperature (°C) during the vegetative period for durum wheat in Ethiopia, and the diurnal temperature range (°C) during the vegetative period for bread wheat in India.

We compared the goodness-of-fit of the model with climatic variables (climate model) against three other models. The first with no covariates (intercept-only model), the second with geolocation, season, planting dates, and soil categories, which represented the experimental design (design model). And the third model with a combination of climatic variables plus geolocation (climate + geolocation model). To compare the models, we calculated a weighted average of pseudo-R^2 (deviance reduction) values across testing seasons (Agresti & Kateri, 2011), using the square root of the sample size as weights (Whitlock, 2005). All this process was done using an alpha version of the R package `gosset` (de Sousa, van Etten, Dumble, et al., 2020). Replication data are available through Dataverse (van Etten et al., 2018)

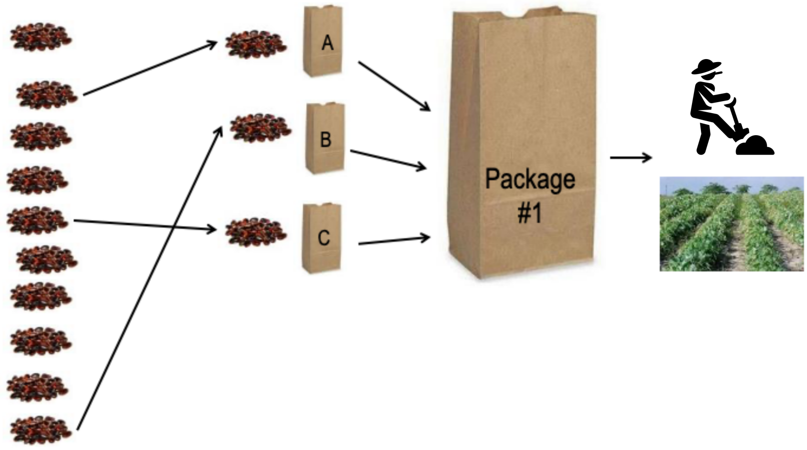

Figure 2: Randomisation and subset allocation in the tricot approach. Three varieties are randomly selected from a larger group and anonymised with the labels A, B, C, participants receive the anonymous subset to evaluate in their farms under their own management practices.

Decentralised genotype selection

We then focussed on the crop variety management approach for decentralised breeding (Paper 4). This research was performed in Ethiopia from 2012 to 2015 with durum wheat. The main question that we addressed was whether a decentralised approach could improve the selection of genotypes for crop breeding targeting challenging crop production environments. We compared a *data-driven decentralised breeding* approach, or 3D-breeding, for short, with a benchmark representing a centralised approach used in current breeding programs (Figure 3). A total of 400 durum wheat genotypes were selected from a representative collection of accessions from the Ethiopian Biodiversity Institute.

Centralised trials were performed in 2012 and 2013 in the districts of Geregera (Amhara) and Hagreselam (Tigray). In 2012, thirty experienced smallholder farmers (15 men and 15 women) were invited to participate in the trial evaluations at the station plots, held concurrently after flowering stage. The farmers had no previous knowledge of the genotypes included in this study to prevent bias in the evaluations. The participants provided appraisal with Likert scales (1 to 5 worse to best) (Likert, 1932) given to genotypes for overall appreciation (OA) (Kidane et al., 2017; Mancini et al., 2017). Research technicians measured grain yield (GY) as grams of grain produced per plot, then converted into $t \cdot ha^{-1}$. Absolute values of GY and OA measured in centralised trials were converted into ordinal rankings.

A total of 1,165 decentralised plots were performed between 2013 and 2015 during three cropping seasons across the regions of Amhara (471), Oromia (399) and Tigray (295) using a subset of the 41 best genotypes identified through farmer evaluation in centralised trials (Mancini et al., 2017). Season 1 (2013) comprised 179 fields, Season 2 (2014) comprised 651 fields, and Season 3 (2015) comprised 335 field. Trial design followed the *tricot* approach as described in the previous section. Farmers reported the overall appreciation and research technicians collected GY measures in farmers' plots after harvesting. The comparison 3D-breeding vs benchmark was done using the subset of 41 genotypes used in both trials.

Genomic DNA was extracted from fresh leaves pooled from five seedlings for each of the accessions in the centralised trials with the GenEluteTM Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, St Louis, USA) following manufacturer's instructions in the Molecular and Biotechnology Laboratory at Mekelle University, Ethiopia. Genotyping was performed on the Infinium 90k wheat chip at TraitGenetics GmbH (Gatersleben, Germany). Single nucleotide polymorphisms (SNPs) were called using the tetraploid wheat pipeline in GenomeStudio V11 (Illumina, Inc., San Diego, CA, USA). Full details on the genotyping are given by Mengistu et al. (2016).

We derived best linear unbiased prediction (BLUP) values from GY and OA measured in centralised trials. The benchmark representing a centralised breeding system was conducted using genomic selection models and marker-based genetic relationship matrices computed on BLUP data. To measure accuracy of genomic selection predictions, we calculated the Kendall's tau coefficient (τ), a measure of similarity of rankings (Kendall, 1938), between predicted values and observed values.

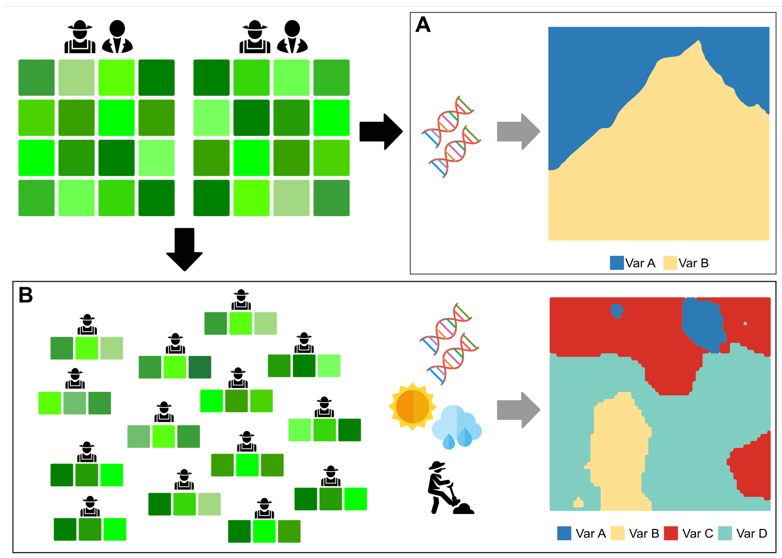

Figure 3: Centralised breeding (A) derives recommendations from breeders' evaluation and possibly participatory assessments in a limited set of stations, using genomics to accelerate the production of varieties that are eventually recommended with coarse spatial resolution. This system may become more efficient if complemented by 3D-breeding (B), a decentralised approach where the best candidate genotypes are tested by farmers in small, blinded and randomized sets. 3D-breeding produces scalable solutions that can be linked to genomics, farmers' knowledge and environmental data, to enhance the local adaptation of the resulting varieties and tailor their recommendation to the landscape.

The statistical model that represented 3D-breeding was developed using the data generated by the citizen science decentralised trials using the Plackett-Luce model. DNA data from SNPs was added into the model as a prior using an additive matrix. To take into account explanatory variables, we created Plackett-Luce trees through Model-Based Recursive Partitioning. Daily temperature and precipitation data was obtained from the NASA Langley Research Center POWER Project funded through the NASA Earth Science Directorate Applied Science Program (https://power.larc.nasa.gov/), using the R package `nasapower` (Sparks, 2018). Climatic variables were obtained using the R package `climatrends` (de Sousa, van Etten, & Solberg, 2020). We selected the most relevant climatic variable using the process described in the earlier section. Which retained the maximum night temperature (°C) during reproductive growth and the minimum night temperature (°C) during the vegetative growth. Replication data and code used in this analysis are available through Dataverse (de Sousa, van Etten, Poland, et al., 2020b).

Results and discussion

Climate awareness and farmers' adaptation decisions

We assessed whether climate awareness could led to sustainable adaptation decisions in smallholder farms in Central America. The 255 interviewed farmers (out of 283) reported to perceived changes in climate patterns over the 10 years prior to the survey (2005–2014). The multiple correspondence analysis of farmers' perceptions versus anomalies from observed data shows partial correlations between farmers' perceptions and the time series data (Figure 4). Farmers who perceived uncertainties in the start/end of the rainy season correlate with observed decrease in heavy precipitation (Rx5day) and increase in the duration of consecutive dry days (MLDS). However, farmers who perceived less annual precipitation correlate with observed increase in heavy precipitation (as result of less distributed rain across the season but concentrated in a short period), while those who perceived more precipitation or heavy precipitation did not correlate with any of the observed precipitation indices.

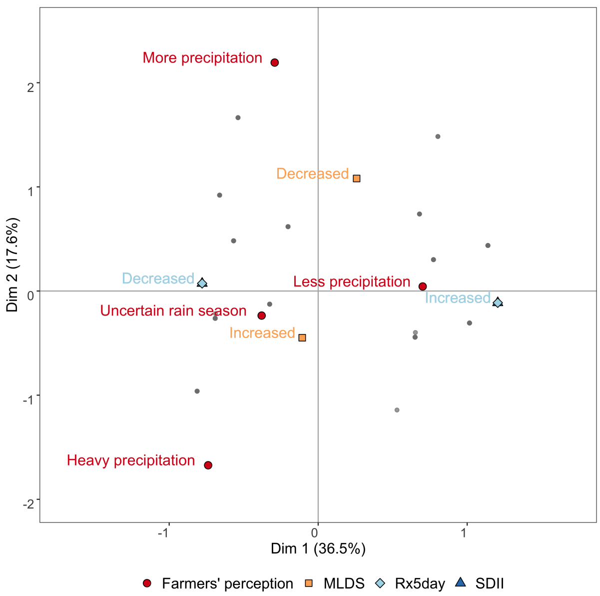

Figure 4: Correlation between farmers' perception on changes in precipitation and observed anomalies in precipitation indices over 2005–2014 in the sampled locations across Central America. MLDS, maximum length of consecutive dry days (< 1 mm); Rx5day, maximum 5-day precipitation (mm); SDII, simple annual precipitation index (mm/day).

The partial correlations in observed and perceived climate may be explained by the difficulty to properly observe the changes as they occur, without the aid of measuring devices (*e.g.* weather station, garden moisture meter), or access to climate services (Bouroncle et al., 2019). Or the result of cognitive bias (Holmgren et al., 2018). However, even if farmers do not perfectly perceive these changes, they do observe reductions in yields, and at times total losses, which draws their attention to climate-related issues and increases their willingness to innovate and try new farm management practices. Previous studies reported this behaviour as a product of experiencing climatic risks or disasters (Bergquist et al., 2019; van Valkengoed & Steg, 2019).

In the case of interviewed farmers in Central America, there was a list of 10 practices that were adapted upon the perception of changes in climate patterns. The *worth* parameters for the adaptation practices show significant differences between the ranked options (Table 3). Practices of *Reforestation and Restoration*, *Introduction of New Crops*, and *Sustainable Soil Management* were reported as the most preferred choices among interviewed farmers, showing higher *worth* than the reference *Production Diversification*, which was selected as reference because if the most recommended adaptation strategy for farmers (Atlin et al., 2017; van Zonneveld et al., 2020). *Change in varieties*, had a lower *worth* than the reference. The other practices were ranked below the reference, with *Leave Farming System* and *Change Agricultural Calendar* on the bottom of preferred practices to adapt to perceived changes in climatic patterns.

Table 3: Bradley-Terry model estimates from farmers' management practices employed to adapt crop systems to perceived changes in climate patterns in Central America.

Adaptation decision	Worth	Std. Error	Pr(>|z|)	Signif.
Reforestation and Restoration	1.5120	0.0811	< 0.0001	***
Introduction of new crops	0.7572	0.0844	< 0.0001	***
Sustainable soil management	0.2554	0.0834	0.0022	***
Production diversification	**0**	–	–	–
Change in varieties	-0.2805	0.0883	0.0015	**
Sustainable water management	-0.6814	0.0919	< 0.0001	***
Use of more fertilizers and pesticides	-0.7658	0.0925	< 0.0001	***
Use of less fertilizers and pesticides	-0.8516	0.0942	< 0.0001	***
Leave farming system	-1.4053	0.1069	< 0.0001	***
Change in agricultural calendar	-1.5276	0.1095	< 0.0001	***

Significance levels: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1.

We found significant differences in preferences for adaptation practices based on formal education level, farm size and ecological region, with the BT model identifying four different sub-groups. Overall, reforestation was the preferred choice among farmers independent of socio-economic profiles. This practice had an interplay with agroforestry, where farmers intentionally planted trees or allowed then to grow through natural regeneration within their crop systems. Several studies have reported agroforestry among the bests climate change adaptation strategies (Blaser et al., 2018; Lipper et al., 2014; Mbow, van Noordwijk, et al., 2014; Mbow, Smith, et al., 2014; Verchot et al., 2007), it includes both mitigation and adaptation by providing carbon sink, microclimate regulation and protection to extreme climate events (Caudill et al., 2015; Holmgren & Scheffer, 2010; Torres et al., 2017). Farmers, however, have a clear preference to few marketable species or that has a clear utility within the crop system (Cerdán et al., 2012; de Sousa et al., 2016; Ordoñez et al., 2014). Some studies have shown that climate change can also affect the suitability of tree species (de Sousa et al., 2017; Lyra et al., 2017) and hinder the benefits of agroforestry as an adaptation strategy.

Change in varieties was the least choice in the top-5 options reported by interviewed farmers in Central America. The low uptake of this approach reflects the weakness of formal and informal seed systems in supporting smallholders to select the right variety (McGuire & Sperling, 2016), mainly the informal seed systems, as it is expected to provide higher contribution for local adaptation (Bellon et al., 2011). Costs and the high risk associated with crop variety managements by smallholder farmers are also among the reasons for the low uptake. Several studies in Africa and Asia investigated the reason for uptake (or non-uptake) of agricultural innovations among smallholder farmers (Elum et al., 2017; Meijer et al., 2015; Senyolo et al., 2018; Singh et al., 2017), the lack of knowledge (Meijer et al., 2015) and high costs (Senyolo et al., 2018), among others, were also pointed out as main reasons for non-uptake of innovations, such as crop variety management. Smallholder farming also have an intrinsic characteristic of being performed in diverse low-input systems (Lowder et al., 2016), which makes the challenge of recommending (or producing new) crop varieties more problematic and riskier.

Cost is a factor that involves multiple factors beyond farmers' control (Chapagain et al., 2020). However, risks in selecting crop varieties could be reduced with tailored advice. Recent experiences in Central America, Africa and Asia provided new evidence that this challenge can be addressed by scaling agricultural experimentation with citizen science (Beza et al., 2017; Steinke et al., 2017; van Etten, Beza, et al., 2019). These studies showed positive prospects that citizen science could support seed systems in tracking the responses of crop systems to the changing climate patterns as they occur in the farms and help farmers in taking the best decisions towards climate adaptation.

Overall, when facing changes in climate, farmers adopt a set of sustainable climate-friendly practices to cope with the negative effects. The utilisation of more fertilizers and pesticides may be a controversial choice but is likely to be associated with farm intensification, that helps smallholders in increasing the productivity (Cassman, 1999), that again is reducing their need to expand the production into new crop areas. The adoption of sustainable practices and farm intensification is likely to be associated with participation in long-term outreach projects. Such projects support farmers in enhancing their learning (Baumann et al., 2020; Gutiérrez-Montes & Ramirez-Aguero, 2015) to adapt their production systems.

The future of coffee and cocoa agroforestry

We then looked how suitable to climate change are the most common trees used in coffee and cocoa agroforestry. Coffee and cocoa are perennial crops with high vulnerability to climate variability. Both are plants which were found in the tropical forests understory in the canopy of bigger trees, they have good development under full sun, but the stress caused by long-term exposure to the sun shortens the lifespan of such plants. Agroforestry management is used to reduce the stress and provide a large lifespan for the plantation, and also provide other benefits such as N-fixing in the soil using legume trees. First, we assessed how suitable are the coffee and cocoa habitats to climate change in Central America. The results showed that, by 2050, between

55–62% of current areas for coffee production will likely become unsuitable (Fig. 5A), especially in mid-altitudinal areas (400–700 m a.s.l.). Highlands (>1,800 m a.s.l.) may partly compensate these losses, where coffee will likely expand up to 9–13%. This result confirms the findings of previous studies on Arabica coffee vulnerability (Bunn et al., 2015; Ovalle-Rivera et al., 2015).

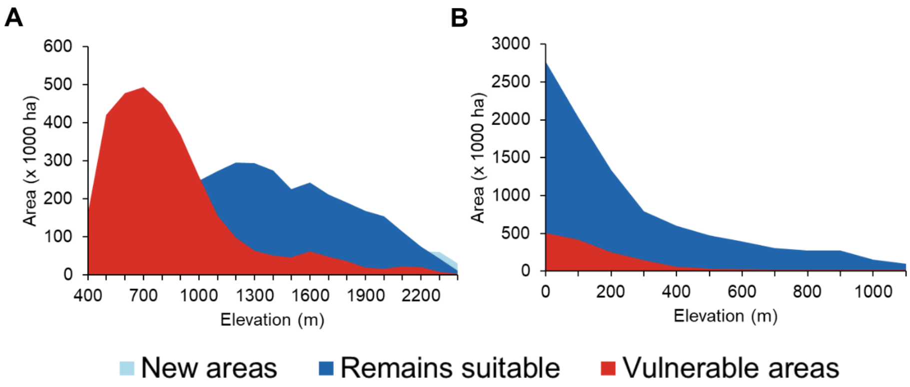

Figure 5: Shifts in suitability due to climate change by 2050 across the altitudinal gradient of (A) coffee (*Coffea arabica* L.) and (B) cocoa (*Theobroma cacao* L.) in Central America

In contrast, cocoa is likely to lose between 13–17% of the current distribution range (Fig. 5B) especially in dry lowland areas (0–300 m a.s.l.), expected to become drier in the next decades (Lyra et al., 2017). Humid areas along the Atlantic coast will remain suitable for cocoa, and have an overlap with a portion of vulnerable coffee areas, showing that cocoa could potentially replace 85% of the vulnerable coffee areas (Fig. 6).

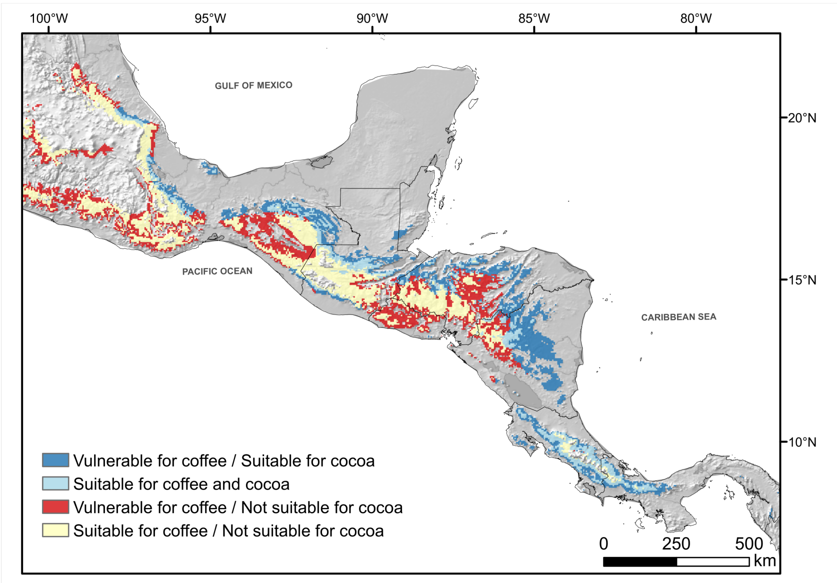

Figure 6: Potential areas where cocoa (*Theobroma cacao* L.) could replace coffee (*Coffea arabica* L.) under climate change. Dark blue indicate vulnerable areas for coffee that can be replaced by cocoa. Light blue indicate areas suitable for coffee and cocoa. Red indicate vulnerable areas for coffee where cocoa is not an alternative under climate change. Light yellow indicate remaining areas for coffee where cocoa is not suitable.

Even as a current trend (Cohen & Castro, 2016; Gross, 2014; Renteria & Rowling, 2016), changing coffee with cocoa is a dramatic alternative for most smallholder farmers in the region. It requires a series of well-structured efforts to reduce the costs of transformation (maybe subsidised by the cocoa industry) and ensure that farmers are well trained to deliver a product that meets the strict market requirements (Levai et al., 2015), to name a few. Additionally, a recent study showed that the impacts of climate change on coffee could be lower than what was projected (DaMatta et al., 2019). For example, coffee could find optimal growth conditions with the increase of CO_2 availability, new varieties with local adaptation are in the breeding pipeline (Arguedas-Ortiz, 2019; Marie et al., 2020; Pruvot-Woehl et al., 2020), and technological changes could sustain coffee production in the future. Is important to note that projections based on ecological niches, as the object of our study, do not take into account site-specific agroecological factors, and should be used as a proxy to identify vulnerable areas and define adaptation strategies, like those discussed by DaMatta et al. (2019).

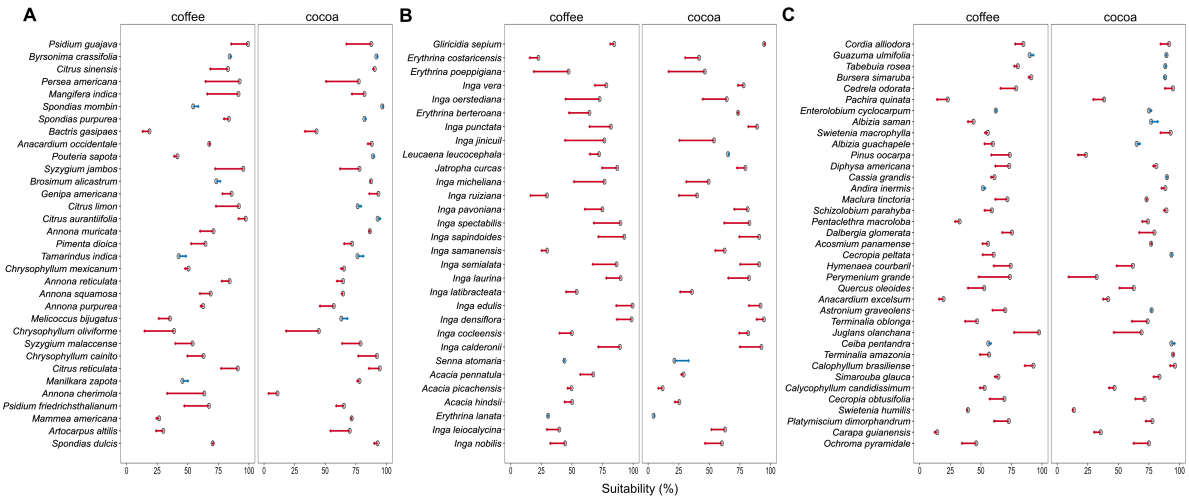

Figure 7: Expected changes in suitability due to climate change of the most common (A) fruit trees, (B) N-fixing trees and (C) timber trees in coffee (*Coffea arabica* L.) and cocoa (*Theobroma cacao* L.) habitats in Central America. Grey dot represents the area of a given species under the current climate conditions. Red arrows (left direction), represent decrease in suitable areas. Blue arrows (right direction) represent increase in suitable areas. Species ordered by main use and by their abundance (from top to bottom) in the inventoried coffee and cocoa farms across Central America.

Alternatively, by managing agroforestry systems, farmers could potentially maintain their current coffee and cocoa plantations using suitable trees to ameliorate microclimatic conditions. This alternative could also prevent the expansion of agricultural activities towards protected areas that are reported to be suitable in the future (Schroth et al., 2015). Although it may require a change in the current agroforestry combinations given our projections showing a high vulnerability of the most preferred tree species (Fig. 7).

Looking at specific tree groups by their main use, we estimate that 20 of the 33 fruit trees will lose more than 15% of their current suitability in coffee areas. The same trend is observed for 14 fruit trees in cocoa suitable areas. High losses (>15%) are expected for 25 of the 30 N-fixing tree species assessed in coffee and for 18 N-fixing tree species in cocoa areas. We also estimate loses of >15% for 22 of 37 species in coffee and 12 tree species in cocoa areas. Most of these losses accounts for the most preferred tree species, such as the timber species cedar (*Cedrela odorata* L.), the fruit species avocado (*Persea americana* Mill.), and the N-fixing poró (*Erythrina poeppigiana* (Walp.) O.F.Cook).

Despite the overall losses in suitability for some of the most popular tree species, our projections suggest that agroforestry could persist as a viable alternative to manage coffee and cocoa plantations in Central America. Approximately 72% of coffee areas (both, remaining and vulnerable) will be suitable for more than 30 tree species. This includes a portfolio of at least 10 species per main use (10 fruit species, 10 N-fixing species and 10 timber species). Most of these tree species are already present in coffee plantations (as identified in the baseline inventories) but mainly in low densities and remain underutilised. Only 9% of coffee areas have very low tree species options (<3 species). The results also suggest that cocoa suitable areas have a higher potential for agroforestry than coffee, with 95% of cocoa areas being suitable for more than 30 tree species. Only 3% of cocoa areas have very low tree species options (<3 species) potentially available

Coffee areas with high potential to select tree species from a portfolio of at least 10 species per main use include the highlands across the Pine-Oak Forests and Petén-Veracruz Moist Forests in Mexico. In Honduras, the Pine-Oak Forests and Mountain Forests, and across the Talamancan Montane Forests in Costa Rica. Areas with high vulnerability (<3 species per main use) are identified across the midlands Pine-Oak Forests and Dry Forests of Honduras, Nicaragua, El Salvador and highlands in Mexico.

Cocoa areas with high potential for selecting trees from a portfolio of more than 30 species (10 per main use) cover all the humid tropical forest at lowlands across the Pacific coast in Costa Rica, Atlantic coast in Nicaragua, Mosquitia in Honduras, Belize, lowlands north of Cobán and south of Sierra Madre in Guatemala and lowlands in the Gulf of Mexico. There is also a high potential for selecting more than 30 species across the transition zones of the Central American Moist Forests and Central American dry forests in Nicaragua, the Dry Forests of El Salvador and the Moist Forests of Costa Rica. Vulnerable cocoa areas with low agroforestry options (< 3 species), are identified across the Dry Forest of Honduras.

The results of our study show that is highly probable that current agroforestry schemes will need to be modified in terms of species composition, since some of the most popular tree species are also vulnerable to future climates. It is particularly concerning the losses in habitat suitability of N-fixing trees since these species make up the most abundant agroforestry trees in coffee and cocoa plantations in the region (Cannavo et al., 2011; Peeters et al., 2003) and have a key role for the management of soil fertility, especially in low-input and smallholder farms (Schnabel et al., 2017).

Rethinking current agroforestry species composition in coffee and cocoa landscapes requires the identification of the best tree species. We found an opportunity for the underutilised species which are present in low densities in coffee and cocoa plantations, and most of them are remnants of previous vegetation (Ordoñez et al., 2014). Expanding the adoption of underutilised species in agroforestry systems will require a deeper understanding of their agronomic performance considering other factors beyond just climate (*e.g.* pests, diseases, soil fertility), crop × tree interactions, farmers' perceptions and local knowledge regarding management and utilisation of these tree species, as well as market incentives to facilitate their wider use. Therefore, selecting the best climate-adapted agroforestry designs is one of the big challenges for the future of cocoa and coffee agroforestry.

Some authors argued that agroforestry (in the case of cocoa) could be less resilient to extreme climates than under full sun (Abdulai et al., 2017), and despite the obvious controversies in the study (Norgrove, 2018), the main message is that a bad agroforestry design may hamper known benefits of agroforestry (Andres et al., 2018; Armengot et al., 2020; Blaser et al., 2018; Schnabel et al., 2017). The work of Padovan et al. (2018, 2015) studying root interactions and water utilisation in coffee agroforestry in a transition area (from dry forest to rainforest) of Nicaragua brought new knowledge into the crop × tree interactions showing that the evergreen tree *Simarouba amara* Aubl. is more more suitable as coffee shade tree compared to the deciduous tree *Tabebuia rosea* (Bertol.) Bertero ex A.DC. due to the competition for water. A study by Cerdán et al. (2012), showed that farmers are aware of such interaction and classify these trees as 'fresh' (suitable for integration) or 'hot' (unsuitable) based on their leaf texture and size, foliage density, crown shape, and root system attributes. Agroforestry is an ancient agricultural practice (Levis et al., 2017; Maezumi et al., 2018; Nair, 1993), but also a new discipline with its first concepts being developed in the late 1970's (Nair, 1993; Spurgeon, 1979), and there is still a number of questions to be explored. More recently, for example, Sauvadet et al. (2020) used an approach with phylogenetic analyses that may help in selecting the most appropriate shade trees in cocoa agroforestry systems. The utilisation of the functional diversity approach (Díaz et al., 2016; Suárez Salazar et al., 2018) linked to farmers knowledge (Cerdán et al., 2012), can also answer a series of open questions on how shade trees interact with crops (during their different phenological stages) and how to best design climate-adapted agroforestry systems.

Overall, the results of our study are just a starting point to develop lines of research that support the re-design of agroforestry schemes and open new venues of research to adapt coffee and cocoa production systems in Central America.

Crop variety management

We then looked on how to define adaptation strategies in seasonal crops. First, we tested whether the model with climatic variables was able to outperform models tested with farmer-generated rankings. The three case studies, with different crops in Nicaragua, Ethiopia and India, provided independent confirmation of the predictive value of the *tricot* trials (Table 4), the model with only climate covariates has the best fit in all cases. Various factors influenced model fit, including farmers' observation skills and environmental variation. We found that most of the differences were among countries, likely due to the different levels of diversity within the sets of varieties.

In Ethiopia, farmers evaluated a large poll of varieties with easily observable differences in performance, while in Nicaragua and India, farmers evaluated a small set of varieties with relatively homogeneous performance.

The results demonstrated the ability of the model with climatic covariates to capture the environmental variability of the sampled environments. This means that the climatic covariates contain unique and substantial information explaining varietal performance. For Nicaragua, we found that common bean variety performance changed when the maximum night temperature exceeded 18.7 $°C$. For durum wheat in Ethiopia, varietal performance was very much related to cold night temperatures during the vegetative period, the performance changed when night temperature exceeded the 8.4 $°C$. For bread wheat in India, varietal performance patterns changed with the diurnal temperature range (DTR) during the vegetative period, which is the difference between minimum and maximum daily temperatures, when DTR exceeded the 14.5 $°C$ and 15.7 $°C$. These findings correspond to the threshold temperature for heat stress or cold acclimatisation reported in the literature for each crop species (Fowler, 2008; Rainey & Griffiths, 2005; Rao et al., 2015).

Table 4: Goodness-of-fit (pseudo-R^2) of Plackett-Luce models linked to explanatory variables. Four models were built using farmer-generated rankings and compared to each other. Intercept-only is a model with no covariates; Design model uses geolocation, season, planting dates, and soil categories, which represents the experimental design; Climate model uses climatic covariates selected with forward selection; and Climate + geolocation is a combination of the Climate model plus geolocation.

Model	Nicaragua	Ethiopia	India
Intercept-only	0.1484	0.3947	0.0381
Design	0.1869	0.4709	0.0721
Climate	**0.1978**	**0.4870**	**0.0882**
Climate + geolocation	0.1977	0.4720	0.0872

We showed that the climatic analysis can improve variety recommendations by incorporating seasonal forecasts, and generate variety recommendations for wider areas trough spatial extrapolation (Fig. 8 A, B, C). Since long-term forecasts were not available by the time of the research, we prepared representative seasonal scenarios of past climate conditions of each site by extracting the last 15 years of seasonal climate data derived from the MODIS dataset. For Nicaragua we show that official variety recommendations fail to identify superior bean varieties that are sufficiently heat tolerant for the study area. In Ethiopia, the findings can improve variety recommendations for durum wheat by uncovering the importance of cold adaptation. In India, the analysis of the *tricot* trial data adds geographic specificity to the existing variety recommendations and suggests that a broader set of bread wheat varieties should be promoted to take into account the climatic differences across the study area. We quantified how much farmers could benefit from the variety recommendations by calculating variety reliability, the probability of outperforming a check variety. For each location, we compared the recommendations produced in the study with the previous recommendations as the check. Reliabilities ranged from 0.59 to 0.65 in Ethiopia, from 0.58 to 0.60 in Nicaragua, and from 0.51 to 0.62 in India (Fig. 8 D, E, F), indicating substantial benefits for large areas.

Overall, the *tricot* citizen science data revealed generalisable relations between seasonal climate variables and crop variety performance that corresponded to known yield-determining factors (Challinor et al., 2014; Villegas et al., 2016). By scaling the agricultural experimentation we were able to track climate trends as they manifest themselves on farms, adjust variety recommendations and recommendation domains, and contribute to understanding how climate affects on-farm varietal performance. The main contribution is that the citizen science data can be linked to seasonal forecasts to provide tailored crop variety recommendations to smallholder farmers in challenging crop production environments. Other opportunities of this approach for climate adaptation have been discussed elsewhere, arguing that it can be used in agroecological intensification (Nelson, 2020), can enhance farmers' access to high-quality germplasm (van Zonneveld et al., 2020), can enhance the accuracy of self-reported data in remote sensing (Paliwal & Jain, 2020), and can become part of decentralised plant breeding strategies for climate adaptation (Ramirez-Villegas et al., 2020). I discuss the later opportunity in the next section.

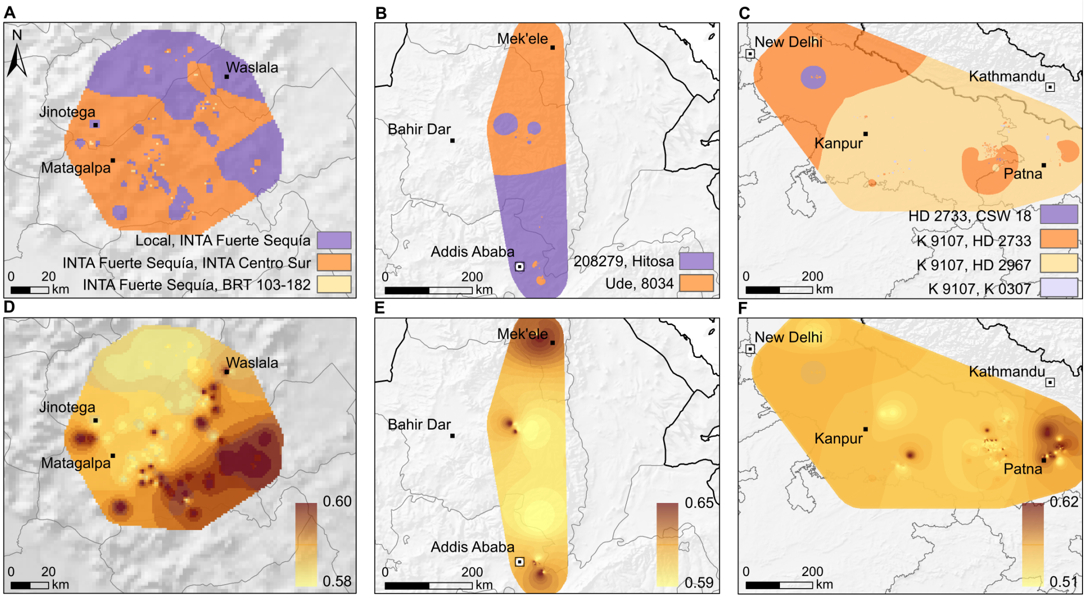

Figure 8: Citizen science can improve variety recommendations. Top two varieties for each area according to their probability of winning over a base period (2002–2016), in (A) Nicaragua, (B) Ethiopia and (C) India. Probability of outperforming (reliability) existing varietal recommendations by using crop varieties recommendations generated with the tricot citizen science approach in (D) Nicaragua, (E) Ethiopia and (F) India.

Genotype selection in challenging crop environments

We explored the opportunity of the *tricot* approach in support decentralised breeding strategies for climate adaptation. We focussed on the durum wheat trials in Ethiopia, linking the farmer-generated data in the *tricot* trials with environmental and DNA data to track $G \times E \times M$ interactions that could support the selection of local-adapted genotypes for crop breeding (Paper 4). We call this approach *data-driven decentralised breeding*, or 3D-breeding. We established a benchmark that represents a centralised breeding approach that is a competitive alternative to 3D-breeding. We focused on grain yield (GY) and farmers' overall appreciation of genotypes (OA), which were both recorded in centralised (station) and decentralised (farm) trials. The goodness-of-fit was assessed using the Kendall τ correlation between observed and predicted rankings in each trial, with higher values meaning a higher ability to predict the rankings.

The results show that 3D-breeding consistently provided higher accuracy than the benchmark for both GY and OA, and also that OA in 3D-breeding produces higher accuracies than GY (Table 1). Previous studies showed that farmer evaluations are able to capture agronomic performance of genotypes in untested locations (Annicchiarico et al., 2019; Kidane et al., 2017). Farmers provided OA according to their own experience and preferences, and it presumably depended on a combination of traits of which GY represented only one dimension (Mancini et al., 2017).

Table 5: Performance of the 3D-breeding compared with the benchmark of a centralised genomic selection. 3D-breeding provides higher across-season goodness-of-fit (Kendall τ) than centralised genomic selection on farmers' overall appreciation (OA) and grain yield (GY).

Approach	OA	GY
Centralised GS		
Season 1 (n=179)	0.134	-0.012
Season 2 (n=651)	0.105	0.076
Season 3 (n=335)	0.183	0.073
	0.141 (± 0.039)	**0.046 (± 0.049)**
3D-breeding		
Season 1 (n=179)	0.270	0.160
Season 2 (n=651)	0.276	0.078
Season 3 (n=335)	0.203	0.119
	0.251 (± 0.040)	**0.109 (± 0.041)**

We show that 3D-breeding can identify genotypes with local adaptation traits. The best three genotypes in each terminal node of the 3D-breeding model (from the Plackett-Luce trees) had a genetic background markedly separated from that of varieties currently recommended for the region, and consistently higher *worth* (Fig. 9). Indeed, the model selected genotypes derived from landraces over improved varieties. We estimated the probability that the model recommendation exceeds the current recommendation in terms of OA. In this assessment, predictions from 3D-breeding outperformed the current varietal recommendations in most of the farmers' fields, with consistently higher probabilities (0.83-0.91), including in marginal areas for which the centralised breeding approach could not provide accurate predictions.

In centralised breeding, the environmental variation of target environments is factored through experimental control or indirectly as an average response across breeding stations as in our benchmark. This makes extrapolation to real farming conditions challenging. The 3D-breeding approach addresses the low correlation between performance in selection environments and production environments, while taking a step forward to fully data-driven breeding and may speed up the turnover of varietal release to address the climate change challenges (Cai et al., 2014; Challinor et al., 2016; Ray et al., 2012). The expansion of the design with the addition of further testing seasons and local management conditions may allow to highlight drivers of local performance of genotypes beyond temperature (Kehel et al., 2016). Indeed, as the model grew in complexity (by adding the DNA data), it was able to retain another environmental variable (maximum night temperature during reproductive growth) differently from our previous case study with only climate variables (van Etten, de Sousa, et al., 2019).

The advantages provided by the approach are clear, phenotyping costs would be divided in much smaller packets, supporting the modular expansion of the breeding effort towards new genetic materials or new locations. In return, each generated data-point would be a better representation of the true farming conditions to which varieties are directed. As farmers are at the centre of the experimental design, varieties deriving from 3D-breeding are more likely to be adopted and suited to local cultivation (Ceccarelli, 2012; Rhoades & Booth, 1982), increasing the effectiveness of breeding efforts. Indeed, we found that farmers' OA was a better predictor than GY in predicting yield realised both in centralised and decentralised trials (Table 5). However, there are a number of open questions in relation to decentralised crop breeding, including how to best motivate new farmers to participate in the evaluation of materials, how much planting material each farmer needs, the logistics of providing farmers with the genetic material, and how to share benefits deriving from the utilisation of farmers' knowledge to produce new varieties.

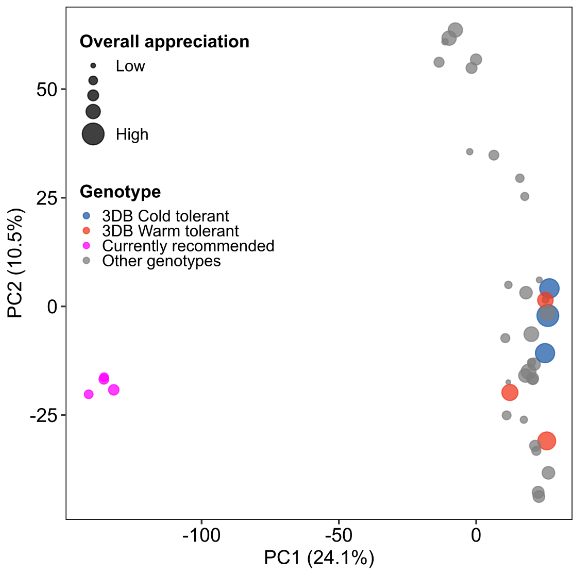

Figure 9: Selection of durum wheat genotypes based on 3D-breeding. Principal component coordinates of the genetic diversity of tested genotypes. Pink dots represent the varieties currently recommended for the area of study. 3DB Cold tolerant (blue) represents the top 3 genotypes selected by 3D-breeding in cold areas (minimum night temperature < 11.5 °C). 3DB Warm tolerant (red) represents the top 3 genotypes selected by 3D-breeding in warm areas (minimum night temperature > 11.5 °C). Size of dots represents the performance of genotypes in farmer fields as overall appreciation (OA).

Overall, we show that the data-driven focus of 3D-breeding enables embracing the complexity of real-world $G \times E \times M$ interactions for the benefit of breeding. Such a multidimensional, collaborative approach calls for best practices in data management and sharing (Leonelli et al., 2017). 3D-breeding is based on a documented set of methods, from experimental design (van Etten, Beza, et al., 2019) to data curation and analysis (de Sousa, van Etten, Dumble, et al., 2020; Turner et al., 2020). Some of these methods were specifically developed to enable the analysis and inference of the *tricot* data, which I explain in the next section. We show that the crowdsourced citizen science approach associated with open-source digital tools makes it possible for breeders and farmers to apply 3D-breeding in new contexts and crops to complement traditional breeding.

A workflow to analyse crowdsourced citizen science data

During the implementation of the activities in this project (Paper 1, 3 and 4) we developed a set of tools and methods to accommodate the *tricot* data into a dynamic workflow in R and that could update the models and provide new insights as the database grows in number of data-points (Paper 5 and 6). These were our first attempts in moving the *tricot* data analysis into a machine learning framework (James et al., 2013). The workflow (Fig. 10) follows the steps: (*A*) Several participants contribute with small tasks, as explained in the *tricot* description, (*B*) Explanatory variables are added (*e.g.* using geographical coordinates and planting dates, or even DNA markers), (*C*) Model selection to find the variables that best explain the data, (*D*) Automated reports can be generated and provide feedback to participants, (*E*) A stable recursive partitioning tree is used for further analysis and inference.

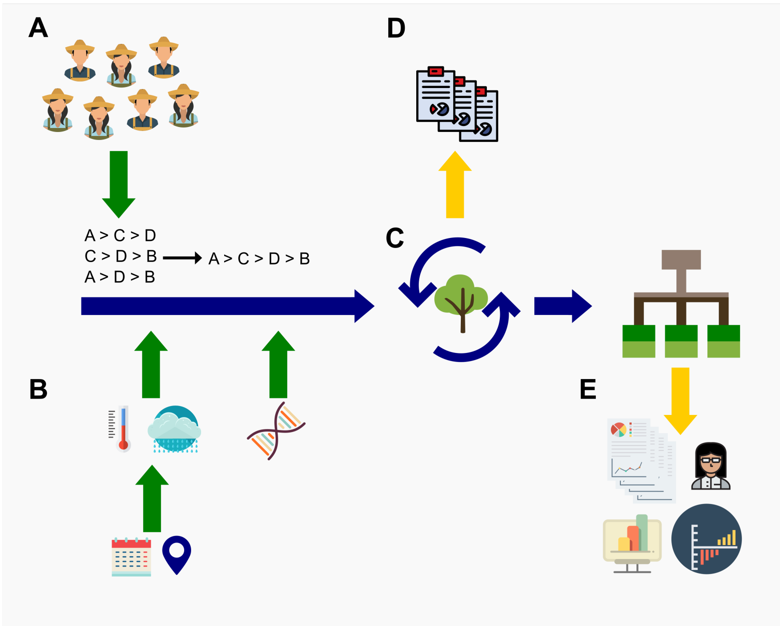

Figure 10: Workflow to analyse crowdsourced citizen science data. (A) Several participants contribute with small tasks; all data is combined using rankings. (B) Covariates are linked to the rankings using georeferenced information and planting dates. (C) Selection of the most relevant covariate(s) using forward selection approach. (D) Automated reports can be generated to give feedback to participants in A. (E) A stable tree is used for further analysis and inference.

The ClimMob platform (https://climmob.net/) storages the *tricot* data in the cloud. To start the workflow, we developed `ClimMobTools` (de Sousa, van Etten, & Madriz, 2020) an application programming interface (API) client that allows the user to fetch their *tricot* data into the R section preserving the original data. This package is designed to work with the ClimMob platform and has no clear application in other domains.

For the other tools, we focused on generalisation, which means that they can be used in other domains outside the proposed workflow. The package `gosset` (de Sousa, van Etten, Dumble, et al., 2020) provides the methods to analyse the *tricot* and other metadata, from collating data (transforming into rankings) to model selection and visualisation. It supports steps A, C, D and E of the workflow. The cross-validation and forward selection procedure, as well as the Akaike weights, pseudo-R^2, Kendal τ, and worst regret computation used in Paper 3 and Paper 4 are implemented in this package. Additionally, it also has visualisation tools that help the explanatory analysis of any rank-based data.

During the data discovery for Paper 3 we tried a number of statistical learning process to analyse the *tricot* data (*e.g.* boosting, bootstrap aggregating, cross-validation, forward selection). We found that the combination of cross-validation + forward selection + Akaike weights was the most reasonable method to work with the data. Afterwards, the scripts used in the analysis were converted into functions in the package `gosset` allowing it to be applied in other research domains using Bradley-Terry, Plackett-Luce or Generalised Linear Models.

To link environmental covariates with the rankings (Fig. 10B) we developed the packages `chirps` (Paper 5) (de Sousa, Sparks, et al., 2020) and `climatrends` (Paper VI) (de Sousa, van Etten, & Solberg, 2020). The package `chirps` offers an API client for the CHIRPS data (Funk et al., 2015) in R. With the package, users can fetch CHIRPS data into the R section and compute rainfall indices that can be linked to the workflow or used in other research domains. The package `climatrends` supports analysis of trends in climate change, ecological and crop modelling by computing temperature and precipitation indices. It started as a set of R scripts used in Paper 1 (de Sousa, Casanoves, et al., 2018) and, as it grew in complexity and demand, it was implemented as a R package. The indices implemented in the package derived from literature showing case studies with applications in a diverse set of areas (Aguilar et al., 2005; Challinor et al., 2016; Kehel et al., 2016; Prentice et al., 1992; Trnka et al., 2014). The climatic data required to compute such indices can be provided by the user (*e.g.* from data loggers) or alternatively using the R package `nasapower` (Sparks, 2018) and `chirps`, further details are given in the dynamic documentation in the packages' website.

One of the steps to better address smallholder farmer adaptation is the evaluation and learning (van Zonneveld et al., 2020). In step D, supported by the package `gosset`, we provide the supporting methods to give feedback to participants in the *tricot* experiments, a procedure that is automated in the ClimMob platform (de Sousa, Dumble, et al., 2020). Overall, these tools are our attempt to provide to the research community a set of open source tools for data analysis, and also to enable the replication of the analysis reported in this thesis.

Conclusions

This thesis provided insights to support climate adaptation strategies in agriculture. We identified how smallholder farmers in Central America reacted to changes in climate patterns and what were the main socio-economic drivers for their adaptation decisions. Overall, farmers had a tendency to select a set practices which is likely related to their participation and training in long-term research for development programs that were conducted across the research areas.

Reforestation and restoration (linked to agroforestry) is the most preferred adaptation decision among the participants in this research. Crop variety management is considered a risk activity and was identified as the least choice among the top-5 adaptation decisions reported by farmers.

Farmers in Central America select a number of tree species to intercrop but had a clear preference in increasing tree density using a small set of species. Based on the projected habitats of the 100 most common trees in coffee and cocoa production systems, we show that some of the most preferred trees are also the most vulnerable to climate change. We argue that re-thinking the design of the current agroforestry schemes is necessary. Farmers have the option to increase the density of underutilised trees already present in most of the current coffee and cocoa systems. Transformation costs and lack of markets for underutilised trees are among the bottlenecks for the adaptation of such crop systems. Modern approaches as phylogenetics and functional traits showed positive prospects to support the selection of species based on crop × tree interactions. The current advances in selecting coffee and cocoa genotypes to develop new locally adapted varieties can also provide, together with agroforestry, a new hope to smallholder farmers that are strongly vulnerable to climate change.

By scaling agricultural experimentation powered with citizen science we developed and validated the *tricot* approach, that can generate tailored recommendations for crop variety management. This can significantly reduce the risks of smallholder farmers in managing crop varieties across different seasons. The approach was validated with data-points from thousands of farmer-managed plots in smallholder farms in three countries (Nicaragua, India and Ethiopia). The outcomes of this study, however, can most likely be applied to a diverse set of regions (beyond tropical areas) and farming systems (beyond smallholder farming systems). We also showed the ability of the *tricot* approach, linked to DNA and environmental data, to support breeding programs to fully track $G \times E \times M$ interactions and select genotypes with local adapted traits, specifically in challenging and diverse crop production environments. The analytical workflow that was developed for this research can be employed to provide new insights as new data becomes available in the future.

Future research and perspective

There is a series of open questions to explore on how to identify new agroforestry designs that both provide benefits to farmers and are resilient to future climates. The increasingly number of datasets being made available by research organisations in open source databases (*e.g.* Dataverse, Zenodo, CGIAR Gardian) may support the meta-analysis of large datasets to infer patterns on crop × tree interactions to deliver recommendations based on the portfolio theory (Blandon, 1985). New studies on phylogenetics (Sauvadet et al., 2020) and farmers' local knowledge (Cerdán et al., 2012) also offers an excellent opportunity to provide data-driven insights for future agroforestry designs.

Future research on crop variety management would explore the other dimensions of the *tricot* data. With Paper 4 we also demonstrated the potential in joining the farmer-generated data with scientist-generated data (DNA, grain yield, satellite data, etc.) to capture some of the layers of information provided by citizen scientists. The data-driven approach could be used to further developed and integrate different topics in agroecology, maybe joining information from plant $\times$ insect interactions and the interplay of using a diversified portfolio of species/varieties and other organisms that are part of the production system. The *tricot* approach can also be further developed to be adjusted to the needs of the civil society, maybe integrated into urban agriculture, digital agricultural markets and extension programs.

Additionally, understanding farmers' planting decisions (based on the recorded planting dates) could provide information on the environmental factors that need to be considered to adjust the crop growing calendar for climate adaptation. We also see a potential to employ farmer-generated data with crop models targeting challenging crop production environments by, for example, using the new indices for physiological stress (Challinor et al., 2016; Trnka et al., 2014) implemented in the package `climatrends`, such as heat stress events, lethal temperatures and crop duration index. Incorporating socio-economic data derived from the Rural Household Multi-Indicator Survey (RHoMIS) (Hammond et al., 2017) could unlock a new factor (S) to the interactions that drive crop variety performance and management decisions, in that way tracking $G \times E \times M \times S$ interactions as they occur in the system. Future research would look for adoption rates of new crop varieties and look for the intertwining of farmers' overall appreciation with other traits (*e.g.* resistance to pests, market value) that are also registered in the *tricot* trials but never explored until now. In the final line of the breeding program, it is also important to consider the consumers' preference and market acceptance on new products, the *tricot* approach could support such assessments and orient breeders and markets on the development and fine tuning of agricultural products for resilient and sustainable food systems.

Funding statement

The research in the papers attached alongside this thesis has been possible thanks to the financial support of multiple funding agencies. We thank the Royal Embassy of Norway (contract no. CAM-2647-12/0004, MAP Norway Second Phase). The US Agency for International Development, which supported this research via the Cooperative Agreement AID-OAA-F-14-00035, made possible by the generous support of the American people. The McKnight Foundation (grant num. CCRP 16–098). The German Federal Ministry for Economic Cooperation and Development (grant num. 81194988). The Nordic Joint Committee for Agricultural and Food Research (grant num. 202100-2817). And Hivos International (grant num. HI-A70D41/A1092).

Acknowledgements

This thesis would be likely impossible without the contribution of a number of people, to be exact, 34 co-authors who contributed in some extent to the development of this work. Thanks to the more than 10,000 farmers citizen scientists who evaluated varieties in Nicaragua, Ethiopia, and India. They made it possible for me to develop a significant part of this thesis. I really appreciate all the contribution. Thanks to Elias de Melo Virginio Filho and Ronnie De Camino for opening the doors for my career in international agricultural research. One special thanks to Jenny Ordoñez, Maarten van Zonneveld, Milena Holmgren, Leida Mercado and Fernando Casanoves who guided me in the first steps that I took towards this project. Your support, advice, encouragement, and brilliant insights are something that I will keep with me for the rest of my life. Thank you!

Thanks to Svein Ø. Solberg for being such an amazing supervisor, proving great support and always great advices that made possible for me to develop this work. Thanks to my co-supervisor, Jacob van Etten, for his brilliant ideas that inspired me during these years. Thanks to Fred H. Johnsen for his great support in my arrival to Norway and the early steps of my PhD project. Thanks to Matteo Dell'Acqua and Mario Enrico Pè for being such a nice hosts during my research stay at the Scuola Superiore Sant'Anna in Italy. Thanks to Øyvind Antonsen and the Department of Biotechnology for hosting me at *Biohus* during a significant part of this project. Thanks to Isolde D. K. Ferraz and Geângelo P. Calvi for the excellent research stay at the National Institute of Amazonian Research (INPA) in Brazil. Thanks to Magne Neby for the nice discussions while we shared the office in Blæstad, the logistic support in moving from Ottestad to Hamar, and the help in editing this thesis. Thanks to Eliana Bontti for all the advices and nice talks that helped me a lot. Thanks to Olivier Devineau for being always supportive in giving statistical advice. Thanks to Henriette Wathne Gelink, Kaja Johnsen, Sarah Louise Loftheim, Elisabeth Røe, Morten Tofastrud, Lucy Walmsnæss, Maria Hörnell Willebrand and Ragnhild Østerhagen for the administrative support and advice. Also thanks to my Department colleagues Thomas Cottis, Hans Christian Endrerud, Dag Jørgensen, Marius Kjønsberg, Stein Amundsen Morthov, Terje Motrøen, Jill Peng and Lars Erik Ruud for all the support and companionship. Thanks to Nora Castañeda-Álvarez for the inputs and comments during the mid-term evaluation that helped in the improvement of this thesis. Thanks to the International Research School in Applied Ecology (IRSAE) for providing learning and networking opportunities via the Summer Schools, Climate Change Group and Courses.

Thanks to my wife Suelen, my son Kairê and my daughter Liana for being always present, supportive and inspiring. Thanks to my parents for the great effort to provide the best possible for the family. Thanks to my friends Jorge Sellare and Juan Pablo Iñamagua for the nice discussions and great moments that we had in the course of these years.

Thank you!

References

Abdulai, I., Vaast, P., Hoffmann, M. P., Asare, R., Jassogne, L., Van Asten, P., Rötter, R. P., & Graefe, S. (2017). Cocoa agroforestry is less resilient to sub-optimal and extreme climate than cocoa in full sun. *Global Change Biology*, *2018*(24), 273–286. https://doi.org/10.1111/gcb.13885

Agresti, A., & Kateri, M. (2011). *Categorical Data Analysis* (M. Lovric, Ed.; pp. 206–208). Springer Berlin Heidelberg. https://doi.org/10.1002/0471249688

Aguilar, E., Peterson, T. C., Obando, P. R., Frutos, R., Retana, J. A., Solera, M., Soley, J., García, I. G., Araujo, R. M., Santos, A. R., Valle, V. E., Brunet, M., Aguilar, L., Álvarez, L., Bautista, M., Castañón, C., Herrera, L., Ruano, E., Sinay, J. J., … Mayorga, R. (2005). Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. *Journal of Geophysical Research*, *110*(D23), D23107. https://doi.org/10.1029/2005JD006119

Altieri, M. A., & Nicholls, C. I. (2017). The adaptation and mitigation potential of traditional agriculture in a changing climate. *Climatic Change*, *140*(1), 33–45. https://doi.org/10.1007/s10584-013-0909-y

Andrade, D., Pasini, F., & Scarano, F. R. (2020). Syntropy and innovation in agriculture. *Current Opinion in Environmental Sustainability*, *45*, 20–24. https://doi.org/10.1016/j.cosust.2020.08.003

Andres, C., Blaser, W. J., Dzahini-Obiatey, H. K., Ameyaw, G. A., Domfeh, O. K., Awiagah, M. A., Gattinger, A., Schneider, M., Offei, S. K., & Six, J. (2018). Agroforestry systems can mitigate the severity of cocoa swollen shoot virus disease. *Agriculture, Ecosystems & Environment*, *252*, 83–92. https://doi.org/10.1016/j.agee.2017.09.031

Annicchiarico, P. (2002). *Genotype x environment interaction: challenges and opportunities for plant breeding and cultivar recommendations*. Food; Agriculture Organization of The United Nations.

Annicchiarico, P., Russi, L., Romani, M., Pecetti, L., & Nazzicari, N. (2019). Farmer-participatory vs. conventional market-oriented breeding of inbred crops using phenotypic and genome-enabled approaches: A pea case study. *Field Crops Research*, *232*, 30–39. https://doi.org/10.1016/j.fcr.2018.11.001

Arguedas-Ortiz, D. (2019). The accident that led to the discovery of climate-change-proof coffee. *MIT Technology Review*. https://www.technologyreview.com/2019/04/24/135937/the-accident-that-led-to-the-discovery-of-climate-change-proof-coffee/

Armengot, L., Ferrari, L., Milz, J., Velásquez, F., Hohmann, P., & Schneider, M. (2020). Cacao agroforestry systems do not increase pest and disease incidence compared with monocultures under good cultural management practices. *Crop Protection*, *130*, 105047. https://doi.org/10.1016/j.cropro.2019.105047

Atlin, G. N., Cairns, J. E., & Das, B. (2017). Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change. *Global Food Security*, *12*, 31–37. https://doi.org/10.1016/j.gfs.2017.01.008

Atlin, G. N., Cooper, M., & Bjørnstad, Å. (2001). A comparison of formal and participatory breeding approaches using selection theory. *Euphytica*, *122*(3), 463–475. https://doi.org/10.1023/A:1017557307800

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*(2), 303–321. https://doi.org/10.1007/s12571-015-0446-9

Baumann, M. D., Zimmerer, K. S., & van Etten, J. (2020). Participatory seed projects and agroecological landscape knowledge in Central America. *International Journal of Agricultural Sustainability*, *18*(4), 300–318. https://doi.org/10.1080/14735903.2020.1775930

Beer, J., Muschler, R., Kass, D., & Somarriba, E. (1998). Shade management in coffee and cacao plantations. *Agroforestry Systems*, *38*(1-3), 139–164. https://doi.org/10.1023/A:1005956528316

Bellon, M. R., Hodson, D., & Hellin, J. (2011). Assessing the vulnerability of traditional maize seed systems in Mexico to climate change. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(33), 13432–13437. https://doi.org/10.1073/pnas.1103373108

Bellwood, P. (2004). *First Farmers: The Origins of Agricultural Societies* (p. 384). Wiley-Blackwell.

Bergquist, M., Nilsson, A., & Schultz, P. W. (2019). Experiencing a Severe Weather Event Increases Concern About Climate Change. *Frontiers in Psychology*, *10*(FEB), 220. https://doi.org/10.3389/fpsyg.2019.00220

Beza, E., Steinke, J., van Etten, J., Reidsma, P., Fadda, C., Mittra, S., Mathur, P., & Kooistra, L. (2017). What are the prospects for citizen science in agriculture? Evidence from three continents on motivation and mobile telephone use of resource-poor farmers. *PLOS ONE*, *12*(5), e0175700. https://doi.org/10.1371/journal.pone.0175700

Blandon, P. (1985). Agroforestry and portfolio theory. *Agroforestry Systems*, *3*(3), 239–249. https://doi.org/10.1007/BF00046957

Blaser, W. J., Oppong, J., Hart, S. P., Landolt, J., Yeboah, E., & Six, J. (2018). Climate-smart sustainable agriculture in low-to-intermediate shade agroforests. *Nature Sustainability*, *1*(5), 234–239. https://doi.org/10.1038/s41893-018-0062-8

Bonilla Zunhiga, G., Caballero, A., Escalante Santos, M. Y., Llanderal Ocampo, T., Lopez, A., Orozco, L., & Zunhiga Pereira, C. (2014). Database Tipologias cafetaleras en Centroamerica. *Harvard Dataverse*. https://doi.org/10.7910/DVN/27826

Booth, T. H. (2018). Why understanding the pioneering and continuing contributions of BIOCLIM to species distribution modelling is important. *Austral Ecology*, *43*(8), 852–860. https://doi.org/10.1111/aec.12628

Bouroncle, C., Müller, A., Giraldo, D., Rios, D., Imbach, P., Girón, E., Portillo, F., Boni, A., Etten, J. van, & Ramirez-Villegas, J. (2019). A systematic approach to assess climate information products applied to agriculture and food security in Guatemala and Colombia. *Climate Services*, *16*, 100137. https://doi.org/10.1016/j.cliser.2019.100137

Bradley, R. A., & Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, *39*(3/4), 324–345. https://doi.org/10.2307/2334029

Bunn, C., Läderach, P., Ovalle Rivera, O., & Kirschke, D. (2015). A bitter cup: climate change profile of global production of Arabica and Robusta coffee. *Climatic Change*, *129*(1-2), 89–101. https://doi.org/10.1007/s10584-014-1306-x

Cai, W., Borlace, S., Lengaigne, M., van Rensch, P., Collins, M., Vecchi, G., Timmermann, A., Santoso, A., McPhaden, M. J., Wu, L., England, M. H., Wang, G., Guilyardi, E., & Jin, F.-F. (2014). Increasing frequency of extreme El Niño events due to greenhouse warming. *Nature Climate Change*, *5*(2), 1–6. https://doi.org/10.1038/nclimate2100

Campbell, B. M. B. M., Vermeulen, S. J. S. J., Aggarwal, P. K. P. K., Corner-Dolloff, C., Girvetz, E., Loboguerrero, A. M. A. M., Ramirez-Villegas, J., Rosenstock, T., Sebastian, L., Thornton, P. K., & Wollenberg, E. (2016). Reducing risks to food security from climate change. *Global Food Security*, *11*, 0–1. https://doi.org/10.1016/j.gfs.2016.06.002

Cannavo, P., Sansoulet, J., Harmand, J.-M., Siles, P., Dreyer, E., & Vaast, P. (2011). Agroforestry associating coffee and Inga densiflora results in complementarity for water uptake and decreases deep drainage in Costa Rica. *Agriculture, Ecosystems & Environment*, *140*(1-2), 1–13. https://doi.org/10.1016/J.AGEE.2010.11.005

Cassman, K. G. (1999). Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(11), 5952–5959. https://doi.org/10.1073/PNAS.96.11.5952

CATIE, & OFI. (2003). *Árboles de Centroamérica* (J. Cordero & D. Boshier, Eds.; 1st ed., p. 1091). Oxford Forestry Institute; Centro Agronomico Tropical de Investigación y Enseñanza. http://orton.catie.ac.cr/repdoc/a11445e/a11445e.pdf

Caudill, S. A., DeClerck, F. J. A., & Husband, T. P. (2015). Connecting sustainable agriculture and wildlife conservation: Does shade coffee provide habitat for mammals? *Agriculture, Ecosystems & Environment*, *199*(0), 85–93. https://doi.org/10.1016/j.agee.2014.08.023

Ceccarelli, S. (2012). *Plant Breeding with Farmers: A technical manual* (p. 125). International Center for Agricultural Research in the Dry Areas (ICARDA). https://doi.org/https://hdl.handle.net/20.500.11766/7745

Ceccarelli, S., & Grando, S. (2019). Participatory plant breeding: Who did it, who does it and where? *Experimental Agriculture*, *56*(1), 1–11. https://doi.org/10.1017/S0014479719000127

Cerda, R., Deheuvels, O., Calvache, D., Niehaus, L., Saenz, Y., Kent, J., Vilchez, S., Villota, A., Martinez, C., & Somarriba, E. (2014). Contribution of cocoa agroforestry systems to family income and domestic consumption: looking toward intensification. *Agroforestry Systems*, *88*(6), 957–981. https://doi.org/10.1007/s10457-014-9691-8

Cerdán, C. R., Rebolledo, M. C., Soto, G., Rapidel, B., & Sinclair, F. L. (2012). Local knowledge of impacts of tree cover on ecosystem services in smallholder coffee production systems. *Agricultural Systems*, *110*(0), 119–130. https://doi.org/10.1016/j.agsy.2012.03.014

Challinor, A. J., Koehler, A.-K., Ramirez-Villegas, J., Whitfield, S., & Das, B. (2016). Current warming will reduce yields unless maize breeding and seed systems adapt immediately. *Nature Climate Change*, *6*(10), 954–958. https://doi.org/10.1038/nclimate3061

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R., & Chhetri, N. (2014). A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*, *4*(4), 287–291. https://doi.org/10.1038/nclimate2153

Chapagain, D., Baarsch, F., Schaeffer, M., & D'haen, S. (2020). Climate change adaptation costs in developing countries: insights from existing estimates. *Climate and Development*, 1–9. https://doi.org/10.1080/17565529.2020.1711698

Cohen, L., & Castro, I. (2016). As climate change threatens CentAm coffee, a cocoa boom is born. *Thomson Reuters*. http://www.reuters.com/article/us-climatechange-cocoa-coffee-idUSKCN0UW1AV

DaMatta, F. M., Rahn, E., Läderach, P., Ghini, R., & Ramalho, J. C. (2019). Why could the coffee crop endure climate change and global warming to a greater extent than previously estimated? *Climatic Change*, *152*(1), 167–178. https://doi.org/10.1007/s10584-018-2346-4

Dawson, J. C., Murphy, K. M., & Jones, S. S. (2008). Decentralized selection and participatory approaches in plant breeding for low-input systems. *Euphytica*, *160*(2), 143–154. https://doi.org/10.1007/s10681-007-9533-0

de Sousa, K., Casanoves, F., Sellare, J., Ospina, A., Suchini, J. G., Aguilar, A., & Mercado, L. (2018). How climate awareness influences farmers' adaptation decisions in Central America? *Journal of Rural Studies*, *64*, 11–19. https://doi.org/10.1016/j.jrurstud.2018.09.018

de Sousa, K., Detlefsen, G., de Melo Virginio Filho, E., Tobar, D., & Casanoves, F. (2016). Timber yield from smallholder agroforestry systems in Nicaragua and Honduras. *Agroforestry Systems*, *90*(2), 207–218. https://doi.org/10.1007/s10457-015-9846-2

de Sousa, K., Dumble, S., Greliche, N., Madriz, B., & van Etten, J. (2020). Workflow for data analysis and report with crowdsourced citizen science-generated data. *Zenodo*, *v1.1-1*. https://doi.org/10.5281/zenodo.3984711 de Sousa, K., Sparks, A., Ashmall, W., van Etten, J., & Solberg, S. (2020). chirps: API Client for the CHIRPS Precipitation Data in R. *Journal of Open Source Software*, *5*(51), 2419. https://doi.org/10.21105/joss.02419

de Sousa, K., van Etten, J., Dumble, S., Greliche, N., & Steinke, J. (2020). *gosset: Modelling Metadata and Crowdsourcing Citizen Science*. https://agrobioinfoservices.github.io/gosset/

de Sousa, K., van Etten, J., & Madriz, B. (2020). ClimMobTools: API Client for the 'ClimMob' Platform. *The Comprehensive R Archive Network*. https://cran.r-project.org/package=ClimMobTools

de Sousa, K., van Etten, J., Poland, J., Kidane, Y. G., Lakew, B. F., Mengistu, D. K., Jannink, J.-L., Solberg, S. Ø., Fadda, C., Pè, M. E., & Dell'Acqua, M. (2020a). Data-driven decentralised breeding increases genetic gain in a challenging crop production environment. *Under Review*.

de Sousa, K., van Etten, J., Poland, J., Kidane, Y. G., Lakew, B. F., Mengistu, D. K., Jannink, J.-L., Solberg, S. Ø., Fadda, C., Pè, M. E., & Dell'Acqua, M. (2020b). Replication data for: "Data-driven decentralized breeding increases genetic gain in challenging crop production environments. *Harvard Dataverse*. https://doi.org/10.7910/DVN/OEZGVP

de Sousa, K., van Etten, J., & Solberg, S. Ø. (2020). climatrends: Climate Variability Indices for Ecological Modelling. *The Comprehensive R Archive Network*. https://cran.r-project.org/package=climatrends

de Sousa, K., van Zonneveld, M., Holmgren, M., Kindt, R., & Ordoñez, J. C. (2018). Replication data for: "The future of coffee and cocoa agroforestry in a warmer Mesoamerica". *Harvard Dataverse*, *v2*. https://doi.org/10.7910/DVN/0O1GW1

de Sousa, K., van Zonneveld, M., Holmgren, M., Kindt, R., & Ordoñez, J. C. (2019). The future of coffee and cocoa agroforestry in a warmer Mesoamerica. *Scientific Reports*, *9*(1), 8828. https://doi.org/10.1038/s41598-019-45491-7

de Sousa, K., van Zonneveld, M., Imbach, P., Kindt, R., & Ordoñez, J. C. (2017). *Suitability of key Central American agroforestry species under future climates: an atlas* (p. 257). World Agroforestry Centre, Occasional Paper 26. http://www.worldagroforestry.org/output/suitability-key-central-american-agroforestry-species-under-future-climates-atlas

Deutsch, C. A., Tewksbury, J. J., Tigchelaar, M., Battisti, D. S., Merrill, S. C., Huey, R. B., & Naylor, R. L. (2018). Increase in crop losses to insect pests in a warming climate. *Science*, *361*(6405), 916–919. https://doi.org/10.1126/science.aat3466

Díaz, S., Kattge, J., Cornelissen, J. H. C., Wright, I. J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Colin Prentice, I., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P. B., Moles, A. T., Dickie, J., Gillison, A. N., Zanne, A. E., … Gorné, L. D. (2016). The global spectrum of plant form and function. *Nature*, *529*(7585), 167–171. https://doi.org/10.1038/nature16489

Elum, Z. A., Modise, D. M., & Marr, A. (2017). Farmer's perception of climate change and responsive strategies in three selected provinces of South Africa. *Climate Risk Management*, *16*, 246–257. https://doi.org/10.1016/j.crm.2016.11.001

Eshed, Y., & Lippman, Z. B. (2019). Revolutions in agriculture chart a course for targeted breeding of old and new crops. *Science*, *366*(6466). https://doi.org/10.1126/science.aax0025

FAO, IFAD, UNICEF, WFP, & WHO. (2019). *The State of Food Security and Nutrition in the World 2019. Safeguarding against economic slowdowns and downturns* (p. 290). Food and Agriculture Organization of The United Nations. http://www.fao.org/state-of-food-security-nutrition

Fowler, D. B. (2008). Cold Acclimation Threshold Induction Temperatures in Cereals. *Crop Science*, *48*(3), 1147. https://doi.org/10.2135/cropsci2007.10.0581

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*, 150066. https://doi.org/10.1038/sdata.2015.66

GBIF.org. (2020). *GBIF Occurrence Download*. https://doi.org/10.15468/dl.cagrkf

Gibbs, H. K., Ruesch, A. S., Achard, F., Clayton, M. K., Holmgren, P., Ramankutty, N., & Foley, J. A. (2010). Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(38), 16732–16737. https://doi.org/10.1073/pnas.0910275107

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., Tilman, D., Schmidhuber, J., Tubiello, F. N., Piesse, J., Thirtle, C., Pretty, J., Balmford, A., Green, R. E., Scharlemann, J. P. W., . . . Smaling, E. M. A. (2010). Food security: the challenge of feeding 9 billion people. *Science*, *327*(5967), 812–818. https://doi.org/10.1126/science.1185383

Gross, M. (2014). Coffee and chocolate in danger. *Current Biology*, *24*(11), R503–R506. https://doi.org/10.1016/j.cub.2014.05.035

Gutierrez-Montes, I., Arguedas, M., Ramirez-Aguero, F., Mercado, L., & Sellare, J. (2020). Contributing to the construction of a framework for improved gender integration into climate-smart agriculture projects monitoring and evaluation: MAP-Norway experience. *Climatic Change*, *158*(1), 93–106. https://doi.org/10.1007/s10584-018-2231-1

Gutiérrez-Montes, I. A., & Ramirez-Aguero, F. (2015). The Mesoamerican Agroenvironmental Program: Critical lessons learned from an integrated approach to achieve sustainable land management. In E. Méndez, C. M. Bacon, R. Cohen, & S. R. Gliessman (Eds.), *Agroecology: A transdisciplinary, participatory and action-oriented approach* (1st ed., pp. 217–229). CRC Press.

Hammond, J., Fraval, S., van Etten, J., Suchini, J. G., Mercado, L., Pagella, T., Frelat, R., Lannerstad, M., Douxchamps, S., Teufel, N., Valbuena, D., & van Wijk, M. T. (2017). The

Rural Household Multi-Indicator Survey (RHoMIS) for rapid characterisation of households to inform climate smart agriculture interventions: Description and applications in East Africa and Central America. *Agricultural Systems*, *151*, 225–233. https://doi.org/10.1016/j.agsy.2016.05.003

Harvey, C. A., Rakotobe, Z. L., Rao, N. S., Dave, R., Razafimahatratra, H., Rabarijohn, R. H., Rajaofara, H., & MacKinnon, J. L. (2014). Extreme vulnerability of smallholder farmers to agricultural risks and climate change in Madagascar. *Philosophical Transactions of the Royal Society B Biological Sciences*, *369*(1639), 20130089. https://doi.org/10.1098/rstb.2013.0089

Hickey, L. T., N. Hafeez, A., Robinson, H., Jackson, S. A., Leal-Bertioli, S. C. M., Tester, M., Gao, C., Godwin, I. D., Hayes, B. J., & Wulff, B. B. H. (2019). Breeding crops to feed 10 billion. *Nature Biotechnology*, 1. https://doi.org/10.1038/s41587-019-0152-9

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*(15), 1965–1978. https://doi.org/10.1002/joc.1276

Hirvonen, K., Bai, Y., Headey, D., & Masters, W. A. (2020). Affordability of the EAT–Lancet reference diet: a global analysis. *The Lancet Global Health*, *8*(1), e59–e66. https://doi.org/10.1016/S2214-109X(19)30447-4

Holmgren, M., Hirota, M., Nes, E. H. van, & Scheffer, M. (2013). Effects of interannual climate variability on tropical tree cover. *Nature Climate Change*, *3*(8), 755–758. https://doi.org/10.1038/nclimate1906

Holmgren, M., Kabanshi, A., Marsh, J. E., & Sörqvist, P. (2018). When A+B < A: Cognitive Bias in Experts' Judgment of Environmental Impact. *Frontiers in Psychology*, *9*, 823. https://doi.org/10.3389/fpsyg.2018.00823

Holmgren, M., & Scheffer, M. (2010). Strong facilitation in mild environments: the stress gradient hypothesis revisited. *Journal of Ecology*, *98*(6), 1269–1275. https://doi.org/10.1111/j.1365-2745.2010.01709.x

Holmgren, M., Scheffer, M., & Huston, M. A. (1997). The interplay of facilitation and competition in plant communities. *Ecology*, *78*(7), 1966–1975. https://doi.org/10.1890/0012-9658(1997)078%5B1966:TIOFAC%5D2.0.CO;2

Hunter, D. R. (2004). MM Algorithms for Generalized Bradley-Terry Models. *The Annals of Statistics*, *32*(1), 384–406. http://www.jstor.org/stable/3448514

Imbach, P., Chou, S. C., Lyra, A., Rodrigues, D., Rodriguez, D., Latinovic, D., Siqueira, G., Silva, A., Garofolo, L., & Georgiou, S. (2018). Future climate change scenarios in Central America at high spatial resolution. *PLOS ONE*, *13*(4), e0193570. https://doi.org/10.1371/journal.pone.0193570

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R* (G. Casella, S. Fienberg, & I. Olkin, Eds.; 1st ed., p. 426).

Springer-Verlag New York. https://doi.org/10.1007/978-1-4614-7138-7

Jarvis, A., Reuter, H. I., Nelson, A., & Guevara, E. (2008). *Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT)*. http://srtm.csi.cgiar.org

Kehel, Z., Crossa, J., & Reynolds, M. (2016). Identifying Climate Patterns during the Crop-Growing Cycle from 30 Years of CIMMYT Elite Spring Wheat International Yield Trials. In A. Bari, A. B. Damania, M. Mackay, & S. Dayanandan (Eds.), *Applied mathematics and omics to assess crop genetic resources for climate change adaptive traits* (pp. 151–174). CRC Press.

Kendall, M. G. (1938). A new measure of ranking correlation. *Biometrika*, *30*(1-2), 81–93. https://doi.org/10.1093/biomet/30.1-2.81

Kidane, Y. G., Mancini, C., Mengistu, D. K., Frascaroli, E., Fadda, C., Pè, M. E., & Dell'Acqua, M. (2017). Genome wide association study to identify the genetic base of smallholder farmer preferences of durum wheat traits. *Frontiers in Plant Science*, *8*, 1230. https://doi.org/10.3389/fpls.2017.01230

Kindt, R. (2018). Ensemble species distribution modelling with transformed suitability values. *Environmental Modelling & Software*, *100*, 136–145. https://doi.org/10.1016/j.envsoft.2017.11.009

King, M., Altdorff, D., Li, P., Galagedara, L., Holden, J., & Unc, A. (2018). Northward shift of the agricultural climate zone under 21st-century global climate change. *Scientific Reports*, *8*(1), 7904. https://doi.org/10.1038/s41598-018-26321-8

Läderach, P., Ramirez–Villegas, J., Navarro-Racines, C., Zelaya, C., Martinez–Valle, A., & Jarvis, A. (2017). Climate change adaptation of coffee production in space and time. *Climatic Change*, *141*(1), 47–62. https://doi.org/10.1007/s10584-016-1788-9

Leonelli, S., Davey, R. P., Arnaud, E., Parry, G., & Bastow, R. (2017). Data management and best practice for plant science. *Nature Plants*, *3*(6), 1–4. https://doi.org/10.1038/nplants.2017.86

Levai, L. D., Meriki, H. D., Adiobo, A., Awa-Mengi, S., Akoachere, J.-F. T. K., & Titanji, V. P. K. (2015). Postharvest practices and farmers' perception of cocoa bean quality in Cameroon. *Agriculture & Food Security*, *4*(1), 28. https://doi.org/10.1186/s40066-015-0047-z

Levis, C., Costa, F. R. C. C., Bongers, F., Peña-Claros, M., Clement, C. R., Junqueira, A. B., Neves, E. G., Tamanaha, E. K., Figueiredo, F. O. G. G., Salomão, R. P., Castilho, C. V., Magnusson, W. E., Phillips, O. L., Guevara, J. E., Sabatier, D., Molino, J.-F. F., Cárdenas López, D., Mendoza, A. M., Pitman, N. C. A. A., … Ter Steege, H. (2017). Persistent effects of pre-Columbian plant domestication on Amazonian forest composition. *Science*, *355*(6328), 925 LP–931. https://doi.org/10.1126/science.aal0157

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 1–18. https://doi.org/10.18637/jss.v025.i01

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22*(140), 55.

Lin, P., He, Z., Du, J., Chen, L., Zhu, X., & Li, J. (2017). Recent changes in daily climate extremes in an arid mountain region, a case study in northwestern China's Qilian Mountains. *Scientific Reports*, *7*(1), 2245. https://doi.org/10.1038/s41598-017-02345-4

Lipper, L., Thornton, P., Campbell, B. M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A., Garrity, D., Henry, K., Hottle, R., Jackson, L., Jarvis, A., Kossam, F., Mann, W., McCarthy, N., Meybeck, A., Neufeldt, H., Remington, T., ... Torquebiau, E. F. (2014). Climate-smart agriculture for food security. *Nature Climate Change*, *4*(12), 1068–1072. https://doi.org/10.1038/nclimate2437

Liu, C., White, M., & Newell, G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography*, *40*(4), 778–789. https://doi.org/10.1111/jbi.12058

Loboguerrero, A. M., Thornton, P., Wadsworth, J., Campbell, B. M., Herrero, M., Mason-D'Croz, D., Dinesh, D., Huyer, S., Jarvis, A., Millan, A., Wollenberg, E., & Zebiak, S. (2020). Perspective article : Actions to reconfigure food systems. *Global Food Security*, *26*(September), 100432. https://doi.org/10.1016/j.gfs.2020.100432

Lowder, S. K., Skoet, J., & Raney, T. (2016). The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Development*, *87*, 16–29. https://doi.org/10.1016/j.worlddev.2015.10.041

Luce, R. D. (1959). *Individual Choice Behavior* (p. 153). Courier Corporation.

Lyra, A., Imbach, P., Rodriguez, D., Chou, S. C., Georgiou, S., & Garofolo, L. (2017). Projections of climate change impacts on Central America tropical rainforest. *Climatic Change*, *141*(1), 93–105. https://doi.org/10.1007/s10584-016-1790-2

Maezumi, S. Y., Alves, D., Robinson, M., de Souza, J. G., Levis, C., Barnett, R. L., Almeida de Oliveira, E., Urrego, D., Schaan, D., & Iriarte, J. (2018). The legacy of 4,500 years of polyculture agroforestry in the eastern Amazon. *Nature Plants*, *4*(8), 540–547. https://doi.org/10.1038/s41477-018-0205-y

Mancini, C., Kidane, Y. G., Mengistu, D. K., Pè, M. E., Fadda, C., & Dell'Acqua, M. (2017). Joining smallholder farmers' traditional knowledge with metric traits to select better varieties of Ethiopian wheat. *Scientific Reports*, *7*(1), 9120. https://doi.org/10.1038/s41598-017-07628-4

Manners, R., & van Etten, J. (2018). Are agricultural researchers working on the right crops to enable food and nutrition security under future climates? *Global Environmental Change*, *53*, 182–194. https://doi.org/10.1016/J.GLOENVCHA.2018.09.010

Marie, L., Abdallah, C., Campa, C., Courtel, P., Bordeaux, M., Navarini, L., Lonzarich, V., Bosselmann, A. S., Turreira-García, N., Alpizar, E., Georget, F., Breitler, J. C., Etienne, H., &

Bertrand, B. (2020). G × E interactions on yield and quality in Coffea arabica: new F1 hybrids outperform American cultivars. *Euphytica*, *216*(5), 1–17. https://doi.org/10.1007/s10681-020-02608-8

Mbow, C., Smith, P., Skole, D., Duguma, L., & Bustamante, M. (2014). Achieving mitigation and adaptation to climate change through sustainable agroforestry practices in Africa. *Current Opinion in Environmental Sustainability*, *6*, 8–14. https://doi.org/10.1016/J.COSUST.2013.09.002

Mbow, C., van Noordwijk, M., Luedeling, E., Neufeldt, H., Minang, P. A., & Kowero, G. (2014). *Agroforestry solutions to address food security and climate change challenges in Africa* (Nos. 0; Vol. 6, pp. 61–67). https://doi.org/10.1016/j.cosust.2013.10.014

McGuire, S., & Sperling, L. (2016). Seed systems smallholder farmers use. *Food Security*, *8*(1), 179–195. https://doi.org/10.1007/s12571-015-0528-8

Meijer, S. S., Catacutan, D., Ajayi, O. C., Sileshi, G. W., & Nieuwenhuis, M. (2015). The role of knowledge, attitudes and perceptions in the uptake of agricultural and agroforestry innovations among smallholder farmers in sub-Saharan Africa. *International Journal of Agricultural Sustainability*, *13*(1), 40–54. https://doi.org/10.1080/14735903.2014.912493

Mengistu, D. K., Kidane, Y. G., Catellani, M., Frascaroli, E., Fadda, C., Pè, M. E., & Dell'Acqua, M. (2016). High-density molecular characterization and association mapping in Ethiopian durum wheat landraces reveals high diversity and potential for wheat breeding. *Plant Biotechnology Journal*, *14*(9), 1800–1812. https://doi.org/10.1111/pbi.12538

Meyer, H., Reudenbach, C., Hengl, T., Katurji, M., & Nauss, T. (2018). Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling and Software*, *101*(2018), 1–9. https://doi.org/10.1016/j.envsoft.2017.12.001

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 21. https://doi.org/10.1038/s41562-016-0021

Nair, P. K. R. (1993). *An introduction to agroforestry* (1st ed., p. 499). Springer Netherlands. https://www.springer.com/gp/book/9780792321347

Nelson, R. (2020). Viewpoint: International agriculture's needed shift from energy intensification to agroecological intensification. *Food Policy*, *91*, 101815. https://doi.org/10.1016/j.foodpol.2019.101815

Niles, M. T., & Mueller, N. D. (2016). Farmer perceptions of climate change: Associations with observed temperature and precipitation trends, irrigation, and climate beliefs. *Global Environmental Change*, *39*, 133–142. https://doi.org/10.1016/j.gloenvcha.2016.05.002

Norgrove, L. (2018). Neither dark nor light but shades in-between: cocoa merits a finer sampling. *Global Change Biology*, *24*(2), 559–560. https://doi.org/10.1111/gcb.14012

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., & Kassem, K. R. (2001). Terrestrial ecoregions of the world: A new map of life on Earth. *BioScience*, *51*(11), 933–938. https://doi.org/10.1641/0006-3568(2001)051%5B0933:TEOTWA%5D2.0.CO;2

Ordoñez, J. C., Luedeling, E., Kindt, R., Tata, H. L., Harja, D., Jamnadass, R., & van Noordwijk, M. (2014). Constraints and opportunities for tree diversity management along the forest transition curve to achieve multifunctional agriculture. *Current Opinion in Environmental Sustainability*, *6*(0), 54–60. https://doi.org/10.1016/j.cosust.2013.10.009

Orozco, L., Somarriba, E., Villalobos, M., Deheuvels, O., & Andrade, H. (2014). Recopilacion de datos de investigacion del Proyecto Cacao Centroamerica (PCC). *Harvard Dataverse*. https://doi.org/10.7910/DVN/25284

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*(4). https://doi.org/10.1371/journal.pone.0124155

Padovan, M. P., Brook, R. M., Barrios, M., Cruz-Castillo, J. B., Vilchez-Mendoza, S. J., Costa, A. N., & Rapidel, B. (2018). Water loss by transpiration and soil evaporation in coffee shaded by Tabebuia rosea Bertol. and Simarouba glauca dc. compared to unshaded coffee in sub-optimal environmental conditions. *Agricultural and Forest Meteorology*, *248*, 1–14. https://doi.org/10.1016/j.agrformet.2017.08.036

Padovan, M. P., Cortez, V. J., Navarrete, L. F., Navarrete, E. D., Deffner, A. C., Centeno, L. G., Munguía, R., Barrios, M., Vílchez-Mendoza, J. S., Vega-Jarquín, C., Costa, A. N., Brook, R. M., & Rapidel, B. (2015). Root distribution and water use in coffee shaded with Tabebuia rosea Bertol. and Simarouba glauca DC. compared to full sun coffee in sub-optimal environmental conditions. *Agroforestry Systems*, *89*(5), 857–868. https://doi.org/10.1007/s10457-015-9820-z

Paliwal, A., & Jain, M. (2020). The accuracy of self-reported crop yield estimates and their ability to train remote sensing algorithms. *Frontiers in Sustainable Food Systems*, *4*, 25. https://doi.org/10.3389/fsufs.2020.00025

Peeters, L. Y. K., Soto-Pinto, L., Perales, H., Montoya, G., & Ishiki, M. (2003). Coffee production, timber, and firewood in traditional and Inga-shaded plantations in Southern Mexico. *Agriculture, Ecosystems & Environment*, *95*(2–3), 481–493. https://doi.org/10.1016/S0167-8809(02)00204-9

Peng, S., Huang, J., Sheehy, J. E., Laza, R. C., Visperas, R. M., Zhong, X., Centeno, G. S., Khush, G. S., & Cassman, K. G. (2004). Rice yields decline with higher night temperature from global warming. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(27), 9971–9975. https://doi.org/10.1073/pnas.0403720101

Pingali, P. L. (2012). Green revolution: Impacts, limits, and the path ahead. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(31), 12302–12308. https://doi.org/10.1073/pnas.0912953109

Plackett, R. L. (1975). The Analysis of Permutations. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *24*(2), 193–202. https://doi.org/10.2307/2346567

Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science*, *360*(6392), 987–992. https://doi.org/10.1126/science.aaq0216

Powers, S. M., & Hampton, S. E. (2019). Open science, reproducibility, and transparency in ecology. *Ecological Applications*, *29*(1), e01822. https://doi.org/10.1002/eap.1822

Prentice, I. C., Cramer, W., Harrison, S. P., Leemans, R., Monserud, R. A., & Solomon, A. M. (1992). Special Paper: A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography*, *19*(2), 117. https://doi.org/10.2307/2845499

Pruvot-Woehl, S., Krishnan, S., Solano, W., Schilling, T., Toniutti, L., Bertrand, B., & Montagnon, C. (2020). Authentication of <i> Coffea arabica <ı> varieties through DNA fingerprinting and its significance for the coffee sector. *Journal of AOAC INTERNATIONAL*, *103*(2), 325–334. https://doi.org/10.1093/jaocint/qsz003

Rainey, K. M., & Griffiths, P. D. (2005). Inheritance of heat tolerance during reproductive development in snap bean (Phaseolus vulgaris L.). *Journal of the American Society for Horticultural Science*, *130*(5), 700–706. https://doi.org/10.21273/jashs.130.5.700

Ramirez-Villegas, J., Khoury, C. K., Achicanoy, H. A., Mendez, A. C., Diaz, M. V., Sosa, C. C., Debouck, D. G., Kehel, Z., & Guarino, L. (2020). A gap analysis modelling framework to prioritize collecting for ex situ conservation of crop landraces. *Diversity and Distributions*, *26*(6), 730–742. https://doi.org/10.1111/ddi.13046

Ramírez, O. A., Somarriba, E., Ludewigs, T., & Ferreira, P. (2001). *Financial returns, stability and risk of cacao-plantain-timber agroforestry systems in Central America* (Nos. 2; Vol. 51, pp. 141–154). Kluwer Academic Publishers. https://doi.org/10.1023/A:1010655304724

Ranjitkar, S., Kindt, R., Sujakhu, N. M., Hart, R., Guo, W., Yang, X., Shrestha, K. K., Xu, J., & Luedeling, E. (2014). *Separation of the bioclimatic spaces of Himalayan tree rhododendron species predicted by ensemble suitability models* (Vol. 1, pp. 2–12). https://doi.org/10.1016/j.gecco.2014.07.001

Rao, B. B., Chowdary, P. S., Sandeep, V. M., Pramod, V. P., & Rao, V. U. M. (2015). Spatial analysis of the sensitivity of wheat yields to temperature in India. *Agricultural and Forest Meteorology*, *200*, 192–202. https://doi.org/10.1016/j.agrformet.2014.09.023

Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., & Foley, J. A. (2012). Recent patterns of crop yield growth and stagnation. *Nature Communications*, *3*. https://doi.org/10.1038/ncomms2296

R Core Team. (2020). *R: A language and environment for statistical computing. version 4.0.2*. CRAN R Project. https://r-project.org/

Renteria, N., & Rowling, M. (2016). Hit by climate change, Central American coffee growers get a taste for cocoa. *Thomson Reuters*. https://www.reuters.com/article/us-centralamerica-climatechange-coffee-c/hit-by-climate-change-central-american-coffee-growers-get-a-taste-for-cocoa-idUSKCN10Z0VX

Rhoades, R. E., & Booth, R. H. (1982). Farmer-back-to-farmer: A model for generating acceptable agricultural technology. *Agricultural Administration*, *11*(2), 127–137. https://doi.org/10.1016/0309-586X(82)90056-5

Ricciardi, V., Ramankutty, N., Mehrabi, Z., Jarvis, L., & Chookolingo, B. (2018). How much of the world's food do smallholders produce? *Global Food Security*, *17*, 64–72. https://doi.org/10.1016/j.gfs.2018.05.002

Sanchez, P. A. (2020). Viewpoint: Time to Increase Production of Nutrient-rich Foods. *Food Policy*, *91*, 101843. https://doi.org/10.1016/j.foodpol.2020.101843

Sauvadet, M., Asare, R., & Isaac, M. E. (2020). Evolutionary distance explains shade tree selection in agroforestry systems. *Agriculture, Ecosystems and Environment*, *304*, 107125. https://doi.org/10.1016/j.agee.2020.107125

Schattman, R. E., Conner, D., & Méndez, V. E. (2016). Farmer perceptions of climate change risk and associated on-farm management strategies in Vermont, northeastern United States. *Elementa: Science of the Anthropocene*, *4*(0), 000131. https://doi.org/10.12952/journal.elementa.000131

Schnabel, F., de Melo Virginio Filho, E., Xu, S., Fisk, I. D., Roupsard, O., & Haggar, J. (2017). Shade trees: a determinant to the relative success of organic versus conventional coffee production. *Agroforestry Systems*, *First Online*, 1–15. https://doi.org/10.1007/s10457-017-0100-y

Schroth, G., Läderach, P., Blackburn Cuero, D. S., Neilson, J., & Bunn, C. (2015). Winner or loser of climate change? A modeling study of current and future climatic suitability of Arabica coffee in Indonesia. *Regional Environmental Change*, *15*(7), 1473–1482. https://doi.org/10.1007/s10113-014-0713-x

Schroth, G., Läderach, P., Martinez-Valle, A. I., Bunn, C., & Jassogne, L. (2016). Vulnerability to climate change of cocoa in West Africa: Patterns, opportunities and limits to adaptation. *Science of the Total Environment*, *556*, 231–241. https://doi.org/10.1016/j.scitotenv.2016.03.024

Senyolo, M. P., Long, T. B., Blok, V., & Omta, O. (2018). How the characteristics of innovations impact their adoption: An exploration of climate-smart agricultural innovations in South

Africa. *Journal of Cleaner Production*, *172*, 3825–3840. https://doi.org/10.1016/J.JCLEPRO.2017.06.019

Sepulveda, N., & Barrios, M. (2016). Ensayo de sistemas agroforestales con cafe, Masatepe, Nicaragua. *Harvard Dataverse*. https://doi.org/10.7910/DVN/27915

Singh, R. K., Zander, K. K., Kumar, S., Singh, A., Sheoran, P., Kumar, A., Hussain, S. M., Riba, T., Rallen, O., Lego, Y. J., Padung, E., & Garnett, S. T. (2017). Perceptions of climate variability and livelihood adaptations relating to gender and wealth among the Adi community of the Eastern Indian Himalayas. *Applied Geography*, *86*, 41–52. https://doi.org/10.1016/J.APGEOG.2017.06.018

Somarriba, E., Cerda, R., Orozco, L., Cifuentes, M., Dávila, H., Espin, T., Mavisoy, H., Ávila, G., Alvarado, E., Poveda, V., Astorga, C., Say, E., & Deheuvels, O. (2013). Carbon stocks and cocoa yields in agroforestry systems of Central America. *Agriculture, Ecosystems and Environment*, *173*(2013), 46–57. https://doi.org/10.1016/j.agee.2013.04.013

Sparks, A. H. (2018). nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R. *Journal of Open Source Software*, *3*(30), 1035. https://doi.org/10.21105/joss.01035

Spurgeon, D. (1979). Agroforestry: new hope for subsistence farmers. *Nature*, *280*, 533–534. https://doi.org/10.1038/280533a0

Steenwerth, K. L., Hodson, A. K., Bloom, A. J., Carter, M. R., Cattaneo, A., Chartres, C. J., Hatfield, J. L., Henry, K., Hopmans, J. W., Horwath, W. R., Jenkins, B. M., Kebreab, E., Leemans, R., Lipper, L., Lubell, M. N., Msangi, S., Prabhu, R., Reynolds, M. P., Sandoval Solis, S., … Sykes, M. (2014). Climate-smart agriculture global research agenda: scientific basis for action. *Agriculture & Food Security*, *3*(1), 11. https://doi.org/10.1186/2048-7010-3-11

Steinke, J., van Etten, J., & Zelan, P. M. (2017). The accuracy of farmer-generated data in an agricultural citizen science methodology. *Agronomy for Sustainable Development*, *37*(4), 32. https://doi.org/10.1007/s13593-017-0441-y

Stevens, J. R. (2017). Replicability and Reproducibility in Comparative Psychology. *Frontiers in Psychology*, *8*, 862. https://doi.org/10.3389/fpsyg.2017.00862

Strobl, C., Wickelmaier, F., & Zeileis, A. (2011). Accounting for individual differences in Bradley-Terry models by means of recursive partitioning. *Journal of Educational and Behavioral Statistics*, *36*(2), 135–153. https://doi.org/10.3102/1076998609359791

Suárez Salazar, J. C., Melgarejo, L. M., Casanoves, F., Di Rienzo, J. A., DaMatta, F. M., & Armas, C. (2018). Photosynthesis limitations in cacao leaves under different agroforestry systems in the Colombian Amazon. *PLOS ONE*, *13*(11), e0206149. https://doi.org/10.1371/journal.pone.0206149

Tester, M., & Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science*, *327*(5967), 818–822. https://doi.org/10.1126/science.1183700

Thornton, P., Aggarwal, P., & Parsons, D. (2017). Editorial: Prioritising climate-smart agricultural interventions at different scales. *Agricultural Systems*, *151*, 149–152. https://doi.org/10.1016/j.agsy.2016.12.007

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(50), 20260–20264. https://doi.org/10.1073/pnas.1116437108

Tollenaar, M., Fridgen, J., Tyagi, P., Stackhouse Jr, P. W., & Kumudini, S. (2017). The contribution of solar brightening to the US maize yield trend. *Nature Clim. Change*, *7*(4), 275–278. https://doi.org/10.1038/nclimate3234

Torres, C. M. M. E., Jacovine, L. A. G., Nolasco de Olivera Neto, S., Fraisse, C. W., Soares, C. P. B., de Castro Neto, F., Ferreira, L. R., Zanuncio, J. C., & Lemes, P. G. (2017). Greenhouse gas emissions and carbon sequestration by agroforestry systems in southeastern Brazil. *Scientific Reports*, *7*(1), 16738. https://doi.org/10.1038/s41598-017-16821-4

Trnka, M., Rötter, R. P., Ruiz-Ramos, M., Kersebaum, K. C., Olesen, J. E., Žalud, Z., & Semenov, M. A. (2014). Adverse weather conditions for European wheat production will become more frequent with climate change. *Nature Climate Change*, *4*(7), 637–643. https://doi.org/10.1038/nclimate2242

Turner, H. L., van Etten, J., Firth, D., & Kosmidis, I. (2020). Modelling rankings in R: the PlackettLuce package. *Computational Statistics*, *2020*(35), 1027–1057. https://doi.org/10.1007/s00180-020-00959-3

van Eeuwijk, F. A., Cooper, M., DeLacy, I. H., Ceccarelli, S., & Grando, S. (2001). Some vocabulary and grammar for the analysis of multi-environment trials, as applied to the analysis of FPB and PPB trials. *Euphytica*, *122*(3), 477–490. https://doi.org/10.1023/A:1017591407285

van Etten, J., Beza, E., Calderer, L., van Duijvendijk, K., Fadda, C., Fantahun, B., Kidane, Y. G., van de Gevel, J., Gupta, A., Mengistu, D. K., Kiambi, D., Mathur, P. N., Mercado, L., Mittra, S., Mollel, M. J., Rosas, J. C., Steinke, J., Suchini, J. G., & Zimmerer, K. S. (2019). First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Experimental Agriculture*, *55*(S1), 275–296. https://doi.org/10.1017/S0014479716000739

van Etten, J., de Sousa, K., Aguilar, A., Barrios, M., Coto, A., Dell'Acqua, M., Fadda, C., Gebrehawaryat, Y., van de Gevel, J., Gupta, A., Kiros, A. Y., Madriz, B., Mathur, P., Mengistu, D. K., Mercado, L., Nurhisen Mohammed, J., Paliwal, A., Pè, M. E., Quirós, C. F., . . . Steinke, J. (2018). Replication data for: "Crop variety management for climate adaptation supported by citizen science". *Harvard Dataverse*. https://doi.org/10.7910/DVN/4ICF6W van Etten, J., de Sousa, K., Aguilar, A., Barrios, M., Coto, A., Dell'Acqua, M., Fadda, C., Gebrehawaryat, Y., van de Gevel, J., Gupta, A., Kiros, A. Y., Madriz, B., Mathur, P., Mengistu, D. K., Mercado, L., Nurhisen Mohammed, J., Paliwal, A., Pè, M. E., Quirós, C. F., . . . Steinke, J. (2019). Crop variety management for climate adaptation supported by citizen science. *Proceedings of the National Academy of Sciences*, *116*(10), 4194–4199. https://doi.org/10.1073/pnas.1813720116

van Valkengoed, A. M., & Steg, L. (2019). Meta-analyses of factors motivating climate change adaptation behaviour. *Nature Climate Change*, *9*(2), 158–163. https://doi.org/10.1038/s41558-018-0371-y

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., & Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change*, *109*(1), 5–31. https://doi.org/10.1007/s10584-011-0148-z

van Zonneveld, M., Turmel, M.-S., & Hellin, J. (2020). Decision-making to diversify farm systems for climate change adaptation. *Frontiers in Sustainable Food Systems*, *4*, 32. https://doi.org/10.3389/FSUFS.2020.00032

Vavilov, N. I. (1992). *Origin and geography of cultivated plants* (p. 498). Cambridge University Press.

Verchot, L. V., van Noordwijk, M., Kandji, S., Tomich, T., Ong, C., Albrecht, A., Mackensen, J., Bantilan, C., Anupama, K. V., & Palm, C. (2007). Climate change: linking adaptation and mitigation through agroforestry. *Mitigation and Adaptation Strategies for Global Change*, *12*(5), 901–918. https://doi.org/10.1007/s11027-007-9105-6

Vermeulen, S. J., Aggarwal, P. K., Ainslie, A., Angelone, C., Campbell, B. M., Challinor, A. J., Hansen, J. W., Ingram, J. S. I., Jarvis, A., Kristjanson, P., Lau, C., Nelson, G. C., Thornton, P. K., & Wollenberg, E. (2012). Options for support to agriculture and food security under climate change. *Environmental Science and Policy*, *15*(1), 136–144. https://doi.org/10.1016/j.envsci.2011.09.003

Vermeulen, S. J., Challinor, A. J., Thornton, P. K., Campbell, B. M., Eriyagama, N., Vervoort, J. M., Kinyangi, J., Jarvis, A., Läderach, P., Ramirez-Villegas, J., Nicklin, K. J., Hawkins, E., & Smith, D. R. (2013). Addressing uncertainty in adaptation planning for agriculture. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(21), 8357–8362. https://doi.org/10.1073/pnas.1219441110

Villegas, D., Alfaro, C., Ammar, K., Cátedra, M. M., Crossa, J., García del Moral, L. F., & Royo, C. (2016). Daylength, Temperature and Solar Radiation Effects on the Phenology and Yield Formation of Spring Durum Wheat. *Journal of Agronomy and Crop Science*, *202*(3), 203–216. https://doi.org/10.1111/jac.12146

Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic*

Bulletin and Review, *11*(1), 192–196. https://doi.org/10.3758/BF03206482

Wan, Z., Hook, S., & Hulley, G. (2015). MYD11A1 MODIS/Aqua Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006 [Data set]. *NASA EOSDIS LP DAAC*. https://doi.org/10.5067/MODIS/MYD11A1.006

Whitlock, M. C. (2005). Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *Journal of Evolutionary Biology*, *18*(5), 1368–1373. https://doi.org/10.1111/j.1420-9101.2005.00917.x

Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., Jonell, M., Clark, M., Gordon, L. J., Fanzo, J., Hawkes, C., Zurayk, R., Rivera, J. A., De Vries, W., Majele Sibanda, L., … Murray, C. J. L. (2019). Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *The Lancet*, *393*(10170), 447–492. https://doi.org/10.1016/S0140-6736(18)31788-4

Zeileis, A., Hothorn, T., & Hornik, K. (2008). Model-based recursive partitioning. *Journal of Computational and Graphical Statistics*, *17*(2), 492–514. https://doi.org/10.1198/106186008X319331

Zeven, A. C. (1998). Landraces: A review of definitions and classifications. *Euphytica*, *104*(2), 127–139. https://doi.org/10.1023/A:1018683119237

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., Huang, M., Yao, Y., Bassu, S., Ciais, P., Durand, J.-L., Elliott, J., Ewert, F., Janssens, I. A., Li, T., Lin, E., Liu, Q., Martre, P., Müller, C., … Asseng, S. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(35), 9326–9331. https://doi.org/10.1073/pnas.1701762114

Zimmerer, K. S., & De Haan, S. (2017). Agrobiodiversity and a sustainable food future. *Nature Plants*, *3*(4), 17047. https://doi.org/10.1038/nplants.2017.47

Zohner, C. M., Mo, L., Renner, S. S., Svenning, J.-c., Vitasse, Y., Benito, B. M., Ordonez, A., Baumgarten, F., Bastin, J.-F., Sebald, V., Reich, P. B., Liang, J., Nabuurs, G.-J., De-Miguel, S., Alberti, G., Antón-Fernández, C., Balazy, R., Brändli, U.-B., Chen, H. Y. H., … Crowther, T. W. (2020). Late-spring frost risk between 1959 and 2017 decreased in North America but increased in Europe and Asia. *Proceedings of the National Academy of Sciences*, *117*(22), 201920816. https://doi.org/10.1073/pnas.1920816117

1

Overview of the agricultural landscape in Copán, Honduras. Credit: K. de Sousa

Journal of Rural Studies 64 (2018) 11–19

Contents lists available at ScienceDirect

Journal of Rural Studies

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

How climate awareness influences farmers' adaptation decisions in Central America?

Kauê de Sousa[a,b,c,*], Fernando Casanoves[d], Jorge Sellare[e], Alejandra Ospina[d], Jose Gabriel Suchini[f], Amilcar Aguilar[g], Leida Mercado[a,**]

[a] Tropical Agriculture and Higher Education Center (CATIE), Research Program in Development, Economy and Environment, Turrialba, Costa Rica
[b] Inland Norway University of Applied Sciences, Department of Agricultural Sciences, Hamar, Norway
[c] Bioversity International, Costa Rica Office, Turrialba, Costa Rica
[d] Tropical Agriculture and Higher Education Center (CATIE), Biostatistics Unit, Turrialba, Costa Rica
[e] University of Goettingen, Department of Agricultural Economics and Rural Development, Goettingen, Germany
[f] Tropical Agriculture and Higher Education Center (CATIE), Guatemala Technical Office, Esquipulas, Guatemala
[g] Tropical Agriculture and Higher Education Center (CATIE), Nicaragua Technical Office, Managua, Nicaragua

ARTICLE INFO

Keywords:
Adaptation to climate change
Bradley-Terry model
Central America dry corridor
Climate change
Climate-smart agriculture
Farmer field schools
Reforestation
Smallholder

ABSTRACT

Central America is one of the regions with the highest vulnerability to climate change, with negative effects projected to affect its economy and food security. To address this issue, an integrative farm management approach such as Climate-Smart Agriculture can help reorient agricultural practices towards climate adaptation and food security. Past studies have shown that several factors can either hinder or encourage the adoptions of Climate-Smart practices, including subjective expectations and perceptions. Building on this literature, we analyze farmers' climate awareness and their perceptions regarding the change in climate patterns as well as their choices of farming practices to adapt to these changes. We show that reforestation was the preferred adaptation strategy among interviewed farmers and that educational profiles and the size of landholdings drive the adoption of this and other practices. Soil management and introduction of new crops are preferred by literate farms with large farmlands, whereas illiterate farmers with smaller farmland tend to move towards farm intensification with an increase in the utilization of external inputs. Our findings provide evidence to support the design of capacity development interventions targeting specific groups of farmers according to their main crop and education profile.

1. Introduction

Trends in greenhouse gases emissions to 2050 indicate a low contribution of Central America to global warming (Marchal et al., 2011), and yet the region is highly vulnerable to the effects of climate change. Several climate-related impacts have been projected for the region, indicating changes in evapotranspiration, temperature, precipitation, species suitability, farm productivity, and forest loss, mainly across the drier zones (Hannah et al., 2017; Lyra et al., 2017). Therefore, promoting farm practices to strengthen resilience and productivity of agricultural systems is crucial to help farmers in Central America adapt to climate change and thus ensure food provision and income generation.

Climate change has increased the risks and uncertainties associated with agriculture, particularly in developing countries (Altieri and Nicholls, 2017; Imbach et al., 2017). Changes in the frequency and intensity of extreme climatic events in the tropics due to climate change have increased the concerns for farm adaptation among scientists (Hannah et al., 2017; Harvey et al., 2014; Mbow et al., 2014) and farmers (Elum et al., 2017; Khatri-Chhetri et al., 2017; Singh et al., 2017). It is argued that the adoption of Climate-Smart Agriculture (CSA) practices will help vulnerable farmers cope with the effects of climate variability and change (Lipper et al., 2014; Steenwerth et al., 2014). Climate-Smart Agriculture is an integrative approach designed to help farmers reorient their agricultural practices to sustainably rise agricultural productivity to ensure increases in farm incomes and food security, while adapting and mitigating climate change. These practices include farm sustainable intensification and diversification of

* Corresponding author. Inland Norway University of Applied Sciences, Department of Agricultural Sciences, Postboks 400, 2418 Elverum, Norway.
** Corresponding author. Tropical Agricultural Research and Higher Education Center, Research Program in Development, Economy and Environment, CATIE, Sede Central, Turrialba, 30501, Costa Rica.
E-mail addresses: k.desousa@cgiar.org, kaue.desousa@inn.no (K. de Sousa), lmercado@catie.ac.cr (L. Mercado).

https://doi.org/10.1016/j.jrurstud.2018.09.018
Received 21 March 2018; Received in revised form 30 August 2018; Accepted 29 September 2018
Available online 08 October 2018

production, agroforestry, varietal selection, plant breeding, ecosystem management, crop patterns identification, and integrated practices to minimize the need of external inputs (FAO, 2010).

The adoption and impact of agricultural practices and technologies has been a focus of study for several years (see Mwangi and Kariuki (2015), for a literature review on adoption, and Ogundari and Bolarinwa (2018), for a recent meta-analysis on the impacts of agricultural technologies). The literature shows that the adoption of technologies by smallholder farmers mostly has a positive effect on welfare and production outcomes, and that adopting technology packages as opposed to individual components can further increase these benefits (Khonje et al., 2018).

Nevertheless, several socio-economic barriers can hinder technology adoption, even in countries that enjoy higher levels of technological innovation and well-established institutions (Long et al., 2016). The presence of certain policies, such as input subsidies (Koppmair et al., 2017), and technology specific characteristics (Senyolo et al., 2018; Wassie and Pauline, 2018) can also influence whether and which technologies farmers adopt. Likewise, intrinsic factors, such as perceptions and knowledge of farmers, play a role on shaping technology adoption (Meijer et al., 2015).

One strain of this body of literature on technology adoption uses the theory of planned behavior (Ajzen, 1991) to understand how perceptions and other underlying psychological constructs affect technology adoption. In a study about the adoption of improved natural grassland in Brazil, Borges et al. (2014) find that farmers' expectations about the benefits of this new technology, their perceptions about social pressure, and their perceptions about their own skills are significantly correlated with the intention to adopt. Similarly, Wauters et al. (2010) show that attitudes towards soil conservation practices are one of the biggest determinants of adoption among Belgium farmers. Regarding sustainable agricultural practices for climate adaptation, several studies conclude farmers' awareness and perceptions of climate change are correlated with adoption (Elum et al., 2017; Niles and Mueller, 2016; Schattman et al., 2016; Singh et al., 2017).

Building on this body of literature, the objective of this study is to understand how farmers' awareness of climate change and their socioeconomic profiles drive the utilization of sustainable farm management practices in Central America. We assess farmers' climate awareness by identifying farmers' perceptions of climate variability and compare it with observed climate anomalies using time series data. Additionally, we implement a Bradley-Terry model to assess how socioeconomic profiles and farm characteristics influence farmers' choices in the adoption of sustainable agriculture practices.

2. Materials and methods

2.1. Study area and household data

We used surveyed data from 283 households participating in the Mesoamerican Environmental Program (MAP), a rural development program conducted in Central America between 2009 and 2017 that used Farmer Field Schools (FSS) to promote CSA practices and gender integration (see Gutierrez-Montes et al. (2018), for details on the methodology applied in the FFS). We used two sets of data: *(i)* a household survey on farmer's perceptions on climate change (Appendix A), and *(ii)* household socioeconomic data and information records of practices adopted by the farmers after participating in FFS obtained from MAP's annual monitoring.

Farmers were located across the two main ecoregions of Central America (Fig. 1): the Central American Dry Corridor (or Dry Forests), corresponding to El Salvador, Guatemala, Honduras, and part of Nicaragua (districts of Jinotega and Matagalpa); and the Central American Rainforests in Nicaragua (districts of Jinotega, Matagalpa, and Atlántico Norte). Farms across the Dry Corridor have an annual average precipitation of 1400 mm (1000–2100 mm), mean annual temperature of 22 °C (14–25 °C) and mean elevation of 750 m a.s.l. (300–1950 m a.s.l.). Farms across the Rainforests present annual average precipitation of 2200 mm (1500–2400 mm), mean annual temperature of 22 °C (19–25 °C) and mean elevation of 570 m a.s.l. (240–1200 m a.s.l.) (Hijmans et al., 2005). Agricultural and livestock production are the main economic activities developed across the research sites.

Precipitation is key for determining the crop seasons in Central America, especially for the annual crops. The first growing season, called *Primera*, starts in May and ends in September, when the second season (*Postrera*) begins. The last growing season, *Apante*, starts in November and ends in January. This season presents a gradual decrease in rainfall until the beginning of the dry season (*Verano*) in January (Fig. 2).

To collect the household data, in 2014, we applied a questionnaire to identify the perceptions of farmers regarding changes in climatic patterns and how they responded to these events in terms of farm management practices. Farmers were questioned about their perceptions regarding changes in precipitation and temperature over the 10 years before the interviews (2005–2014). Farmers who reported to have felt changes in climatic patterns were asked to list the farm management practices they have adopted in their crop systems to cope with such changes. These practices were ranked by the order they were mentioned by the farmers. In Table 1 we show descriptive statistics of the socioeconomic data from the 283 households disaggregated by ecoregion.

2.2. Retrieving environmental data to validate farmers' perceptions

We took farmers' perceptions of changes in climatic patterns and compared them to a gridded time series precipitation database from the Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) (Funk et al., 2015). This database incorporates global daily rainfall data since 1983 with a resolution of 2.5 arc-min ($\sim 5\,km^2$), which is obtained by weather stations and combined with remote sensing. Changes in precipitation were assessed by calculating three extreme precipitation indices relevant for Central America (Aguilar et al., 2005): *(i)* SDII, simple daily intensity index (precipitation amount/rainy days ≥ 1 mm); *(ii)* Rx5day, maximum 5-day precipitation (days); and *(iii)* MLDS, maximum length of consecutive dry days (< 1 mm). Information on temperature was not assessed due to the lack of consistent high-resolution time series data for Central America. We performed a multiple correspondence analysis for quantitative and categorical variables (Lê et al., 2008) to identify the association of observed changes in precipitation (based on CHIRPS data) and farmers' perceptions.

2.3. Ranking farmers' strategies to cope with climate variability

We analyzed the strategies each farmer claimed to have adopted to cope with perceived changes in climate patterns by using a Bradley-Terry model (Bradley and Terry, 1952; Turner and Firth, 2012) to create partial ranks of 5 (the five first strategies mentioned by each farmer). The Bradley-Terry model estimates the "worth parameter" or the relative importance of the different strategies in pairwise comparisons and, under the Model-Based Recursive Partitioning approach, identifies sub-groups of farms with similar choices (Hothorn and Zeileis, 2015; Strobl et al., 2011).

We added six variables to the splitting algorithm: *(i)* the ecoregion (Dry or Rainforest), *(ii)* the Progress Out of Poverty Index (PPI), *(iii)* the literacy level of the head of household, *(iv)* the area of the main crop system (ha), *(v)* the age of the head of household, and *(vi)* the number of practices adopted by the farmers after participating in the FFS. Under this approach, if the difference in chosen strategies was significant ($\alpha < 0.05$), then the model would create different groups. Based on practices reported by farmers, we ranked 10 options: *(i)* Change in Agricultural Calendar, *(ii)* Change in Varieties, *(iii)* Production

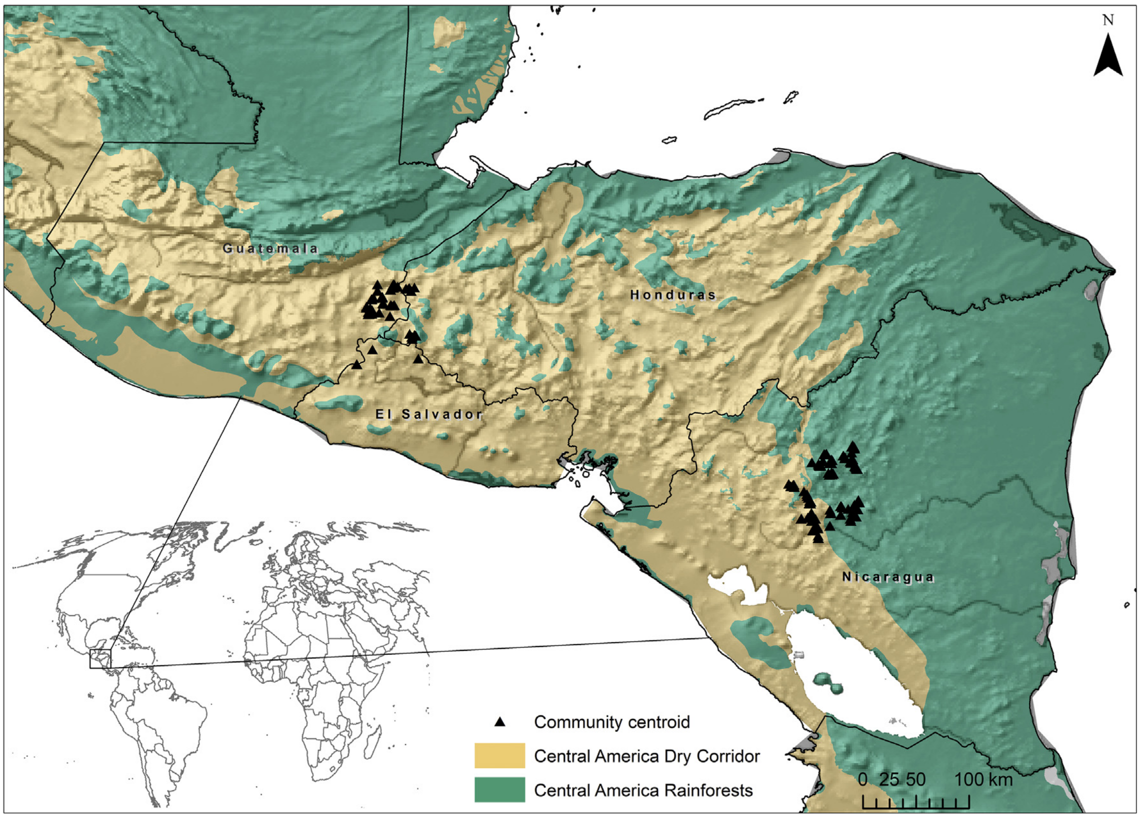

Fig. 1. Research sites across Central America.

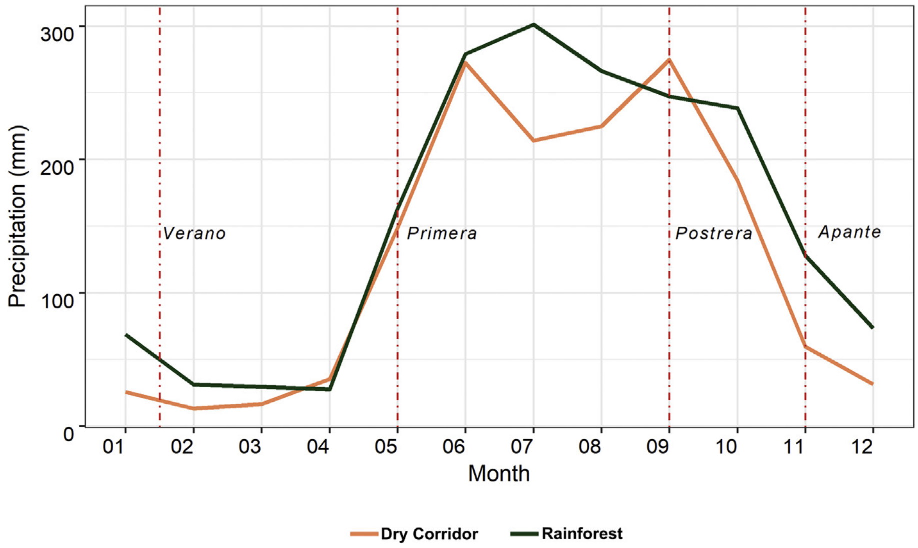

Fig. 2. Average monthly precipitation between 1891 and 2016 per crop season across the research sites in Central America.

Diversification, *(iv)* Introduction of New Crops, *(v)* Less Fertilizers and Pesticides, *(vi)* Reforestation and Restoration, *(vii)* Sustainable Soil Management, *(viii)* Sustainable Water Management, *(ix)* Leave Farming System, and *(x)* More Fertilizers and Pesticides. These practices vary in terms of effort, costs, and information level required for its implementation (for details see FAO (2013)). We used *Production Diversification* as a reference in the Bradley-Terry model, since this is one of the main strategies to reduce risks of food insecurity and climate vulnerability (Campbell et al., 2016). Finally, the likelihood of farmers using these practices was assessed by analyzing the relationship of the farmers' main crop system and their list of reported practices (Theus and Urbanek, 2008).

3. Results

3.1. Farmers perceived changes in precipitation with some accuracy

From the group of 283 interviewed farmers, 255 (90%) felt changes in climate patterns over the 10 years prior to the survey (2005–2014). Trends during this period in the precipitation time series data show statistical differences in all three precipitation indices used in this analysis. The frequency of heavy precipitation in Rx5day was progressively reduced over the period of 2005–2014 across both ecoregions (Fig. 3). The negative anomaly (historical mean minus year mean) in Rx5day is seen in most of the observed years, with significant

Table 1
Socioeconomic characteristics of interviewed households by ecoregion.

Variables	Dry Corridor		Rainforests	
	Mean	S.D.	Mean	S.D.
Age of the HH head	51.69	13.19	50.89	12.85
Level of education of the HH head				
Illiterate (1/0)	0.320		0.280	
Primary school (1/0)	0.600		0.700	
Secondary school (1/0)	0.080		0.020	
Number of HH members above 60 years	1.490	0.570	1.380	0.490
Number of HH members between 15 and 60 years	3.880	1.950	3.810	1.850
Number of HH members between 5 and 15 years	1.910	0.910	2.040	1.080
Production diversity*	2.760	1.060	4.510	1.610
PPI**	37.67	16.20	36.63	15.54
Farm area (ha)	5.380	12.05	10.17	12.13
Area of main system (ha)	5.640	55.40	1.070	0.830
N	159		124	

Note: HH, household. *Number of crops cultivated in the farmland. **PPI, Progress Out of Poverty Index.

decreases in the Rainforests. The daily precipitation intensity (SDII) shows important changes across the Rainforests, with no significant changes across the Dry Corridor. This index also indicates strong negative anomalies in the Rainforests, mainly in 2014. Both ecoregions had gradual increment on the length of consecutive dry days (MLDS), with significant changes occurring in the Rainforests (Fig. 3).

The multiple correspondence analysis of farmers' perceptions versus observed anomalies shows partial correlations between farmers' perceptions and observed time series data (Fig. 3). Farmers who perceived uncertainty regarding the start/end of the rainy season correlate with observed decrease in heavy precipitation (Rx5day), decrease in daily precipitation intensity (SDII), and increase of the length of consecutive dry days (MLDS). Farmers who perceived less annual precipitation correlate with observed increase in SDII and Rx5day. Finally, those who perceived more precipitation or heavy precipitation are not correlated with any of the observed changes from the time series data (Fig. 4).

3.2. Socioeconomic factors led to the utilization of new practices

The worth estimates for ranked practices from the Bradley-Terry model show significant differences between practices employed to adapt with perceived changes in climatic patterns across the research sites (Table 2). Worth estimates for *Reforestation and Restoration*, *Introduction of New Crops*, and *Sustainable Soil Management* are significantly higher than the reference *Production Diversification*. The other practices are ranked below the reference, with *Leave Farming System* and *Change Agricultural Calendar* on the bottom of ranked practices to cope with perceived changes in in climatic patterns (Table 2).

The recursive partitioning algorithm split the data in four sub-groups by the following variables: ecoregion, literacy level and farm area (Fig. 5). Overall, *Reforestation and Restoration* was the first choice in the four sub-groups. The first group includes those farmers living in the Dry Corridor, illiterates and with farm area ≤0.5 ha. Additionally to reforestation, farmers from this sub-group chose practices such as *Sustainable Soil Management*, *Introduction of New Crops*, *Use of More Fertilizers and Pesticides* and *Production Diversification* as the main practices to respond to the effects of perceived climate variability.

The second splitting group comprises the farmers living in the Dry Corridor, illiterates and with farm area > 0.5 ha. In this sub-group, the main chosen practices were *Sustainable Soil Management*, *Leave Farming System*, and *Use of Less Fertilizers and Pesticides*. In the third sub-group, we identify literate farmers (primary or secondary degree) living in the Dry Corridor who chose, additional to reforestation, the *Introduction of New Crops*, *Sustainable Soil Management* and *Production Diversification.* Farmers living in the Rainforests corresponds to the fourth sub-group

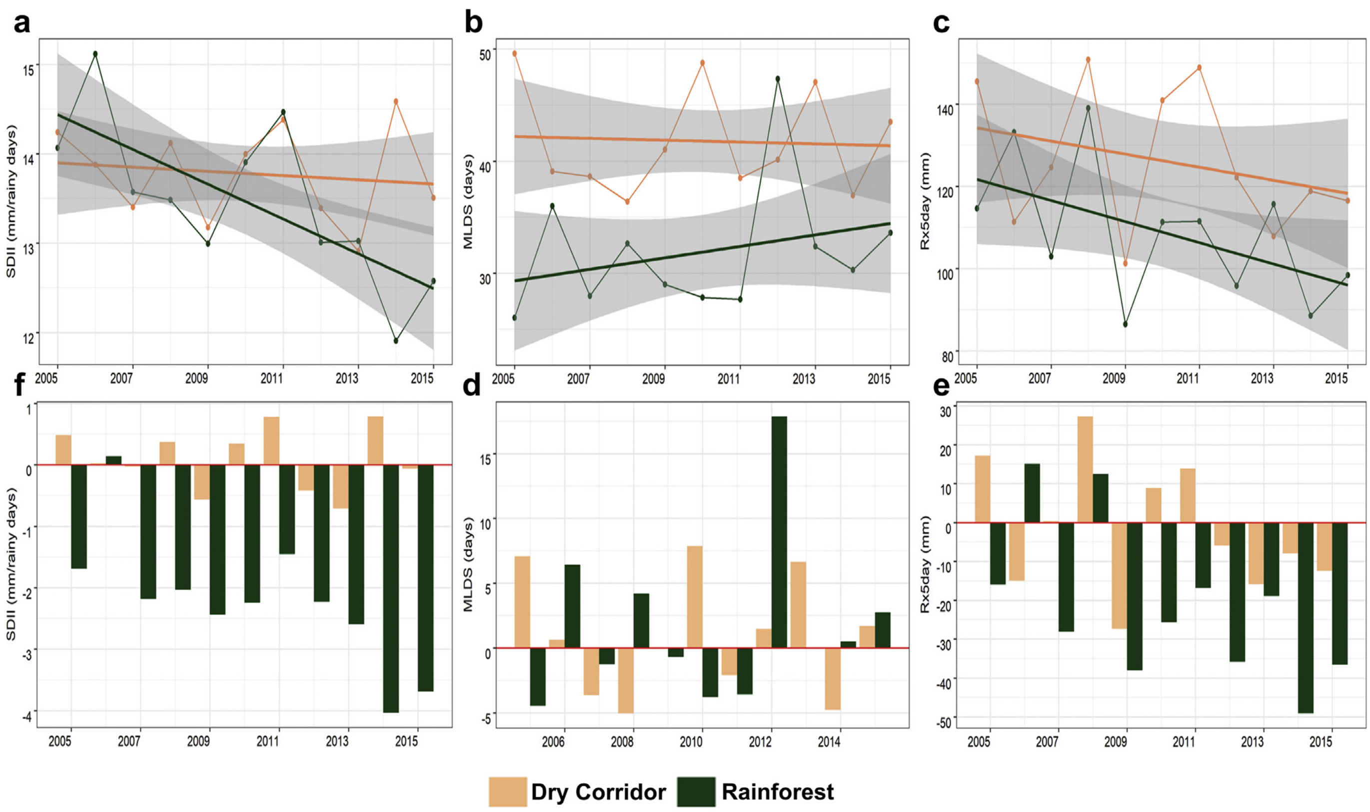

Fig. 3. Trends in precipitation indices (a, b, c) and anomaly (d, e, f) from 2005 to 2014 across the Central America Dry Corridor and Rainforests. SDII, simple annual precipitation index (mm/rainy days); Rx5day, maximum 5-day precipitation (mm); MLDS, maximum length of consecutive dry days (< 1 mm).

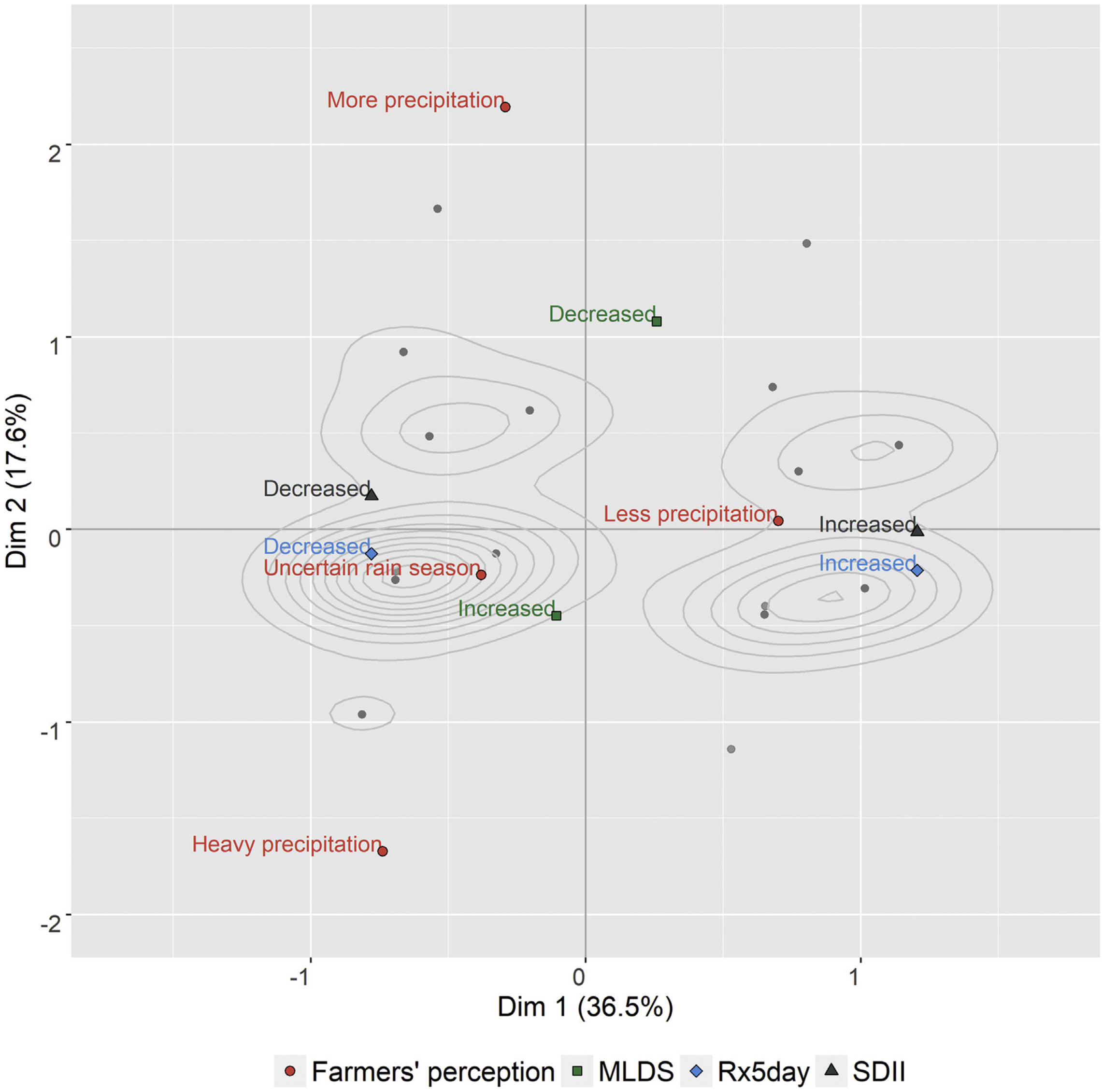

Fig. 4. Correspondence between farmers' perception on changes in precipitation and observed anomalies in precipitation indices over 2005–2014 across the Central America Dry Corridor and Rainforests. MLDS, maximum length of consecutive dry days (< 1 mm); Rx5day, maximum 5-day precipitation (mm); SDII, simple annual precipitation index (mm/rainy days).

whose preferred practices for climate adaptation were *Introduction of New Crops* and *Change Varieties*.

3.3. Choices in practices influenced by the type of crop system

The type of farming system also influenced how farmers chose to adapt to changes in perceived climate patterns. Interviewed cocoa growers showed higher likelihood to use *Change in Agricultural Calendar*, *Introduction of New Crops,* and *Leave Farming System*, as well as a lower likelihood to implement *Sustainable Soil Management* and *Use of Less Fertilizer and Pesticides*. Similarly, farmers who cultivate fruit trees have a higher likelihood to use *Production Diversification* and *Reforestation and Restoration*. On the other hand, livestock farmers are likely to use *Change in Varieties* (livestock grass varieties) and less likely to adopt *Sustainable Practices for Soils and Water Management*. Farmers whose main crop system is vegetables show a higher likelihood to use *Sustainable Soil and Water Management* and *Less Fertilizers and Pesticides*, with low preferences for *Reforestation and Restoration*, *Production Diversification,* and *Change in Varieties* (Fig. 6).

4. Discussion

We show that Central American farmers are aware of the change in climate patterns caused by climate change, with partial correlations between farmers' perceptions and the historical precipitation data. These partial correlations may be explained by the difficulty to properly observe the changes as they occur without the aid of measuring devices (e.g. weather station, garden moisture meter) or without up-to-date weather information from other sources. However, even if farmers do not perfectly perceive these changes in climate patterns, they do observe reductions in their yields and at times losses of their crops, which draws their attention to climate-related problems and increases their willingness to innovate and try new farm management practices.

Reforestation was the preferred choice among farmers independent of education profiles, farm size, and ecoregion. This practice is advocated as the best way to cope with the effects of climate change, since it includes both mitigation and adaptation by providing carbon sink, microclimate regulation and protection to extreme climate events (Caudill et al., 2015; Locatelli et al., 2015; Torres et al., 2017). Farmers demonstrated high willingness to adopt reforestation despite low governmental incentives, which often can act as disincentives given the

Table 2
Model estimates from farmers' management practices employed to adapt to perceived changes in climate patterns in Central America.

Practices	Estimate	Std. Error	z value	Pr (> |z|)	Signif.
Reforestation and Restoration	1.5120	0.0811	18.6470	< 0.0001	***
Introduction of new crops	0.7572	0.0844	8.9680	< 0.0001	***
Sustainable soil management	0.2554	0.0834	3.0620	0.0022	***
Production diversification	**0.0000**	–	–	–	–
Change in varieties	−0.2805	0.0883	−3.1770	0.0015	**
Sustainable water management	−0.6814	0.0919	−7.4140	< 0.0001	***
Use of more fertilizers and pesticides	−0.7658	0.0925	−8.2820	< 0.0001	***
Use of less fertilizers and pesticides	−0.8516	0.0942	−9.0400	< 0.0001	***
Leave farming system	−1.4053	0.1069	−13.1440	< 0.0001	***
Change in agricultural calendar	−1.5276	0.1095	−13.9520	< 0.0001	***

Significance levels: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1.

restrictions and bureaucratic regulations for the utilization of trees outside forests (mainly for timber) in many Central American countries (Detlefsen and Scheelje, 2012). Despite the lack on incentives to grow trees, we show that across the Rainforest, agroforestry (reforestation + introduction of new crops) was the first approach employed by farmers to adapt their systems, which is in accordance with the recent analysis conducted by Somarriba et al. (2017) in this region. Considering, however, the expected impacts of climate change on distribution and suitability of the most common tree species used in Central America (de Sousa et al., 2017), it is necessary to increase farmer's awareness to select the best climate suited trees for their farms.

Illiterate farmers with small landholdings living in the Dry Corridor chose a set of approaches to adapt their systems and intensify the production that includes the adoption of new crops, soil management, and increased use of fertilizers. These practices, when integrated and well managed, can help smallholders to achieve high yields (Cassman, 1999) while reducing the need to expand the production to new crop areas. However, two concerns arise for this group. First, it is not clear if the increased utilization of fertilizers is employed under an optimal level to ensure sustainability and soil conservation, considering the crop and soil requirements. Second, the adoption of this technological package could, in the long run, lead to a high dependency of external inputs, a non-desired outcome in the concept of Climate-Smart Agriculture. To avoid this risk, farmers could employ integrated nutrient practices such as the utilization of nitrogen-fixing plants and green manures (Kang, 1997), which could be utilized as the only approach or integrated with a reduced amount of synthetic inputs.

Farmers living in the Dry Corridor with large farmland also selected reforestation and sustainable soil management as adaptation approaches. However, this group considered leaving the farm system as the third best adaptation strategy, which raises concerns about the future sources of food and household income to these families. The insufficient family workforce (~4 people with 15–60 years-old per family) in a large family farmland may drive farmers to this alternative. An approach for this group could be the intensification of small parts of their farms and utilization of intercropping systems such as *quesungual*, a high advocated alternative for drylands in Central America (Ayarza et al., 2010; Kang, 1993).

Changing agricultural calendar was one of the least preferred choices among interviewed farmers, which is unfortunate, as it is one of the simplest approaches to adapt to the effects of climate variability (Yegbemey et al., 2014). By adopting this approach, farmers can adjust the planting season to operate in a time-efficient manner and avoid extreme climatic events during sensitive growing phases, such as flowering (Sacks et al., 2010). The low preference for this approach may be the result of the scarce up-to-date agroclimatic information and forecasts on upcoming growing seasons, which are also in accordance with the partial correlations between farmers perceptions and the historical data observed in our analysis. The establishment of information services and early warning systems to provide seasonal forecasting and agroclimatic information can help farmers make the best decisions to adapt their systems under seasonal climate variability.

We show that the participation in long-term outreach projects can influence farmers' decision to adopt sustainable practices (Gutierrez-Montes et al., 2018; Mercado et al., 2017). In this study, we provide

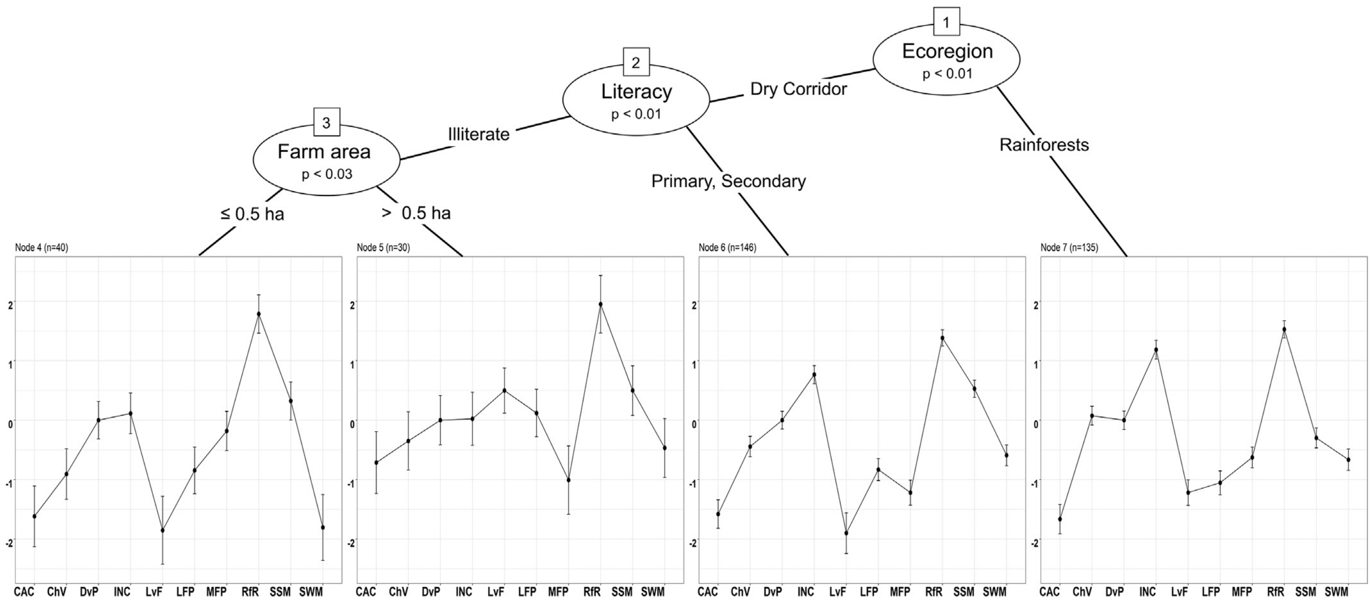

Fig. 5. Recursive partitioning of Bradley-Terry model of farmers' management practices employed to adapt to perceived changes in climate patterns in Central America. Intervals show quasi-standard errors. CAC = Change in agricultural calendar, Chv = Change in varieties, Dvp = Production diversification, INC = Introduction of new crops, Lvf = Leave farming system, LFP = Use of less fertilizers and pesticides, MFP = Use of more fertilizers and pesticides, RfR = Reforestation and restoration, SSM = Sustainable soil management, SWM = Sustainable water management.

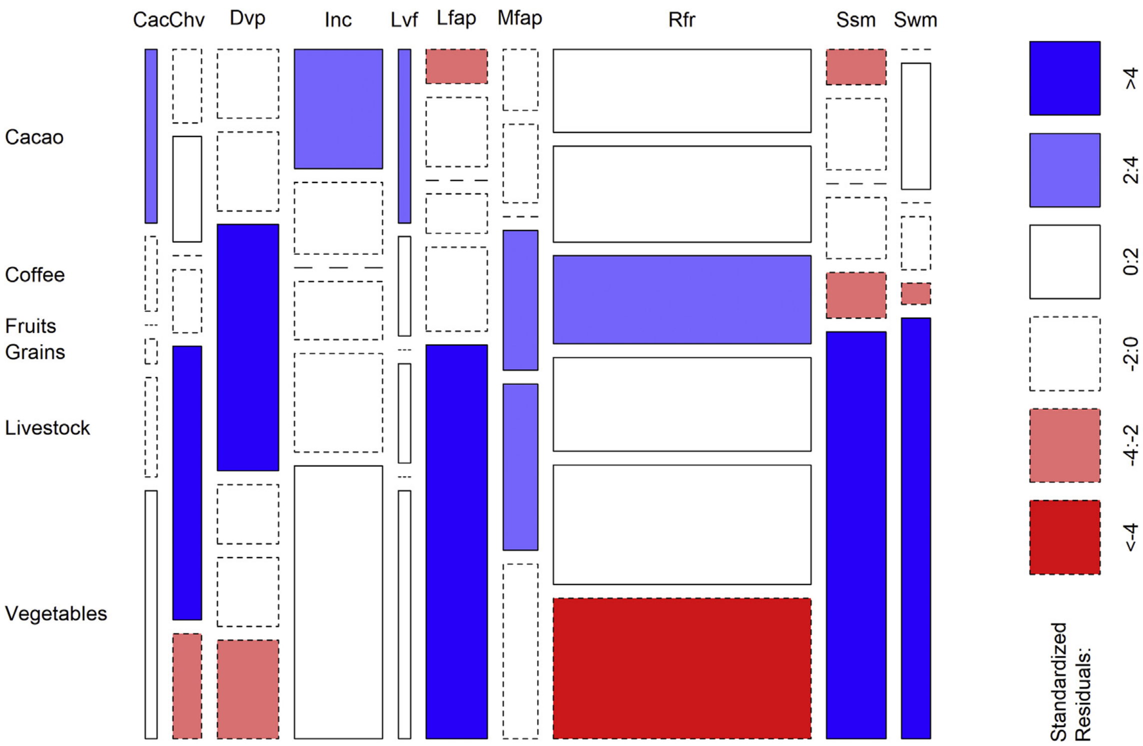

Fig. 6. Relationship between preferred adaptation practices and the main crop systems across the Central America Dry Corridor and Rainforests. CAC = Change in agricultural calendar, Chv = Change in varieties, Dvp = Production diversification, INC = Introduction of new crops, Lvf = Leave farming system, LFP = Use of less fertilizers and pesticides, MFP = Use of more fertilizers and pesticides, RfR = Reforestation and restoration, SSM = Sustainable soil management, SWM = Sustainable water management. Blue color indicate that the observed value is higher than the expected value if the data were random. Red color indicate that the observed value is lower than the expected value if the data were random. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

evidence to support the design and implementation of outreach projects oriented for specific groups of farmers according to their main livelihood, ecoregion, and education profile. For example, when dealing with livestock and illiterate farmers, these findings are very important since they are more likely to increase the use of fertilizers and pesticides and reduce practices for soil and water management. Also, we identified that the preference of farm practices is closely related with the main crop produced by the farmer. For example, the utilization of *Reforestation and Restoration* in farms producing fruits is increased by climate variability, while it is not a preferred option in farms producing vegetables. This finding demonstrates the importance of tailoring the Farmer Field Schools curricula to the farmers' characteristics and the main crop they produce. For example, the need to learn about climate-smart practices related to reforestation may be lower when regarding tree growers.

5. Conclusions

Our study provides an overview of farmers' perception of the changes in climate patterns in Central America and we argue that these perceptions to some extent drive the adoption of Climate-Smart Agriculture practices across the region. We demonstrate the relationship between farmers' awareness of climate variability and their responses through the use of climate-smart practices. Overall, farmers demonstrated self-motivation to adapt their systems to climate variability. Nevertheless, most of them require technical guidance to adopt sustainable practices for sustainable agriculture. The participation in Farmer Field Schools can help farmers make the best decisions to adapt their agricultural systems to climate variability.

As we have shown, there is a strong correlation between some socioeconomic characteristics and the adoption of specific technological packages. Illiterate farmers, for instance, adopted a set of practices that includes the utilization of more fertilizers, which may affect farmers in the long term by increasing their dependency on external inputs and increase financial risks. Therefore, we recommend tailoring the Farmer Field Schools curricula to the needs of each specific group, taking into account their farm size, educational level and main crop.

Although farmers demonstrated awareness to climate change and to its effects the lack of up-do-date agroclimatic information is still an issue that hinders making the best decision regarding crop management, especially for the annual crops. The promotion of community weather stations can help farmers obtain accurate information regarding the climate and thus close this information gap. Furthermore, local and international development agencies and NGOs should make use of the weather information and models already available to foster the adoption of short and long-term technological packages tailored to specific ecoregions.

Given the uncertainties of the multiple effects of climate change in agriculture (Howden et al., 2007; Vermeulen et al., 2013), farmers and stakeholders must be constantly updated about the latest recommendations for each climatic region and for each crop activity. Recent experiences with citizen-science in Central America, Africa and Asia (Beza et al., 2017; Mancini et al., 2017; Steinke et al., 2017; Steinke and van Etten, 2017; van Etten et al., 2016) showed that farmers and decision-makers can track the responses of crop systems to the changing climate patterns as they occur in the farm and take the best decision towards climate adaptation. Therefore, it is important to stay in the loop and understand that adaptation requires constant evaluations on the state of farming system and on the outcomes of employed practices in terms of climate adaptation and productivity.

Acknowledgments

This work has been possible thanks to the financial support from the

Royal Embassy of Norway (contract no. CAM-2647-12/0004, MAP Norway Second Phase). We would like to thank the farmers for their contributions in data collection and to Mirna Barrios, Danilo Padilla, Oscar Mattus, Ana Cristela Gutiérrez, Sayra Talelo, Amada Olivas, and Liseth Hernandez for their support in logistics across the research sites.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jrurstud.2018.09.018.

References

Aguilar, E., Peterson, T.C., Obando, P.R., Frutos, R., Retana, J.A., Solera, M., Soley, J., García, I.G., Araujo, R.M., Santos, A.R., Valle, V.E., Brunet, M., Aguilar, L., Álvarez, L., Bautista, M., Castañón, C., Herrera, L., Ruano, E., Sinay, J.J., Sánchez, E., Oviedo, G.I.H., Obed, F., Salgado, J.E., Vázquez, J.L., Baca, M., Gutiérrez, M., Centella, C., Espinosa, J., Martínez, D., Olmedo, B., Espinoza, C.E.O., Núñez, R., Haylock, M., Benavides, H., Mayorga, R., 2005. Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. J. Geophys. Res. 110, D23107. https://doi.org/10.1029/2005JD006119.

Ajzen, I., 1991. The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50, 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

Altieri, M.A., Nicholls, C.I., 2017. The adaptation and mitigation potential of traditional agriculture in a changing climate. Climatic Change 140, 33–45. https://doi.org/10.1007/s10584-013-0909-y.

Ayarza, M., Huber-Sannwald, E., Herrick, J.E., Reynolds, J.F., García-Barrios, L., Welchez, L.A., Lentes, P., Pavón, J., Morales, J., Alvarado, A., Pinedo, M., Baquera, N., Zelaya, S., Pineda, R., Amézquita, E., Trejo, M., 2010. Changing human–ecological relationships and drivers using the Quesungual agroforestry system in western Honduras. Renew. Agric. Food Syst. 25, 219–227. https://doi.org/10.1017/S1742170510000074.

Beza, E., Steinke, J., van Etten, J., Reidsma, P., Fadda, C., Mittra, S., Mathur, P., Kooistra, L., 2017. What are the prospects for citizen science in agriculture? Evidence from three continents on motivation and mobile telephone use of resource-poor farmers. PloS One 12, e0175700. https://doi.org/10.1371/journal.pone.0175700.

Borges, J.A.R., Oude Lansink, A.G.J.M., Marques Ribeiro, C., Lutke, V., 2014. Understanding farmers' intention to adopt improved natural grassland using the theory of planned behavior. Livest. Sci. 169, 163–174. https://doi.org/10.1016/J.LIVSCI.2014.09.014.

Bradley, R.A., Terry, M.E., 1952. Rank analysis of incomplete block designs: I. The method of paired comparisons. Biometrika 39, 324–345. https://doi.org/10.2307/2334029.

Campbell, B.M., Vermeulen, S.J., Aggarwal, P.K., Corner-Dolloff, C., Girvetz, E., Loboguerrero, A.M., Ramirez-Villegas, J., Rosenstock, T., Sebastian, L., Thornton, P., Wollenberg, E., 2016. Reducing risks to food security from climate change. Glob. Food Sec. 0–1. https://doi.org/10.1016/j.gfs.2016.06.002.

Cassman, K.G., 1999. Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. Proc. Natl. Acad. Sci. U. S. A 96, 5952–5959. https://doi.org/10.1073/PNAS.96.11.5952.

Caudill, S.A., DeClerck, F.J.A., Husband, T.P., 2015. Connecting sustainable agriculture and wildlife conservation: does shade coffee provide habitat for mammals? Agric. Ecosyst. Environ. https://doi.org/10.1016/j.agee.2014.08.023.

de Sousa, K., van Zonneveld, M., Imbach, P., Kindt, R., Ordoñez, J.C., 2017. Suitability of Key Central American Agroforestry Species under Future Climates: an Atlas. World Agroforestry Centre, Lima, Peru Occasional Paper 26.

Detlefsen, G., Scheelje, M., 2012. Las normativas legales y el aprovechamiento de la madera en fincas. In: Detlefsen, G., Somarriba, E. (Eds.), Producción de Madera En Sistemas Agroforestales de Centroamérica. Centro Agronómico Tropical de Investigación y Enseñansa, Turriaba, Costa Rica, pp. 211–243.

Elum, Z.A., Modise, D.M., Marr, A., 2017. Farmer's perception of climate change and responsive strategies in three selected provinces of South Africa. Clim. Risk Manag. 16, 246–257. https://doi.org/10.1016/j.crm.2016.11.001.

FAO (Food and Agriculture Organization of The United Nations), 2013. Climate-smart Agriculture: Sourcebook, first ed. Food and Agriculture Organization of The United Nations, Rome, Italy.

FAO (Food and Agriculture Organization of The United Nations Italy), 2010. Climate-Smart Agriculture: policies, practices and financing for food security, adaptation and mitigation. In: Hague Conf. Agric. Food Secur. Clim. Chang.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Sci. Data 2, 150066. https://doi.org/10.1038/sdata.2015.66.

Gutierrez-Montes, I., Arguedas, M., Ramirez-Aguero, F., Mercado, L., Sellare, J., 2018. Contributing to the construction of a framework for improved gender integration into climate-smart agriculture projects monitoring and evaluation: MAP-Norway experience. Climatic Change 1–14. https://doi.org/10.1007/s10584-018-2231-1.

Hannah, L., Donatti, C.I., Harvey, C.A., Alfaro, E., Rodriguez, D.A., Bouroncle, C., Castellanos, E., Diaz, F., Fung, E., Hidalgo, H.G., Imbach, P., Läderach, P., Landrum, J.P., Solano, A.L., 2017. Regional modeling of climate change impacts on smallholder agriculture and ecosystems in Central America. Climatic Change 141, 29–45. https://doi.org/10.1007/s10584-016-1867-y.

Harvey, C.A., Rakotobe, Z.L., Rao, N.S., Dave, R., Razafimahatratra, H., Rabarijohn, R.H., Rajaofara, H., MacKinnon, J.L., 2014. Extreme vulnerability of smallholder farmers to agricultural risks and climate change in Madagascar. Philos. Trans. R. Soc. B Biol. Sci. https://doi.org/10.1098/rstb.2013.0089.

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. Int. J. Climatol. 25, 1965–1978. https://doi.org/10.1002/joc.1276.

Hothorn, T., Zeileis, A., 2015. Partykit: a modular toolkit for recursive partytioning in R. J. Mach. Learn. Res. 16, 3905–3909.

Howden, S.M., Soussana, J.-F., Tubiello, F.N., Chhetri, N., Dunlop, M., Meinke, H., 2007. Adapting agriculture to climate change. Proc. Natl. Acad. Sci. U. S. A 104, 19691–19696. https://doi.org/10.1073/pnas.0701890104.

Imbach, P., Beardsley, M., Bouroncle, C., Medellin, C., Läderach, P., Hidalgo, H., Alfaro, E., Van Etten, J., Allan, R., Hemming, D., Stone, R., Hannah, L., Donatti, C.I., 2017. Climate change, ecosystems and smallholder agriculture in Central America: an introduction to the special issue. Climatic Change 141, 1–12. https://doi.org/10.1007/s10584-017-1920-5.

Kang, B.T., 1997. Alley cropping—soil productivity and nutrient recycling. For. Ecol. Manage. 91, 75–82. https://doi.org/10.1016/S0378-1127(96)03886-8.

Kang, B.T., 1993. Alley cropping: past achievements and future directions. Agrofor. Syst. 23, 141–155. https://doi.org/10.1007/BF00704912.

Khatri-Chhetri, A., Aggarwal, P.K., Joshi, P.K., Vyas, S., 2017. Farmers' prioritization of climate-smart agriculture (CSA) technologies. Agric. Syst. 151, 184–191. https://doi.org/10.1016/j.agsy.2016.10.005.

Khonje, M.G., Manda, J., Mkandawire, P., Tufa, A.H., Alene, A.D., 2018. Adoption and welfare impacts of multiple agricultural technologies: evidence from eastern Zambia. Agric. Econ. https://doi.org/10.1111/agec.12445.

Koppmair, S., Kassie, M., Qaim, M., 2017. The influence of farm input subsidies on the adoption of natural resource management technologies. Aust. J. Agric. Resour. Econ. 61, 539–556. https://doi.org/10.1111/1467-8489.12220.

Lê, S., Josse, J., Husson, F., 2008. FactoMineR: an R package for multivariate analysis. J. Stat. Software 25, 1–18. https://doi.org/10.18637/jss.v025.i01.

Lipper, L., Thornton, P., Campbell, B.M., Baedeker, T., Braimoh, A., Bwalya, M., Caron, P., Cattaneo, A., Garrity, D., Henry, K., Hottle, R., Jackson, L., Jarvis, A., Kossam, F., Mann, W., McCarthy, N., Meybeck, A., Neufeldt, H., Remington, T., Sen, P.T., Sessa, R., Shula, R., Tibu, A., Torquebiau, E.F., 2014. Climate-smart agriculture for food security. Nat. Clim. Change 4, 1068–1072. https://doi.org/10.1038/nclimate2437.

Locatelli, B., Catterall, C.P., Imbach, P., Kumar, C., Lasco, R., Marín-Spiotta, E., Mercer, B., Powers, J.S., Schwartz, N., Uriarte, M., 2015. Tropical reforestation and climate change: beyond carbon. Restor. Ecol. 23, 337–343. https://doi.org/10.1111/rec.12209.

Long, T.B., Blok, V., Coninx, I., 2016. Barriers to the adoption and diffusion of technological innovations for climate-smart agriculture in Europe: evidence from The Netherlands, France, Switzerland and Italy. J. Clean. Prod. 112, 9–21. https://doi.org/10.1016/J.JCLEPRO.2015.06.044.

Lyra, A., Imbach, P., Rodriguez, D., Chou, S.C., Georgiou, S., Garofolo, L., 2017. Projections of climate change impacts on central America tropical rainforest. Climatic Change 141, 93–105. https://doi.org/10.1007/s10584-016-1790-2.

Mancini, C., Kidane, Y.G., Mengistu, D.K., Pè, M.E., Fadda, C., Dell'Acqua, M., 2017. Joining smallholder farmers' traditional knowledge with metric traits to select better varieties of Ethiopian wheat. Sci. Rep. 7, 9120. https://doi.org/10.1038/s41598-017-07628-4.

Marchal, V., Dellink, R., van Vuuren, D., Clapp, C., Château, J., Lanzi, E., Magné, B., van Vliet, J., 2011. Chapter 3: climate change. In: OECD Enrironmental Outlook to 2050.

Mbow, C., Smith, P., Skole, D., Duguma, L., Bustamante, M., 2014. Achieving mitigation and adaptation to climate change through sustainable agroforestry practices in Africa. Curr. Opin. Environ. Sustain. 6, 8–14. https://doi.org/10.1016/J.COSUST.2013.09.002.

Meijer, S.S., Catacutan, D., Ajayi, O.C., Sileshi, G.W., Nieuwenhuis, M., 2015. The role of knowledge, attitudes and perceptions in the uptake of agricultural and agroforestry innovations among smallholder farmers in sub-Saharan Africa. Int. J. Agric. Sustain. 13, 40–54. https://doi.org/10.1080/14735903.2014.912493.

Mercado, L., Aguilar, A., Padilla, D., Cerda, R., Ospina, A., Arguedas, M., 2017. Programa Ambiental Mesoamericano (MAP) Noruega 2013-2017: Principales Resultados Y Lecciones Aprendidas. Turriaba, Costa Rica.

Mwangi, M., Kariuki, S., 2015. Factors determining adoption of new agricultural technology by smallholder farmers in developing countries. J. Econ. Sustain. Dev. 6, 208–216.

Niles, M.T., Mueller, N.D., 2016. Farmer perceptions of climate change: associations with observed temperature and precipitation trends, irrigation, and climate beliefs. Global Environ. Change 39, 133–142. https://doi.org/10.1016/j.gloenvcha.2016.05.002.

Ogundari, K., Bolarinwa, O.D., 2018. Impact of agricultural innovation adoption: a meta-analysis. Aust. J. Agric. Resour. Econ. 62, 217–236. https://doi.org/10.1111/1467-8489.12247.

Sacks, W.J., Deryng, D., Foley, J.A., Ramankutty, N., 2010. Crop planting dates: an analysis of global patterns. Global Ecol. Biogeogr. 19 no-no. https://doi.org/10.1111/j.1466-8238.2010.00551.x.

Schattman, R.E., Conner, D., Méndez, V.E., 2016. Farmer perceptions of climate change risk and associated on-farm management strategies in Vermont, northeastern United States. Elem. Sci. Anthr. 4, 000131. https://doi.org/10.12952/journal.elementa.000131.

Senyolo, M.P., Long, T.B., Blok, V., Omta, O., 2018. How the characteristics of innovations impact their adoption: an exploration of climate-smart agricultural innovations in South Africa. J. Clean. Prod. 172, 3825–3840. https://doi.org/10.1016/J.JCLEPRO.2017.06.019.

Singh, R.K., Zander, K.K., Kumar, S., Singh, A., Sheoran, P., Kumar, A., Hussain, S.M.,

Riba, T., Rallen, O., Lego, Y.J., Padung, E., Garnett, S.T., 2017. Perceptions of climate variability and livelihood adaptations relating to gender and wealth among the Adi community of the Eastern Indian Himalayas. Appl. Geogr. 86, 41–52. https://doi.org/10.1016/J.APGEOG.2017.06.018.
Somarriba, E., Carreño-Rocabado, G., Amores, F., Caicedo, W., Oblitas Gillés de Pélichy, S., Cerda, R., Ordóñez, J.C., 2017. Trees on Farms for Livelihoods, Conservation of Biodiversity and Carbon Storage: Evidence from Nicaragua on This "Invisible" Resource. Springer, Cham, pp. 369–393. https://doi.org/10.1007/978-3-319-69371-2_15.
Steenwerth, K.L., Hodson, A.K., Bloom, A.J., Carter, M.R., Cattaneo, A., Chartres, C.J., Hatfield, J.L., Henry, K., Hopmans, J.W., Horwath, W.R., Jenkins, B.M., Kebreab, E., Leemans, R., Lipper, L., Lubell, M.N., Msangi, S., Prabhu, R., Reynolds, M.P., Sandoval Solis, S., Sischo, W.M., Springborn, M., Tittonell, P., Wheeler, S.M., Vermeulen, S.J., Wollenberg, E.K., Jarvis, L.S., Jackson, L.E., Smith, J., Smith, P., Sykes, M., 2014. Climate-smart agriculture global research agenda: scientific basis for action. Agric. Food Secur. 3, 11. https://doi.org/10.1186/2048-7010-3-11.
Steinke, J., van Etten, J., 2017. Gamification of farmer-participatory priority setting in plant breeding: design and validation of "AgroDuos. J. Crop Improv. 1–23. https://doi.org/10.1080/15427528.2017.1303801.
Steinke, J., van Etten, J., Zelan, P.M., 2017. The accuracy of farmer-generated data in an agricultural citizen science methodology. Agron. Sustain. Dev. 37, 32. https://doi.org/10.1007/s13593-017-0441-y.
Strobl, C., Wickelmaier, F., Zeileis, A., 2011. Accounting for individual differences in bradley-terry models by means of recursive partitioning. J. Educ. Behav. Stat. 36, 135–153. https://doi.org/10.3102/1076998609359791.
Theus, M., Urbanek, S., 2008. Interactive Graphics for Data Analysis: Principles and Examples. Taylor & Francis.
Torres, C.M.M.E., Jacovine, L.A.G., Nolasco de Olivera Neto, S., Fraisse, C.W., Soares, C.P.B., de Castro Neto, F., Ferreira, L.R., Zanuncio, J.C., Lemes, P.G., 2017. Greenhouse gas emissions and carbon sequestration by agroforestry systems in southeastern Brazil. Sci. Rep. 7, 16738. https://doi.org/10.1038/s41598-017-16821-4.
Turner, H., Firth, D., 2012. Bradley-terry models in R: the BradleyTerry2 package. J. Stat. Software 48, 1–21.
van Etten, J., Beza, E., Calderer, L., van Duijvendijk, K., Fadda, C., Fantahun, B., Kidane, Y.G., van de Gevel, J., Gupta, A., Mengistu, D.K., Kiambi, D., Mathur, P.N., Mercado, L., Mittra, S., Mollel, M.J., Rosas, J.C., Steinke, J., Suchini, J.G., Zimmerer, K.S., 2016. First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). Exp. Agric. 1–22. https://doi.org/10.1017/S0014479716000739.
Vermeulen, S.J., Challinor, A.J., Thornton, P.K., Campbell, B.M., Eriyagama, N., Vervoort, J.M., Kinyangi, J., Jarvis, A., Läderach, P., Ramirez-Villegas, J., Nicklin, K.J., Hawkins, E., Smith, D.R., 2013. Addressing uncertainty in adaptation planning for agriculture. Proc. Natl. Acad. Sci. U. S. A 110, 8357–8362. https://doi.org/10.1073/pnas.1219441110.
Wassie, A., Pauline, N., 2018. Evaluating smallholder farmers' preferences for climate smart agricultural practices in Tehuledere District, northeastern Ethiopia. Singapore J. Trop. Geogr. 39, 300–316. https://doi.org/10.1111/sjtg.12240.
Wauters, E., Bielders, C., Poesen, J., Govers, G., Mathijs, E., 2010. Adoption of soil conservation practices in Belgium: an examination of the theory of planned behaviour in the agri-environmental domain. Land Use Pol. 27, 86–94. https://doi.org/10.1016/J.LANDUSEPOL.2009.02.009.
Yegbemey, R.N., Kabir, H., Awoye, O.H.R., Yabi, J.A., Paraïso, A.A., 2014. Managing the agricultural calendar as coping mechanism to climate variability: a case study of maize farming in northern Benin, West Africa. Clim. Risk Manag. 3, 13–23.

2

Coffee agroforestry system in El Cuá, Nicaragua. Credit: K. de Sousa

SCIENTIFIC REPORTS

OPEN

The future of coffee and cocoa agroforestry in a warmer Mesoamerica

Received: 6 December 2018
Accepted: 10 June 2019
Published online: 20 June 2019

Kauê de Sousa[1,2], Maarten van Zonneveld[2,3], Milena Holmgren[4], Roeland Kindt[5] & Jenny C. Ordoñez[6,7]

Climate change threatens coffee production and the livelihoods of thousands of families in Mesoamerica that depend on it. Replacing coffee with cocoa and integrating trees in combined agroforestry systems to ameliorate abiotic stress are among the proposed alternatives to overcome this challenge. These two alternatives do not consider the vulnerability of cocoa and tree species commonly used in agroforestry plantations to future climate conditions. We assessed the suitability of these alternatives by identifying the potential changes in the distribution of coffee, cocoa and the 100 most common agroforestry trees found in Mesoamerica. Here we show that cocoa could potentially become an alternative in most of coffee vulnerable areas. Agroforestry with currently preferred tree species is highly vulnerable to future climate change. Transforming agroforestry systems by changing tree species composition may be the best approach to adapt most of the coffee and cocoa production areas. Our results stress the urgency for land use planning considering climate change effects and to assess new combinations of agroforestry species in coffee and cocoa plantations in Mesoamerica.

Adapting agricultural systems to climate change is particularly challenging for perennial crops that take long before farmers fully benefit from their management decisions. Yet, a sense of urgency has developed among farmers, scientists and policy makers across the tropics as climate warming and extreme weather events compromise the productivity of major perennial crops[1]. In Mesoamerica – the area comprising Panama to central Mexico – the productivity of Arabica coffee (*Coffea arabica* L.) is expected to drastically decline as suitable growing areas shift[2], and pests and pathogens incidence increases under unfavourable climate conditions[3,4].

Since the first reports of potential impacts of climate change on coffee suitability[2] an ever growing number of news and blogs from private sector, NGO's and research organisations are reporting the replacement of coffee by cocoa in zones under 600 m a.s.l. (above the sea level) mainly in Mesoamerica (supplementary information Table S1). According to these sources the drivers of this shift are trends in recent years of increasing coffee production costs and large loses due to pests and diseases (leaf rust crisis)[4] at low altitudes, attributed to climate change and fuelled by differences in coffee and cocoa prices. All in all, replacing coffee by cocoa has become one of the main strategies for climate change adaptation for producers in low elevation areas[5], already taking place in Nicaragua, Honduras and El Salvador. Moreover, this strategy is strongly advocated by large NGO's and development agencies active across the region, under the assumption that areas not suitable for coffee can become unequivocally suitable for cocoa[6]. Nevertheless, there is no quantitative assessment of the feasibility of such strategy, starting from considering that cocoa is vulnerable to climate change itself[7,8], plus other limitations for transformation of cropping systems.

On the other hand, agroforestry – the deliberate and simultaneous management of trees within crop or livestock systems[9,10] –, is considered another key strategy to increase the resilience of agricultural systems to climate change[11–13]. Currently, most coffee and cocoa production in Mesoamerica occurs in agroforestry systems[14,15]. Under proper management, agroforestry trees can improve microclimatic conditions that reduce abiotic stress and facilitate the performance of understory crops[16,17]. In addition, farmers can benefit from agroforestry systems

[1]Department of Agricultural Sciences, Inland Norway University of Applied Sciences, 2322, Hamar, Norway. [2]Bioversity International, 30501, Turrialba, Costa Rica. [3]World Vegetable Center, 741, Shanhua, Taiwan. [4]Resource Ecology Group, Wageningen University, 6708, Wageningen, The Netherlands. [5]World Agroforestry Centre, 30677, Nairobi, Kenya. [6]Latin America Regional Office, World Agroforestry Centre, 1558, Lima, Peru. [7]Facultad de Ingenieria Agroindustrial, Universidad de las Américas, 170125, Quito, Ecuador. Correspondence and requests for materials should be addressed to K.d.S. (email: kaue.desousa@inn.no)

by its capacity to provide a number of ecological services, such as water and soil conservation, maintenance of soil fertility and biodiversity conservation[18]. Nevertheless, climate change can also affect the future ecological niches of several tree species[19,20] and may restrain the prospects of agroforestry as a viable approach for climate adaptation.

To evaluate these two alternatives, shifting coffee-cocoa plantations or maintaining and promoting crops-agroforestry, we assessed the vulnerability of both coffee and cocoa under climate change and the potential impacts of climate change on the habitat suitability for 100 of the most common tree species in coffee and cocoa plantations across Mesoamerica. We modelled current and future climatic niches with ensemble modelling algorithms[21] using bioclimatic information[22], downscaled from 17 General Circulation Models, under two Representative Concentration Pathways scenarios of climate change[23]. We selected the intermediate scenario RCP 4.5, which predicts an average temperature increase of 1.4 °C (0.9–2.0 °C), and a scenario with high emissions RCP 8.5, which predicts an average temperature increase of 2.0 °C (1.4–2.6 °C) by 2050 (period 2046–2065). We focus on climate projections for the 2050 s to align with the United Nations framework of global challenges in agriculture and food security[13]. For simplicity, we focus the results in the intermediate scenario and included the variation between the two scenarios assessed here into the main text, the full results for climate change scenario with high emissions are available as supplementary information.

Results

Coffee is more vulnerable to climate change than cocoa. Between 55–62% of current areas for coffee production will no longer be suitable by 2050 (Fig. 1a) especially in mid-altitudinal areas (400–700 m a.s.l.). Highlands (>1,800 m a.s.l.) may partly compensate these losses, where coffee will likely expand up to 9–13%. In contrast, cocoa production will probably lose between 13–17% of the current distribution range (Fig. 1a) especially in some lowland areas (0–300 m a.s.l.), expected to become drier in the next decades[19]. Our model projections show that 83–87% of current cocoa areas will remain suitable, especially in the humid areas along the Atlantic coast (0–300 m a.s.l.) (Fig. 1b; Supplementary Fig. S1, Text S1).

Cocoa could potentially replace 85% of the vulnerable coffee areas under climate change in moist regions at elevations under 400 m a.s.l. and 53% at elevations between 400–700 m a.s.l. Areas to be replaced decrease sharply with altitude with no possibility beyond 1,200 m.a.s.l under RCP 4.5 and 1,600 m.a.s.l under RCP 8.5 (Fig. 2, Supplementary Fig. S2).

Agroforestry trees: winners and losers. The distribution range of 79% of the tree species assessed in coffee areas and 62% of the tree species assessed in cocoa areas will drastically shrink or become unsuitable in both remaining and vulnerable areas for coffee and cocoa. Major losses are expected for the most popular trees used for fruits, *N*-fixing and timber in mid-altitudinal coffee areas (400–700 m a.s.l.) and lowland cocoa areas (0–300 m a.s.l.; Fig. 3).

Looking at specific tree groups by their main use, we estimate that 20 of the 33 fruit trees will lose more than 15% of their current suitability in coffee areas. The same trend is observed for 14 fruit trees in cocoa suitable areas. The common fruit trees in coffee and cocoa plantations, *Persea americana* (avocado), *Psidium guajava* (guava) and *Manguifera indica* (mango) are among the most vulnerable species with average loss of 53% in suitable areas. Major gains (>15%), however, are found for species such as *Spondias mombin* (jobo) and *Manilkara zapota* (sapodilla) in coffee, *Melicoccus bijugatus* (mamon) in cocoa and *Tamarindus indica* (tamarind) in both coffee and cocoa areas (Fig. 4a, Supplementary Fig. S3).

High losses (>15%) are expected for 25 of the 30 *N*-fixing tree species assessed in coffee and for 18 *N*-fixing tree species in cocoa areas (Fig. 4b, Supplementary Fig. S4). Most common *N*-fixing trees currently growing in coffee and cocoa plantations, such as *Erythrina poeppigiana* (poró), *Inga oerstediana*, *I. ruiziana* and *I. jinicuil* (guama) are the most vulnerable to expected climate change, with losses of 56% in suitable areas. Only two species, of the selected, may expand their suitability in >26% across cocoa areas, *Inga laurina* (guama) and *Senna atomaria* (vainillo), but only up to 4% in future coffee areas.

In the case of timber trees, we estimate loses of >15% for 22 of 37 species in coffee and 12 tree species in cocoa areas. The most vulnerable timber species include the widely common *Cedrela odorata* (cedar), as well as, the locally important timber species *Perymenium grande* (tatascán) and *Pachira quinata* (pochote), in both coffee and cocoa areas (Fig. 4c, Supplementary Fig. S5). Marginal gains (~5%) are expected for *Albizia saman* (carreto), *Ceiba pentandra* (ceiba) and *Guazuma ulmifolia* (guácimo) in both coffee and cocoa areas.

Prospects for future coffee and cocoa under agroforestry. Despite the overall losses in suitability for some of the most popular tree species, our projections suggest that agroforestry could persist as a viable alternative to manage coffee and cocoa plantations in Mesoamerica under climate change. By 2050, approximately 72% of coffee areas (both, remaining and vulnerable) will be suitable for more than 30 tree species. This includes a portfolio of at least 10 species per main use (10 fruit species, 10 *N*-fixing species and 10 timber species). Most of these tree species are already present in coffee plantations but mainly in low densities and remain underutilised. Only 9% of coffee areas have very low tree species options (≤3 species).

Our results suggest that cocoa suitable areas have a higher potential for agroforestry than coffee. By 2050, 95% of cocoa areas will be suitable for more than 30 tree species. Only 3% of cocoa areas have very low tree species options (≤3 species) potentially available (Supplementary Fig. S6).

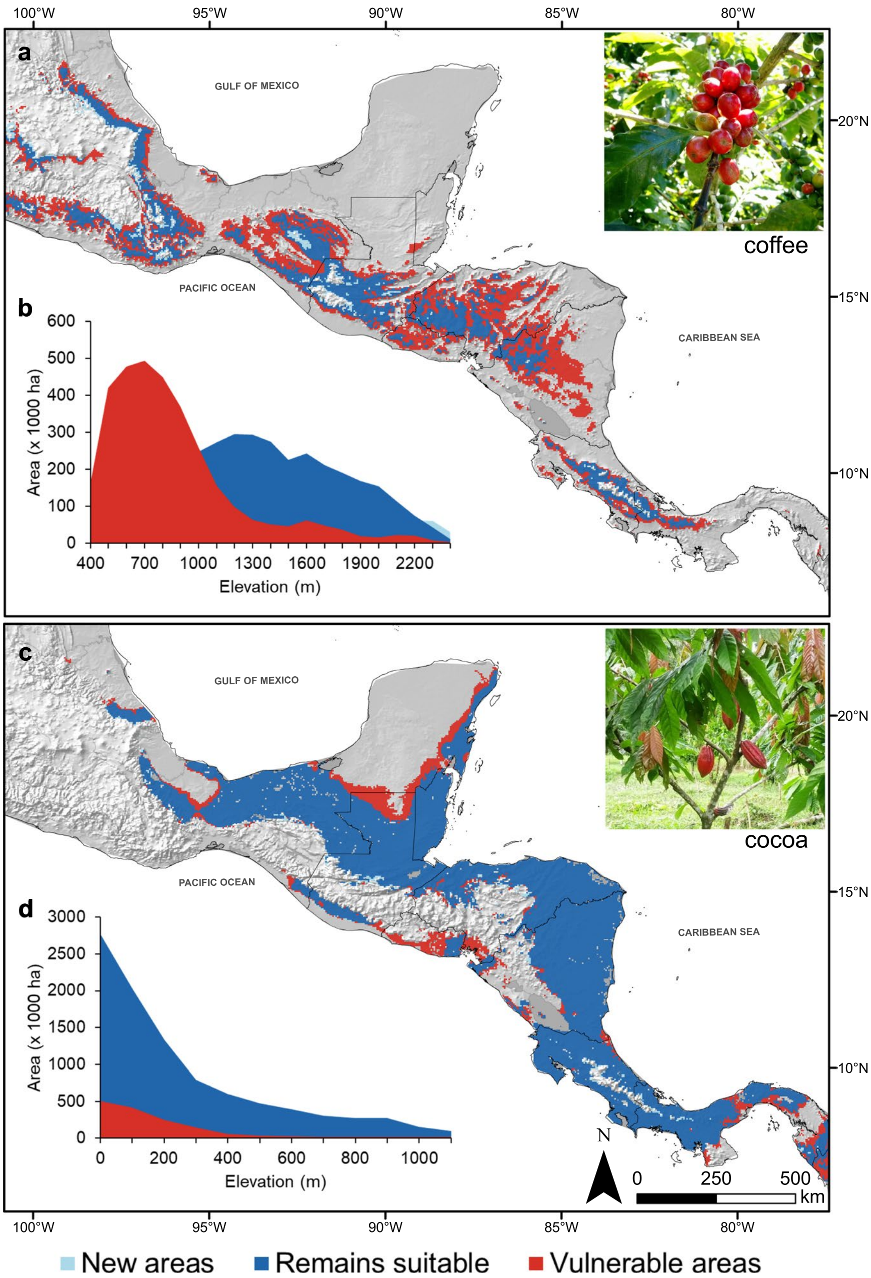

Figure 1. Shifts in suitability due to climate change (RCP 4.5) by 2050 for (**a**) coffee (*Coffea arabica* L.) and (**c**) cocoa (*Theobroma cacao* L.) in Mesoamerica. In (**b**,**d**), shifts in suitability are shown for the altitudinal gradient covered by coffee and cocoa within the continent. Light blue indicate new areas for coffee/cocoa by 2050. Dark blue indicate areas where coffee/cocoa will remain suitable under climate change. Red indicate areas expected to be no longer suitable (vulnerable) for coffee/cocoa under climate change.

Discussion

Our results stress the urgency for land use planning that considers potential climate change impacts to define the best areas and growing systems for production of coffee and cocoa under agroforestry management. These results suggest that important changes in tree species composition will be needed for agroforestry systems to remain as the best alternative for climate adaptation of coffee and cocoa fields.

Large areas are highly suitable for cocoa production in Mesoamerica under current climatic conditions and this suitability remains under climate change in 2050, opposing to the trends reported for the current largest cocoa production countries in West Africa[24]. In fact, the total area potentially suitable for cocoa in 2050 in the region could be four times the current world's cocoa producing area (11 M ha)[25] stressing the comparative advantage of the region for cocoa production. Despite this large potential, currently Mesoamerica is a minor player in the global cocoa supply chain (providing <1% total world cocoa production in 2017). In general, cocoa production systems in the region include smallholders, with low levels of input use, old plantations and low yields

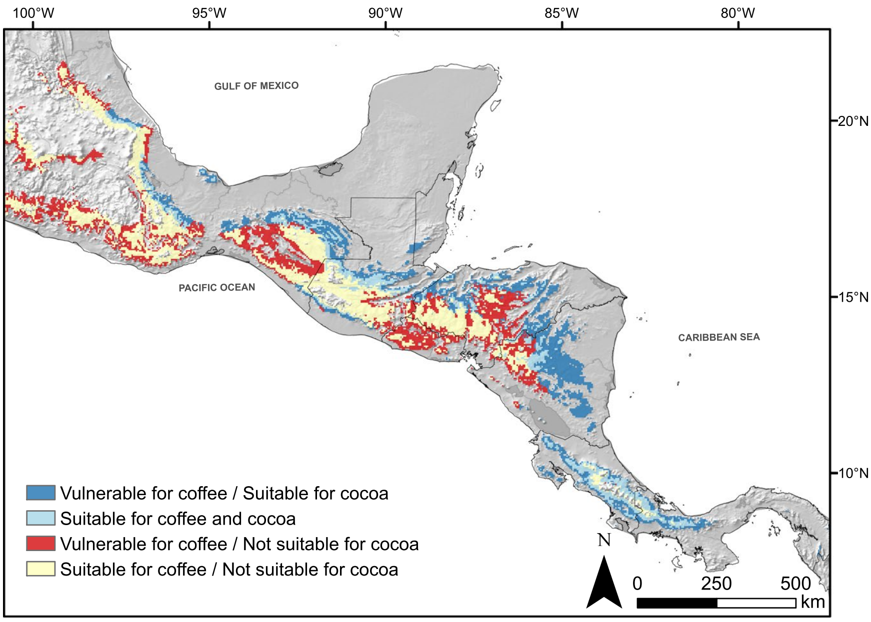

Figure 2. Potential areas in Mesoamerica where cocoa (*Theobroma cacao* L.) can replace coffee (Coffea arabica L.) under climate change (RCP 4.5). Dark blue indicate vulnerable areas for coffee that can be replaced by cocoa. Light blue indicate areas suitable for coffee and cocoa. Red indicate vulnerable areas for coffee where cocoa is not an alternative under climate change. Light yellow indicate remaining areas for coffee where cocoa is not suitable.

(60–328 kg ha^{-1} $year^{-1}$)[26]. It is argued that this panorama could change substantially if, for instance, farmers used to the management of a specialised perennial crop such as coffee, turn their efforts to cocoa production.

Only considering the coffee vulnerable areas to climate change that will be suitable for cocoa in 2050 (a modest 18% of the total suitable area), there could be 7.5 M ha in Mesoamerica available for cocoa production. Even at the extremely low yields typical of the region, these potentially new producing areas could add 1.5 million tons of cocoa to the global supply. In reality the actual coffee areas that can be replaced by cocoa will be lower than these estimated areas, because farmers may lack financial capacities to transform their coffee plantations[27] and the capacity to meet the strict existing quality standards. Still, the potential of the region remains large, but fuelling cocoa expansions will require well-structured efforts to i) reduce barriers to transformation, ii) ensure coupling of production to markets and iii) adequate land use planning to avoid expansion of cocoa into natural forests[28,29] (cocoa suitable areas do coincide with various protected areas within the Mesoamerican Biological Corridor).

Alternatively, by managing agroforestry systems, farmers could potentially maintain their current coffee and cocoa plantations using suitable trees to ameliorate microclimatic conditions. This alternative could also prevent the expansion of agricultural activities towards protected areas that are reported to be suitable in the future[30]. However, it seems highly probable that current agroforestry schemes will need to be modified in terms of species composition, since some of the most popular tree species are also vulnerable to future climate. It is particularly concerning the losses in habitat suitability of *N*-fixing trees such as *E. poeppigiana* (poró) and the majority of *Inga* species. These species make up the most abundant agroforestry trees in coffee and cocoa plantations in Mesoamerica[31,32], and have a key role for the management of soil fertility and sustain more stable productivity[33,34], especially in low-input and small farming plantations[35]. Therefore, our results anticipate a serious threat for future coffee and cocoa plantations if alternatives for *N*-fixing species are not promptly identified.

Rethinking current agroforestry species composition in coffee and cocoa landscapes requires the identification of the best tree species. Currently, farmers have a clear preference towards few species such as *C. odorata* (cedar), *E. poeppigiana* (poró), *Inga* spp., *M. indica* (mango), *P. americana* (avocado) and *P.guajava* (guava), all widespread in agricultural fields or open areas and of easy regeneration and propagation. We found that some currently underutilised tree species in coffee and cocoa plantations could potentially maintain or even increase their suitable distribution ranges under future climate, such as the fruit trees *M. sapota*, *S. dulcis*, *Brosimum alicastrum*, and the timber trees *Simarouba glauca* and *Ceiba pentandra*. These species are present in low densities in coffee and cocoa plantations, and most of them are remnants of previous vegetation[36].

Expanding the use of underutilised species in agroforestry systems will require a deeper understanding of their agronomic performance considering other factors beyond just climate (e.g. pest, diseases, soil fertility), ecological interactions[37–39], farmers' perceptions and local knowledge regarding management and utilisation of these tree species, as well as market incentives to facilitate their wider use. In our assessment, we employed a species distribution modelling (SDM) approach disregarding these aspects. Therefore, the interpretation of our results is driven by the expected changes in biophysical conditions characterised here as changes in extreme precipitation and temperature events. The evidence has shown that these changes are particularly important for

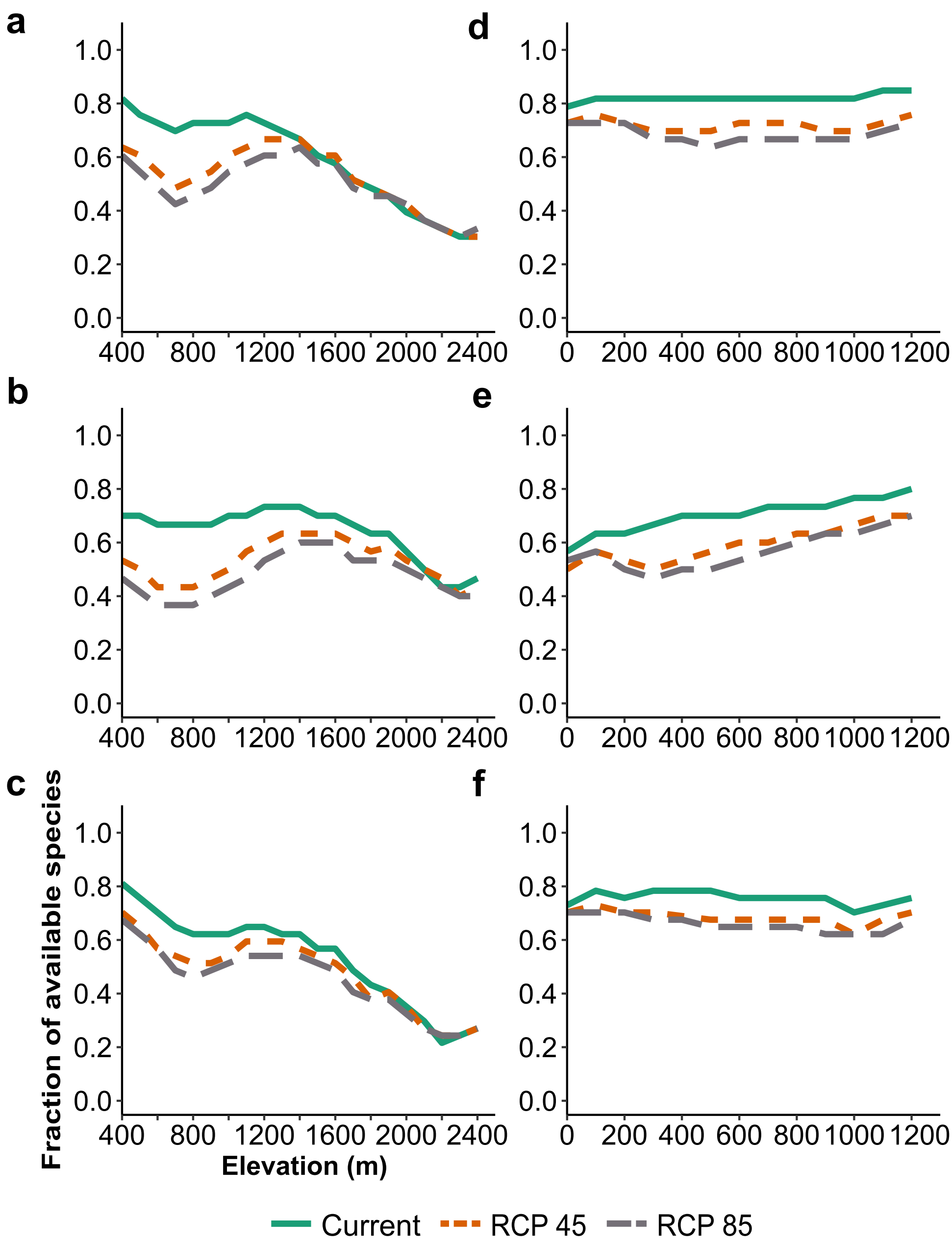

Figure 3. Changes in suitability of the 100 most common tree species in coffee (*Coffea arabica* L.) and cocoa (*Theobroma cacao* L.) agroforestry over the altitudinal gradient in Mesoamerica. Panels a, b and c shows the shifts for fruit, N-fixing and timber trees in coffee areas, respectively. Panels d, e and f shows the shifts for fruit, N-fixing and timber trees in cocoa areas, respectively.

agroecosystems in Mesoamerica, and other regions affected by El Niño Southern Oscillation, in which this phenomenon shapes the ecosystem productivity[20,40], not only across dry regions but also in rainforests[19].

Here we show that coffee systems are more vulnerable than cocoa systems to climate change. Not only is coffee more sensitive than cocoa to future climate, but also the tree species commonly used in coffee plantations are more vulnerable to the expected climate change. Cocoa as an alternative to coffee could potentially occur in most of the vulnerable coffee areas, but this will require addressing other ecological constraints, the impacts of pest and diseases, costs of technological change and market requirements to determine the real potential of cocoa to replace coffee. Adapting coffee and cocoa to changing climates can benefit from agroforestry systems with a new set of currently underutilised tree species already present in coffee and cocoa plantations. The results of this study are a starting point to develop lines of research that support the re-design of agroforestry schemes and open new venues of research to adapt coffee and cocoa production systems in Mesoamerica.

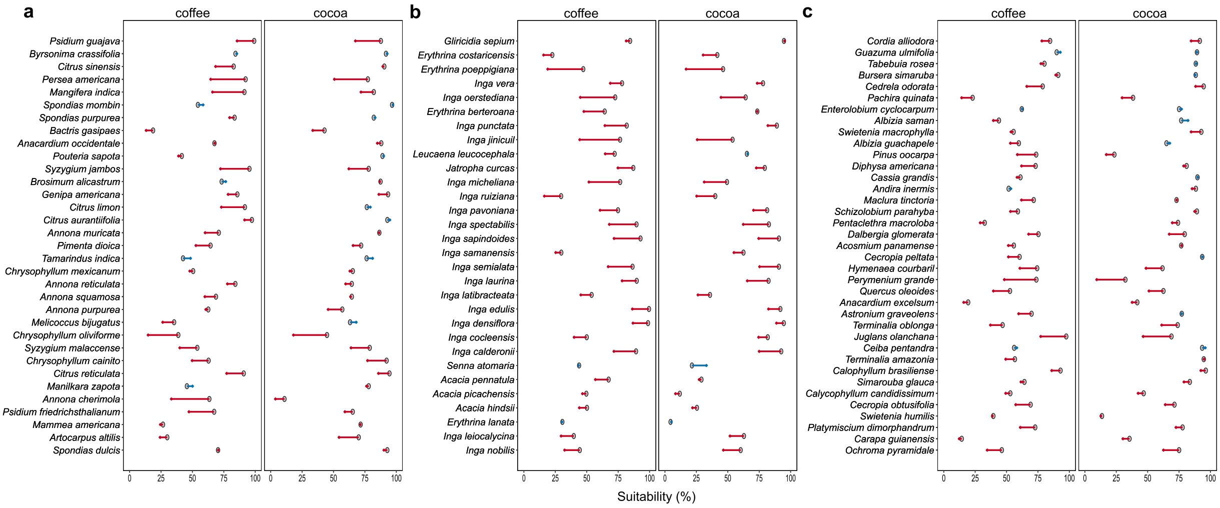

Figure 4. Expected changes in suitability due to climate change (RCP 4.5; expressed as % of current suitable areas) of the most common **a** fruit trees, **b** N-fixing trees and **c** timber trees in coffee (*Coffea arabica* L.) and cocoa (*Theobroma cacao* L.) plantations in Mesoamerica. Grey dot represent the area of a given species under the current climate conditions; Red arrows (left direction), represent decrease in suitable areas; Blue arrows (right direction) represent increase in suitable areas. Species ordered by main use and by their abundance (from to top to bottom) in the inventoried coffee and cocoa farms across Mesoamerica.

Methods

Selection of tree species. We selected 100 of the most commonly used tree species in cocoa and coffee plantations across Mesoamerica (Supplementary Table S2) using three criteria: (i) abundance assessed from compiled inventories of shade species in smallholder farms across the region[41–43]; (ii) ecological and economic services identified by farmers[44,45]; and, (iii) availability of a minimum of 60 records to ensure accurate modelling results[46].

From these 100 species, 30 are mainly used due to their potential to improve soil conditions by fixing nitrogen, 37 species mainly used for timber products (within the farm and potentially marketable) and 33 species mainly used as fruit trees[44,45]. The selected species belong to 27 botanical families and most (91 species) are native of the neotropics; the others are economically important species and naturalised fruit trees in Mesoamerica (Supplementary Table S2).

Compilation and validation of presence location points. We compiled presence location points of selected tree species (including coffee and cocoa) from the Global Biodiversity Information Facility (GBIF)[47], MAPFORGEN[48] and from the database of farm inventories used to select the tree species. No distinction was made between locations from natural forests or farms because this information was not always available in the original sources.

Records with no geographic information or with obvious errors such as incomplete coordinates, locations in the ocean and mismatches between administrative data and coordinates were excluded from the analysis. For this, we compared the collected presence data and information on administrative boundaries with information from the DIVA-GIS database[49], removing the mismatches. Presence locations from 1959 or before were also removed to meet the current baseline climate used. Finally we reduced the possible effects of sampling bias and spatial autocorrelation through systematic sampling[50]. This approach consists in create a grid of a defined cell size (in our case 2.5 arc-min) and randomly sample one presence points per grid cell. In the Fourcade *et al.*[50] assessment, the approach showed well performance among the other tested approaches irrespective the species and bias type, which is our case.

The final dataset with validated and unbiased presence locations comprised 130,480 occurrences for the 100 tree species combined (Supplementary Table S2), 2,194 location points for coffee and 1,241 location points for cocoa. Since absence locations were not available, for each species, we allocated 1,000 random pseudo-absence locations within the study area, which were sampled (without replacement) using the R[51] package dismo[52].

Climate data. We used bioclimatic predictors (baseline period of ~1960–1990) from WorldClim[22] at a spatial resolution of 2.5 arc-min. The bioclimatic variables include extreme or limiting factors that are ecologically important based on the variation in precipitation and temperature. We selected the least correlated variables applying an analysis of variance-inflation factors (VIF)[53], whereby the variables with the highest correlation ($VIF > 10$) were removed, resulting in nine bioclimatic predictors. Which were: (i) bio02, mean diurnal range; (ii) bio03, isothermality; (iii) bio08, mean temperature of wettest quarter; (iv) bio09, mean temperature of driest quarter, (v) bio13, precipitation of wettest month; (vi) bio14, precipitation of driest month; (vii) bio15, precipitation seasonality; (viii) bio18, precipitation of warmest quarter; and, (ix) bio19, precipitation of coldest quarter.

We based the projections of future distribution in 2050 s on two Representative Concentration Pathways scenarios (RCPs) of climate change from the Intergovernmental Panel on Climate Change (IPCC)[23]. We selected the intermediate scenario RCP 4.5, which predicts an average temperature increase of 1.4 °C (0.9–2.0 °C), and a scenario with very high emissions RCP 8.5, which predicts an average temperature increase of 2.0 °C (1.4–2.6 °C) by 2050 (period 2046–2065). We focus on climate projections for 2050 to align with the United Nations framework of global challenges in agriculture[13]. For each selected scenario, we predicted species suitability using the 17 General Circulation Models (GCM) available for both RCP scenarios (Supplementary Table S3).

Data analysis. We modelled the distribution of all species within the longitudes −101 and −77, and the latitudes 7 and 22. All analyses were done in R[51] using a consensus method for species distribution modelling (SDM) compiled by the package BiodiversityR[21], which calculate ensemble suitability as a weighted average of probabilities predicted by 17 SDM algorithms (Supplementary Table S4). Previous studies have shown that the consensus method based on weighted averages can significantly increase the accuracy of SDM[54].

For the model calibration, we performed a 4-fold cross-validation by randomly assigning (without replacement) location data to four bins. The performance of different SDM algorithms was evaluated for each bin separately after algorithms were calibrated with data from the other three bins. The SDM performance was assessed by the area under the curve (AUC[55]) criterion computed by the R package PresenceAbsence[56]. Although some authors tend to criticise this method, the evidence[57] has shown that AUC has strong correlation with the presence-absence threshold that makes sensitivity equal to specificity and remains a valid measure of relative model performance. Considering that, predictions from each of the 17 SDM algorithms were transformed to AUC weights by dividing each by the total of all AUC predictions. We selected the SDM algorithms with AUC weights >0.05, which means at least 5% of contribution to the consensus predictivity[21], and recalculated weights to sum to one[53]. The AUC values for the selected SDM models are shown in supplementary information Fig. S7.

Therefore, selected SDM algorithms were used to obtain the suitability model for coffee, cocoa and the 100 tree species. We then applied the derived suitability model to each of the 17 downscaled GCMs to predict the distribution of suitability by the 2050 s. For each species, ensemble suitability maps for baseline and future climates were converted in absence-presence maps with the recommended threshold method of maximum sensitivity (true positive) + specificity (true negative)[58,59].

Since there are no criteria to assess which of the GCMs best predict future climate, by incorporating all 17 GCMs we included all plausible changes in the distribution of the focal species. The results of the 17 GCMs presence-absence layers were integrated into a single layer, using the criterion of likelihood scale[60], which requires at least 66% of agreement among GCMs to keep the predicted presence or absence in a given grid cell.

Organising the datasets relied on R packages magrittr[61] and tidyverse[62]. Layers were processed using the R packages maptools[63], raster[64], rgeos[65] and rgdal[66]. To produce Figs 3, 4, S3, S4, S5 and S7, the R packages ggplot2[67] and svglite[68] were used.

Data Availability

Data and R code used is available through Dataverse[69]. The full project replication workflow is available through GitHub https://github.com/agrobioinfoservices/enm_agroforestry.

References

1. Altieri, M. A. & Nicholls, C. I. The adaptation and mitigation potential of traditional agriculture in a changing climate. *Clim. Change* **140**, 33–45 (2017).
2. Ovalle-Rivera, O., Läderach, P., Bunn, C., Obersteiner, M. & Schroth, G. Projected shifts in Coffea arabica suitability among major global producing regions due to climate change. *PLoS One* **10** (2015).
3. Rosner, H. Saving Coffee. *Sci. Am.* **311**, 68–73 (2014).
4. Avelino, J. *et al*. The coffee rust crises in Colombia and Central America (2008–2013): impacts, plausible causes and proposed solutions. *Food Secur.* **7**, 303–321 (2015).
5. Läderach, P. *et al*. Climate change adaptation of coffee production in space and time. *Clim. Change* **141**, 47–62 (2017).
6. Renteria, N. & Rowling, M. Hit by climate change, Central American coffee growers get a taste for cocoa. *Thomson Reuters Foundation* (2016).
7. Schmitz, H. & Shapiro, H.-Y. The Future of Chocolate. *Sci. Am.* 28–33, https://doi.org/10.1038/scientificamericanfood0615-28 (2015).
8. Wessel, M. & Quist-Wessel, P. M. F. Cocoa production in West Africa, a review and analysis of recent developments. *NJAS - Wageningen J. Life Sci.* **74–75**, 1–7 (2015).
9. Spurgeon, D. Agroforestry: new hope for subsistence farmers. *Nature* **280**, 533–534 (1979).
10. Nair, P. K. R. An introduction to agroforestry. 499 (1993).
11. Zomer, R. J. *et al*. Global Tree Cover and Biomass Carbon on Agricultural Land: The contribution of agroforestry to global and national carbon budgets. *Sci. Rep.* **6**, 29987 (2016).
12. Jones, H. P., Hole, D. G. & Zavaleta, E. S. Harnessing nature to help people adapt to climate change. *Nat. Clim. Chang.* **2**, 504–509 (2012).
13. Lipper, L. *et al*. Climate-smart agriculture for food security. *Nat. Clim. Chang.* **4**, 1068–1072 (2014).
14. Somarriba, E. *et al*. Carbon stocks and cocoa yields in agroforestry systems of Central America. *Agriculture, Ecosystems and Environment* **173**, 46–57 (2013).
15. Beer, J., Muschler, R., Kass, D. & Somarriba, E. Shade management in coffee and cacao plantations. *Agroforestry Systems* **38**, 139–164 (1998).
16. Holmgren, M. & Scheffer, M. Strong facilitation in mild environments: the stress gradient hypothesis revisited. *J. Ecol.* **98**, 1269–1275 (2010).
17. Blaser, W. J. *et al*. Climate-smart sustainable agriculture in low-to-intermediate shade agroforests. *Nat. Sustain.* **1**, 234–239 (2018).
18. Ellison, D. *et al*. Trees, forests and water: Cool insights for a hot world. *Glob. Environ. Chang.* **43**, 51–61 (2017).
19. Lyra, A. *et al*. Projections of climate change impacts on central America tropical rainforest. *Clim. Change* **141**, 93–105 (2017).
20. Holmgren, M., Hirota, M., van Nes, E. H. & Scheffer, M. Effects of interannual climate variability on tropical tree cover. *Nat. Clim. Chang.* **3**, 755–758 (2013).
21. Kindt, R. Ensemble species distribution modelling with transformed suitability values. *Environ. Model. Softw.* **100**, 136–145 (2018).
22. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
23. van Vuuren, D. P. *et al*. The representative concentration pathways: an overview. *Clim. Change* **109**, 5–31 (2011).

24. Schroth, G., Läderach, P., Martinez-Valle, A. I., Bunn, C. & Jassogne, L. Vulnerability to climate change of cocoa in West Africa: Patterns, opportunities and limits to adaptation. *Science of The Total Environment* **556**, 231–241 (2016).
25. FAO (Food and Agriculture Organization of The United Nations). FAOSTAT statistics database. Available at: http://www.fao.org/faostat/en/ (2019).
26. Orozco-Aguilar L, Olivier, D., Marilyn, V.-R. & Eduardo, S. *El sector cacao en Centroamérica. Estado de desarrollo en el año 2007*. (CATIE, 2015).
27. Howden, S. M. *et al*. Adapting agriculture to climate change. *Proc. Natl. Acad. Sci. USA* **104**, 19691–19696 (2007).
28. Ruf, F., Schroth, G. & Doffangui, K. Climate change, cocoa migrations and deforestation in West Africa: What does the past tell us about the future? *Sustain. Sci.* **10**, 101–111 (2015).
29. Fountain, A. & Huetz-Adams, F. *Cocoa barometer 2018*. (Barometer Consortium 2018).
30. Schroth, G., Läderach, P., Blackburn Cuero, D. S., Neilson, J. & Bunn, C. Winner or loser of climate change? A modeling study of current and future climatic suitability of Arabica coffee in Indonesia. *Reg. Environ. Chang.* **15**, 1473–1482 (2015).
31. Peeters, L. Y. K., Soto-Pinto, L., Perales, H., Montoya, G. & Ishiki, M. Coffee production, timber, and firewood in traditional and Inga-shaded plantations in Southern Mexico. *Agriculture, Ecosystems & Environment* **95**, 481–493 (2003).
32. Cannavo, P. *et al*. Agroforestry associating coffee and Inga densiflora results in complementarity for water uptake and decreases deep drainage in Costa Rica. *Agric. Ecosyst. Environ.* **140**, 1–13 (2011).
33. Rosenstock, T. *et al*. Agroforestry with N2-fixing trees: sustainable development's friend or foe? *Curr. Opin. Environ. Sustain.* **6**, 15–21 (2014).
34. Liu, C.-A., Nie, Y., Zhang, Y.-M., Tang, J.-W. & Siddique, K. H. M. Introduction of a leguminous shrub to a rubber plantation changed the soil carbon and nitrogen fractions and ameliorated soil environments. *Sci. Rep.* **8**, 17324 (2018).
35. Schnabel, F. *et al*. Shade trees: a determinant to the relative success of organic versus conventional coffee production. *Agrofor*. *Syst*. 1–15 https://doi.org/10.1007/s10457-017-0100-y (2017).
36. Ordonez, J. C. *et al*. Constraints and opportunities for tree diversity management along the forest transition curve to achieve multifunctional agriculture. *Current Opinion in Environmental*. *Sustainability* **6**, 54–60 (2014).
37. Abdulai, I. *et al*. Cocoa agroforestry is less resilient to sub-optimal and extreme climate than cocoa in full sun. *Glob*. *Chang*. *Biol*. https://doi.org/10.1111/gcb.13885 (2017).
38. Holmgren, M., Scheffer, M. & Huston, M. A. The interplay of facilitation and competition in plant communities. *Ecology* **78**, 1966–1975 (1997).
39. Wu, J., Liu, W. & Chen, C. Below-ground interspecific competition for water in a rubber agroforestry system may enhance water utilization in plants. *Sci. Rep.* **6**, 19502 (2016).
40. Holmgren, M., Scheffer, M., Ezcurra, E., Gutiérrez, J. R. & Mohren, G. M. J. El Niño effects on the dynamics of terrestrial ecosystems. *Trends Ecol. Evol.* **16**, 89–94 (2001).
41. Bonilla Zunhiga, G. *et al*. Database Tipologias cafetaleras en Centroamerica https://doi.org/10.7910/DVN/27826 (2014).
42. Orozco, L., Somarriba, E., Villalobos, M., Deheuvels, O. & Andrade, H. Recopilacion de datos de investigacion del Proyecto Cacao Centroamerica (PCC), https://doi.org/10.7910/DVN/25284 (2014).
43. Sepulveda, N. & Barrios, M. Ensayo de Sistemas Agroforestales con Cafe, Masatepe, Nicaragua, https://doi.org/10.7910/DVN/27915 (2016).
44. CATIE (Centro Agronómico Tropical de Investigación y Enseñanza Costa Rica) & OFI (Oxford Forestry Institute United Kingdom). *Árboles de Centroamérica*. (CATIE 2003).
45. Chízmar-Fernández, C. *et al*. *Plantas comestibles de Centroamérica*. (INBio 2009).
46. Varela, S., Anderson, R. P., García-Valdés, R. & Fernández-González, F. Environmental filters reduce the effects of sampling bias and improve predictions of ecological niche models. *Ecography* **37**, 1084–1091 (2014).
47. GBIF (Global Biodiversity Information Facility Denmark). Global Biodiversity Information Facility. (2018). Available at: http://www.gbif.org/occurrence. (Accessed: 5th May 2018)
48. MAPFORGEN. Atlas for the conservation of forest genetic resources. (2013). Available at: http://www.mapforgen.org. (Accessed: 20th August 2015)
49. Hijmans, R., Guarino, L., Cruz, M. & Rojas, E. Computer tools for spatial analysis of plant genetic resources data: 1. DIVA-GIS. *Plant Genet. Resour. Newsl.* **2001**, 15–19 (2001).
50. Fourcade, Y., Engler, J. O., Rödder, D. & Secondi, J. Mapping Species Distributions with MAXENT Using a Geographically Biased Sample of Presence Data: A Performance Assessment of Methods for Correcting Sampling Bias. *PLoS One* **9**, e97122 (2014).
51. R Core Team. R: A language and environment for statistical computing. version 3.3.3. (2017).
52. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. Species Distribution Modeling. R package version 1.0–15 (2016).
53. Ranjitkar, S. *et al*. Climate modelling for agroforestry species selection in Yunnan Province, China. *Environmental Modelling & Software* **75**, 263–272 (2016).
54. Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R. K. & Thuiller, W. Evaluation of consensus methods in predictive species distribution modelling. *Divers. Distrib.* **15**, 59–69 (2009).
55. Bradley, A. P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition* **30**, 1145–1159 (1997).
56. Freeman, E. A. & Moisen, G. PresenceAbsence: An R package for presence absence analysis. *J. Stat. Softw.* **23**, 31 (2008).
57. Jiménez-Valverde, A. Threshold-dependence as a desirable attribute for discrimination assessment: implications for the evaluation of species distribution models. *Biodiversity and Conservation* **23**, 369–385 (2014).
58. Liu, C., Berry, P. M., Dawson, T. P. & Pearson, R. G. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography (Cop.)*. **28**, 385–393 (2005).
59. Liu, C., White, M. & Newell, G. Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of Biogeography* **40**, 778–789 (2013).
60. Mastrandrea, M. D. *et al*. *Guidance Note for Lead Authors of the IPCC Fifth Assessment Report on Consistent Treatment of Uncertainties*. (Intergovernmental Panel on Climate Change (IPCC) 2010).
61. Bache, S. M. & Wickham, H. magrittr: A Forward-Pipe Operator for R. R package version 1.5. (2014).
62. Wickham, H. tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1 (2017).
63. Bivand, R. *et al*. maptools: Tools for Reading and Handling Spatial Objects. *R package version* **0**, 8–39 (2016).
64. Hijmans, R. J. *et al*. raster: Geographic Data Analysis and Modeling. *R package version* **2**, 5–8 (2015).
65. Bivand, R., Rundel, C., Pebesma, E. & Hufthammer, K. O. rgeos: Interface to Geometry Engine - Open Source (GEOS). *R package version* **0**, 3–19 (2016).
66. Bivand, R., Keitt, T. & Rowlingson, B. rgdal: Bindings for the Geospatial Data Abstraction Library. *R package version* **1**, 1–10 (2016).
67. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag, 2009).
68. Wickham, H., Henry, L., Luciani, T. J., Decorde, M. & Lise, V. svglite: An 'SVG' Graphics Device. (2017).
69. de Sousa, K., van Zonneveld, M., Holmgren, M., Kindt, R. & Ordoñez, J. C. Replication data for: 'The future of coffee and cocoa agroforestry in a warmer Mesoamerica' [Harvard Dataverse]. https://doi.org/10.7910/DVN/0O1GW1

Acknowledgements
Part of this research was supported by cooperative agreement HI-A70D41/A1092 that was made possible by the support of HIVOS International. This work was implemented as part of the CGIAR Research Program on Forests, Trees and Agroforestry (FTA) and the CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), which are carried out with support from the CGIAR Trust Fund and through bilateral funding agreements. For details, please visit https://www.cgiar.org/funders. The views expressed in this document cannot be taken to reflect the official opinions of these organisations.

Author Contributions
J.C.O., M.v.Z., K.d.S., R.K. and M.H. designed research; K.d.S., M.v.Z. and J.C.O. collected the data; K.d.S. and M.v.Z. analysed and processed the data; All authors contributed to writing.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-45491-7.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

3

Farmer evaluating a common bean plot under the *tricot* approach, Honduras. Credit: J. Steinke

Crop variety management for climate adaptation supported by citizen science

Jacob van Etten[a,1,2], Kauê de Sousa[a,b,1], Amílcar Aguilar[c], Mirna Barrios[c], Allan Coto[a], Matteo Dell'Acqua[d], Carlo Fadda[e], Yosef Gebrehawaryat[e], Jeske van de Gevel[f], Arnab Gupta[g,3], Afewerki Y. Kiros[h], Brandon Madriz[a], Prem Mathur[g], Dejene K. Mengistu[e,h], Leida Mercado[i], Jemal Nurhisen Mohammed[h], Ambica Paliwal[g], Mario Enrico Pè[d], Carlos F. Quirós[a], Juan Carlos Rosas[j], Neeraj Sharma[g,4], S. S. Singh[k], Iswhar S. Solanki[l], and Jonathan Steinke[a,m]

[a]Bioversity International, 30501 Turrialba, Costa Rica; [b]Department of Agricultural Sciences, Inland Norway University of Applied Sciences, 2318 Hamar, Norway; [c]Tropical Agricultural Research and Higher Education Center, 61000 Matagalpa, Nicaragua; [d]Scuola Superiore Sant'Anna, Institute of Life Sciences, 56124 Pisa, Italy; [e]Bioversity International, 1000 Addis Ababa, Ethiopia; [f]Bioversity International, 00100 Nairobi, Kenya; [g]Bioversity International, Delhi 110012, India; [h]Department of Dryland Crop and Horticultural Sciences, Mekelle University, 7000 Mekelle, Ethiopia; [i]Tropical Agricultural Research and Higher Education Center, 30501 Turrialba, Costa Rica; [j]Zamorano Panamerican Agricultural School, 11101 Tegucigalpa, Honduras; [k]Indian Council for Agricultural Research–Indian Institute of Wheat and Barley Research, Karnal (Haryana) 132001, India; [l]Indian Agricultural Research Institute, Samastipur (Bihar) 848125, India; and [m]Department of Agricultural Economics, Humboldt University of Berlin, 10099 Berlin, Germany

Edited by B. L. Turner, Arizona State University, Tempe, AZ, and approved January 9, 2019 (received for review August 9, 2018)

Crop adaptation to climate change requires accelerated crop variety introduction accompanied by recommendations to help farmers match the best variety with their field contexts. Existing approaches to generate these recommendations lack scalability and predictivity in marginal production environments. We tested if crowdsourced citizen science can address this challenge, producing empirical data across geographic space that, in aggregate, can characterize varietal climatic responses. We present the results of 12,409 farmer-managed experimental plots of common bean (*Phaseolus vulgaris* L.) in Nicaragua, durum wheat (*Triticum durum* Desf.) in Ethiopia, and bread wheat (*Triticum aestivum* L.) in India. Farmers collaborated as citizen scientists, each ranking the performance of three varieties randomly assigned from a larger set. We show that the approach can register known specific effects of climate variation on varietal performance. The prediction of variety performance from seasonal climatic variables was generalizable across growing seasons. We show that these analyses can improve variety recommendations in four aspects: reduction of climate bias, incorporation of seasonal climate forecasts, risk analysis, and geographic extrapolation. Variety recommendations derived from the citizen science trials led to important differences with previous recommendations.

climate adaptation | genotype × environment interactions | crop variety evaluation | citizen science | crowdsourcing

Crop improvement is important to increase agricultural productivity and to contribute to food and nutrition security. The need for new crop varieties is exacerbated by climate change. Farmers need to replace crop varieties with better-adapted ones to match rapidly evolving climate conditions (1–4). Where suitable modern varieties do not exist, suitable farmer varieties are needed instead ("variety" is applied to all cultivated materials here) (4). The variety replacement challenge has yet to be effectively addressed. One proposed solution is to increase variety supply by accelerating crop breeding, removing older varieties from the seed supply chain, and assiduously promoting new varieties for farmers (2). Supply-driven variety replacement requires that new varieties are locally adapted and acceptable, but varieties are often recommended without prior geographic analysis to determine recommendation domains (5) on the basis of trials that do not adequately represent local production conditions (6–8). Therefore, a supply-driven approach may introduce varieties that perform worse than locally grown varieties. Demand-oriented approaches address this issue but also fall short of a solution. They involve farmers directly in the selection of crop varieties in on-farm experiments (6). Farmer-participatory selection stimulates local interest in new varieties and produces information on variety performance that is immediately relevant to local climate adaptation. This local focus is a strength as well as a limitation. Scaling is constrained by the resource-intensive nature of current participatory experimental methods and the incompatibility of datasets across different efforts (9). The resulting paucity of data is a problem, because variety trials need to capture spatiotemporal environmental variation to characterize climatic responses.

A solution could come from a more scalable type of participatory research: citizen science using digital "crowdsourcing" approaches (10–12). This has already shown its potential to engage large numbers of volunteering citizen scientists who jointly generate sizable datasets that allow for geospatial analysis of climate change impact (for example, on cross-continental

Significance

Climate adaptation requires farmers to adjust their crop varieties over time and use the right varieties to minimize climate risk. Generating variety recommendations for farmers working in marginal, heterogeneous environments requires variety evaluation under farm conditions. On-farm evaluation is difficult to scale with conventional methods. We used a scalable approach to on-farm participatory variety evaluation using crowdsourced citizen science, assigning small experimental tasks to many volunteering farmers. We generated a unique dataset from 12,409 trial plots in Nicaragua, Ethiopia, and India, a participatory variety evaluation dataset of large size and scope. We show the potential of crowdsourced citizen science to generate insights into variety adaptation, recommend adapted varieties, and help smallholder farmers respond to climate change.

Author contributions: J.v.E., M.D., C.F., Y.G., J.v.d.G., A.G., P.M., D.K.M., L.M., M.E.P., J.C.R., N.S., S.S.S., I.S.S., and J.S. designed research; J.v.E., A.A., M.B., Y.G., A.G., A.Y.K., D.K.M., J.N.M., N.S., and I.S.S. performed research; A.C., B.M., and C.F.Q. contributed new reagents/analytic tools; J.v.E., K.d.S., and A.P. analyzed data; and J.v.E. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The data reported in this paper have been deposited in the Dataverse database, https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/4ICF6W.

[1]J.v.E. and K.d.S. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: j.vanetten@cgiar.org.

[3]Present address: Welthungerhilfe, Pathein, Myanmar.

[4]Present address: International Potato Center, Hanoi, Vietnam.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1813720116/-/DCSupplemental.

AGRICULTURAL SCIENCES

SUSTAINABILITY SCIENCE

bird migration) (13). In a similar way, farmer citizen scientists could provide information about crop variety performance, which would feed into a demand-driven, scalable solution to varietal climate adaptation.

To test this idea, we applied a recently developed citizen science approach tricot—triadic comparisons of technologies (14, 15). In tricot variety evaluation, each farmer plants seeds from a personal test package of three varieties, which are randomly assigned from a larger pool of tested varieties. Farmers' independent on-farm observations are compiled and analyzed centrally. A simple ranking-based feedback format allows even farmers with low literacy skills to contribute their evaluation data through various channels, including mobile telephones (15). Pilots with the tricot approach have established its potential to produce accurate data (16) and to engage motivated farmers as citizen scientists (17).

The question that we address is if tricot trials can provide robust, actionable information on varietal climate adaptation. We organized tricot trials to obtain a dataset covering 842 plots of common bean in Nicaragua, 1,090 plots of durum wheat in Ethiopia, and 10,477 plots of bread wheat in India (Fig. 1). The trials captured environmental variation through broad sampling both spatially (many fields distributed across the landscape) and temporally (different seasons and planting dates). We linked farmers' observations via their geographic coordinates and planting dates to agroclimatic and soil variables. We modeled the influence of the environmental variables on the probability that varieties outperform the other varieties in the trials. We evaluated whether seasonal climate adequately predicts variety performance in the tricot trials. Then, we explored if climatic analysis of tricot trial data improves variety recommendations.

Characterizing Variety Performance

Cross-validation showed that the tricot trials uncovered statistically robust differences in variety performance (Table 1). From a previous pilot study, we expected consistently positive, but low to moderate, pseudo-R^2 values (16). In this study, model fit was comparatively low for bread wheat in India (0.04–0.09), moderate for common bean in Nicaragua (0.15–0.20), and high for durum wheat in Ethiopia (0.39–0.48). The three case studies each provide independent confirmation of the predictive value of the tricot trials. Various factors influenced model fit, including farmers' observation skills and environmental variation. The largest differences were between countries, which were probably due to the different levels of diversity within the sets of varieties. Indian and Nicaraguan farmers evaluated a small, carefully selected group of modern varieties with relatively homogeneous performance. In Ethiopia, farmers tested a diverse set of modern and farmer varieties drawn from a wide area and evidently found easily observable differences in performance between varieties.

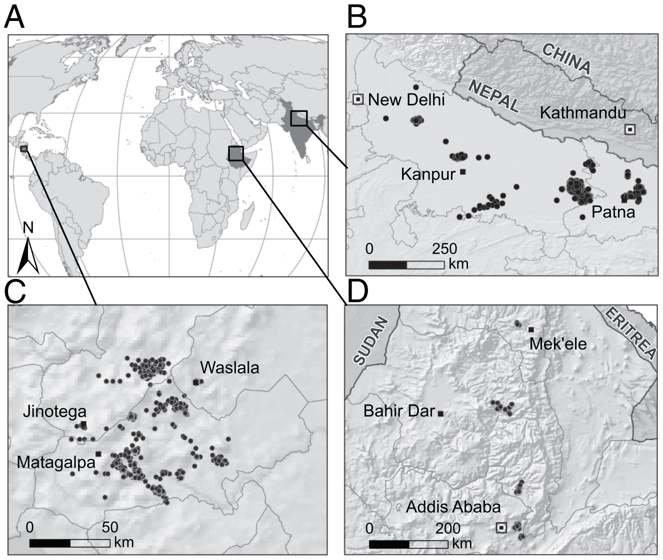

Fig. 1. Research sites: (*A*) overview, (*B*) India, (*C*) Nicaragua, and (*D*) Ethiopia. Farms included in the trials are indicated as dots.

Table 1. Goodness of fit (pseudo-R^2) of PLTs determined with 10-fold cross-validation

PLT model	Nicaragua	Ethiopia	India
No covariates	0.1484	0.3947	0.0381
Design	0.1869	0.4709	0.0721
Climate	**0.1978**	**0.4870**	**0.0882**
Climate + geolocation	0.1977	0.4720	0.0872

The model with only climate covariates has the best fit in all cases (indicated in bold).

For each country, we modeled the environmental influence on variety performance. We were specifically interested in models with covariates derived from seasonal climatic conditions (climate in Table 1), because these covariates can potentially enhance extrapolation of variety performance predictions across time and space. In all cases, these models had indeed a better fit than the respective model without environmental covariates (no covariates in Table 1). The next question that we addressed was if the models with climatic variables captured the main environmental factors or missed important aspects. Therefore, we compared these models with two other types of models. One type of model includes covariates that represent the experimental design and are known in advance: geolocation, season, planting dates, and soil categories (design in Table 1). These models reflect how multilocation trials are often analyzed and capture variation in terms of the trial structure but not in terms of the underlying climatic causal factors, hence limiting the potential of extrapolation beyond the trial. In all cases, the models with climatic covariates slightly outperformed the models with trial design covariates. This means that the climatic covariates contain unique and substantial information explaining varietal performance. A second comparison was with models that include the climatic covariates together with additional covariates that represent geographic structure (climate + geolocation in Table 1). This comparison tested if important local factors are being overlooked that are not covered by the climatic covariates. Adding these geolocational variables did not improve the models, however, and even slightly degraded them. This implies that no large-scale geographical structure remained after accounting for seasonal climate. From this analysis, it is clear that the models with climatic covariates captured a large part of the environmental variation in variety performance. Therefore, in subsequent analyses, we focused on models with climatic covariates only.

We generated generalizable models that afford extrapolation across seasons of variety performance predictions by selecting those climatic variables that contribute to predictivity across seasons. The variable selection procedure retained one climatic variable in each case (Fig. 2 and *SI Appendix*, Fig. S1). We discuss the results for each case study.

For Nicaragua, Fig. 2 shows the Plackett–Luce tree (PLT) with the retained variable of the generalizable model for common bean. We found that bean variety performance changed when the maximum night temperature exceeded 18.7 °C. This finding corresponds to the threshold temperature for heat stress reported in the literature of 20 °C at night (18). Our estimate is slightly lower than the reported threshold but refers to land surface temperature rather than air temperature. Three

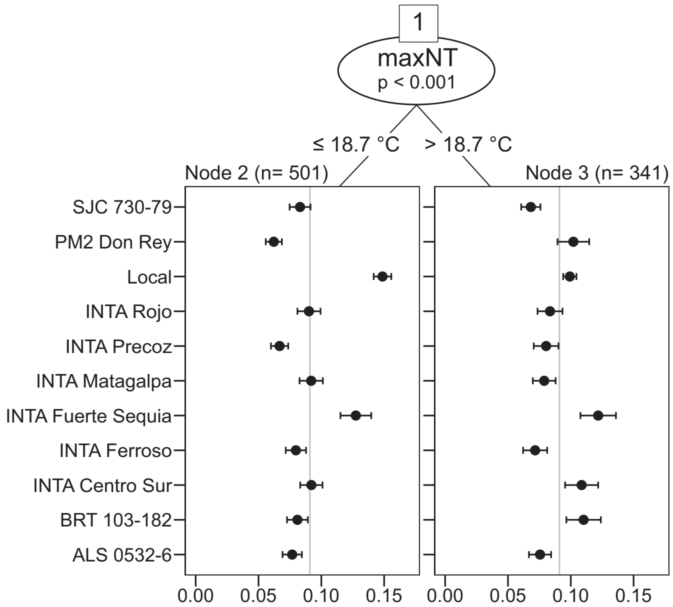

Fig. 2. Plackett–Luce trees of tricot trial data and associated climatic data for common bean in Nicaragua. The horizontal axis of each panel is the probability of winning of varieties. Error bars show quasi-SEs. The gray vertical lines indicate the average probability of winning (1/number of varieties). In this case, the model selected maxNT, the maximum night temperature (degrees Celsius) during the vegetative and flowering periods, as the covariate. Equivalent figures for the trials in Ethiopia and India are shown in *SI Appendix*, Figs. S2 and S3.

bean varieties, INTA Fuerte Sequía, BRT 103–182, and INTA Centro Sur, performed better than the local variety above the heat stress threshold. These three varieties are known to be heat tolerant. Contrary to our expectations, another heat-tolerant variety, SJC 730–79, did not show a performance advantage above 18.7 °C. Above higher-heat stress thresholds, however, this variety did outperform heat-sensitive varieties. The group of local varieties has a small quasi-SE, despite the diverse nature of this group, which contains all varieties that farmers currently grow. This gives a reference on farmers' overall appreciation of the tested varieties in comparison with their own varieties. The local varieties were outperformed under heat stress but not under cool conditions.

For durum wheat in Ethiopia, varietal differences in performance were related to the lowest night temperature during the vegetative period (*SI Appendix*, Fig. S2). Performance patterns changed when at least one 8-day period had average night temperatures under 8.4 °C. This temperature corresponds to the threshold temperatures for vernalization and cold acclimation induction (19). Under warm conditions, vernalization-requiring varieties will delay flowering. Under cold conditions, cold-sensitive varieties will reduce their yield due to chilling or frost damage. Most of the varieties tested in Ethiopia were farmer varieties and likely adapted to their original environments, which may have led to differences in adaptiveness between varieties. To test the effect of local adaptation, we compared cold-adapted varieties with cold-sensitive farmer varieties as detected by the tricot trials (*Materials and Methods*). Cold-adapted varieties came from higher elevations (2,483 ± 113 meters above sea level) than cold-sensitive ones (2,101 ± 485 meters), a significant difference $[t(594) = 16.1, P < 2.2 \cdot 10^{-16}]$. Our results indicate that cold tolerance is a main geographic adaptation factor for durum wheat in the Ethiopian highlands.

For bread wheat in India, varietal performance patterns changed with the diurnal temperature range (DTR) during the vegetative period, which is the difference between minimum and maximum daily temperatures (*SI Appendix*, Fig. S3). Splits occurred at DTR values of 14.5 °C and 15.7 °C. Between these two values, the varieties showed very similar performance. Many varieties that performed above average under high DTR performed below average under low DTR and vice versa. Some varieties performed well under both high and low DTR, especially HD 2967. Our interpretation is that low and high ranges of DTR are related to different sets of stresses, while the middle range has relatively low stress. DTR has an impact on crop yield through several mechanisms: high DTR is associated with increased heat or cold stress, and low DTR is associated with high cloud coverage, low solar radiation, and high rainfall. Consistent with our results, a study has shown that DTR explains a substantial share of wheat yield variation in India (20). This same study found that DTR has a negative correlation with wheat yields in some areas and a positive correlation in other areas, in line with high and low DTRs having an association with different types of crop stress.

Improving Variety Recommendations

We examined four ways in which climatic analysis afforded by tricot trials can improve variety recommendations. First, a potential improvement is that climatic analysis corrects the climatic sampling bias, a bias that occurs when trials are performed under unrepresentative seasonal climate conditions, thereby degrading variety recommendations. To assess the importance of climatic sampling bias, we followed the cross-validation procedure used to generate the generalizable models but did not use the seasonal climate data for predictions. Instead, we predicted variety performance for a representative 15-y base period of seasonal climate data and averaged the results (average season in Table 2). The averaged prediction had slightly higher pseudo-R^2 values than the "no covariates" model in all cases. This analysis shows that, even when climatic sampling bias is low, correction can help to further improve predictions.

Second, climatic analysis can improve variety recommendations by incorporating seasonal forecasts. Perfect forecast in Table 2 shows that the pseudo-R^2 values increase further when observed climate information is available for prediction. The improvement gained from a perfect forecast was substantially larger than the improvement from sampling bias correction. It requires additional work to quantify the improvement of variety recommendations with a realistic climate forecast skill. It is clear, however, that variety recommendations derived from tricot trials can benefit from seasonal forecasts.

Third, climatic analysis can support risk analysis. Table 3 shows the expected probability of outperforming all other varieties, which is a metric of average performance, and a risk metric, worst regret (21)—the largest underperformance of the recommended variety relative to the best variety. These two metrics produced divergent variety recommendations in all three cases (indicated in bold in Table 3). In principle, risk analysis for variety choice is also possible without explicit climatic analysis, but this produces results that are difficult to interpret in terms of climatic causality and requires trials during a large number of

Table 2. Goodness of fit (pseudo-R^2) of generalizable PLT models

Model	Nicaragua	Ethiopia	India
No covariates	0.1533	0.4280	0.0611
Average season	0.1536	0.4290	0.0694
Perfect forecast	0.1749	0.4442	0.1065

Model average season corrects for climatic sampling bias by averaging predictions over a base period of seasonal climate data. Model perfect forecast uses observed climatic covariates in the predicted seasons. Values represent cross-validated pseudo-R^2 values averaged across blocks and weighted with the square root of the sample size of each block.

Table 3. Expected probability of winning (average of all farms over the base period) and worst regret measures of a subset of the varieties

Case study and variety	Probability of winning	Worst regret
Common bean (Nicaragua)		
Local variety	**0.130**	0.023
INTA Fuerte Sequía	0.125	**0.021**
INTA Centro Sur	0.098	0.057
BRT 103-182	0.092	0.068
INTA Rojo	0.088	0.082
INTA Matagalpa	0.087	0.057
Durum wheat (Ethiopia)		
208279	**0.059**	0.062
Hitosa	0.049	**0.035**
208304	0.041	0.048
8034	0.030	0.053
Ude	0.025	0.063
222360	0.023	0.061
Bread wheat (India)		
K 9107 (Deva)	**0.077**	0.051
HD 2967	0.068	0.047
HD 2733	0.066	**0.036**
K 0307 (Shatabadi)	0.063	0.095
CSW 18	0.042	0.073
HI 1563 (Pusa Prachi)	0.041	0.093

The results show how different criteria of variety selection can lead to different recommendations (best value according to each criterion is indicated in bold). Using the probability of winning as a criterion maximizes the average performance but ignores risk. Minimizing worst regret (the loss under the worst possible outcome) is a criterion that takes a conservative approach to risk.

seasons to avoid sampling bias and to characterize probability distributions accurately (22).

Fourth, climatic analysis of tricot trial data can generate variety recommendations for wider areas through geospatial extrapolation. To illustrate this, we generated maps of varieties recommendations based on "average season" model predictions (Fig. 3). In all three cases, geographical patterns of variety adaptation have no relationship to administrative boundaries or agroecological zones, which are commonly used to delineate recommendation domains.

To assess what the tricot trial results mean in practice, we contrast our results with existing recommendations. For Nicaragua, we compare the results of the tricot trials with the recommendations of a recent national variety catalog (23). The catalog recommends INTA Rojo and INTA Matagalpa for the study area, but these varieties performed worse than the local varieties in the tricot trials (Fig. 3*A*). However, the tricot trials identified INTA Fuerte Sequía and INTA Centro Sur as top varieties (Table 3), but the variety catalog recommends them for warm areas outside our study area. In the tricot trials, INTA Fuerte Sequía and INTA Centro Sur outperformed other varieties, especially under heat stress, which apparently occurs with more frequency in our study area than assumed by current variety recommendations. In Nicaragua, then, the tricot trial results show that official variety recommendations fail to identify superior bean varieties that are sufficiently heat tolerant for the study area.

For Ethiopia, the *Wheat Atlas* of the International Maize and Wheat Improvement Center (CIMMYT) recommends modern varieties Hitosa, Ude, and Assassa for all of the Ethiopian highlands, which it classifies as a single "mega-environment" (24). The tricot approach produced geographically more specific recommendations (Fig. 3*B*). With this, we confirm the results of a previous analysis based on multilocational trial data that showed the benefits of location-specific recommendation domains for durum wheat in Algeria, and we show that such an analysis can also be done with tricot data (25). The tricot results confirmed the superiority of farmer varieties 8208 and 208304 (Table 3), which were approved for official variety release in March 2017 (on the basis of other field trials) (26). Farmer variety 208279 also has a high probability of winning, but it has a high value of worst regret (Table 3). Our analysis suggests that 208279 could be considered for the coldest areas as shown in Fig. 3*B*. In Ethiopia, the tricot trial findings improve variety recommendations for durum wheat by uncovering the importance of cold adaptation.

For India, we compare our findings with the front-line demonstrations of the Indian Institute for Wheat and Barley Research (IIWBR); the 1-ha plots demonstrate new varieties by comparing them with a check variety. IIWBR promoted the variety HD 2967 for the North-Eastern Plain Zone during 2016–2017 (27). HD 2967 was indeed the top variety in the tricot trial among the varieties considered by the IIWBR (Table 3). In the tricot trials, however, K 9107 (a variety released in 1996) outperformed HD 2967 (released in 2011), with a comparable level of worst regret (Table 3). The tricot trials also showed that another variety, HD 2733, outperformed HD 2967 in a large part of the study area (Table 3). In the IIWBR front-line demonstrations, HD 2733 was included as a check variety in four areas and was outyielded by HD 2967 in only one of four areas, while in the other three, the yield difference was not significant (27). Our analysis shows that HD 2733 generally does better than HD

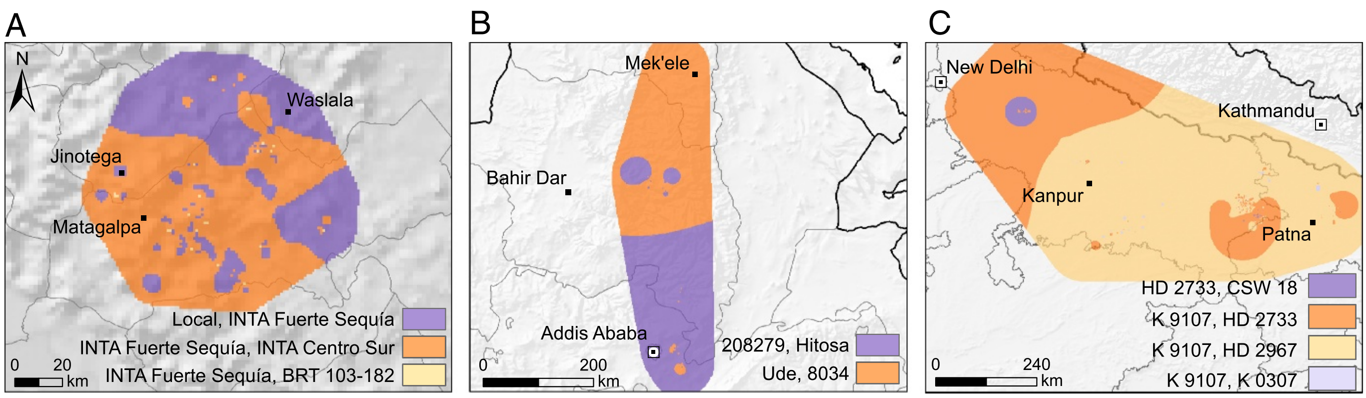

Fig. 3. Variety recommendations based on average season predictions from PLTs using climatic variables for (*A*) common bean in Nicaragua (Apante season), (*B*) durum wheat in Ethiopia (Meher season), and (*C*) bread wheat in India (Rabi season). Map categories show the top two varieties for each area according to their probability of winning over a base period (2002–2016).

2967 in areas with a low average DTR during the growing season (Fig. 3*C*). In India, the analysis of the tricot trial data adds geographic specificity to the existing variety recommendations and suggests that a broader set of wheat varieties should be promoted to take into account the climatic differences across the study area.

We quantified how much farmers can benefit from tricot-based variety recommendations by calculating variety reliability, the probability of outperforming a check variety (Eq. **2** in *Materials and Methods*). For each location, we compared the tricot-recommended variety (Fig. 3) with the best-performing variety from the previous recommendations as the check. Reliabilities ranged from 0.59 to 0.65 in Ethiopia, from 0.58 to 0.60 in Nicaragua, and from 0.51 to 0.62 in India (*SI Appendix*, Fig. S4), indicating substantial benefits for large areas.

Conclusions

The main question that we addressed is whether on-farm participatory crop trials, scaled through a farmer citizen science approach, can generate insights into climate adaptation of varieties. Citizen science data revealed generalizable relations between seasonal climate variables and crop variety performance that corresponded to known yield-determining factors. Climatic analyses of these data were shown to improve variety recommendations. Our study demonstrates that, in vulnerable, low-income areas, climatic analysis of variety performance is possible with trial data generated directly by farmer citizen scientists on farms. Arguably, similar results could be achieved by a combination of existing approaches (target environment characterization, multilocation trials, participatory variety selection, variety dissemination). The unique contribution of the tricot approach is that it integrates aspects of these approaches into a simple trial format that addresses the challenge of variety replacement for climate adaptation in a way that is, at the same time, scalable and demand led. Tricot trials can track climate trends as they manifest themselves on farms, adjust variety recommendations and recommendation domains, and contribute to understanding how climate affects on-farm varietal performance. Trial analysis combines insights in climatic adaptation mechanisms with a comprehensive evaluation of variety performance from the perspective of farmers, the end users of the seeds. Results can, therefore, be directly translated into actionable information for climate adaptation on the ground. The findings can serve to create variety portfolios that diminish climate risk (22), can feed into climate information services in combination with seasonal forecasts (28), and can become part of decentralized plant breeding strategies for climate adaptation (8). Combining the tricot trial data with other data could generate additional insights into variety performance and acceptability as influenced by environmental (11), socioeconomic (29), and genomic (30) factors.

The tricot approach facilitates engaging large numbers of farmers in citizen science trials with large sets of varieties. Scaling does not only involve an expansion in terms of numbers and scope, however, but also, it implies new institutional arrangements. Carefully designed strategies should foster communication between providers and users of information (31). Wide-ranging collaborations are needed for climate adaptation in crop variety management, involving farmers, extension agents, seed retailers, seed producers, plant breeders, and climate information providers. The tricot approach can help to cut across these different domains, because it is able to link climatic and varietal information directly to farmer decision making. With appropriate institutional support and investment, citizen science can potentially make an important contribution to farmers' adaptive capacity and to the mobilization of crop genetic diversity for climate adaptation.

Materials and Methods

Crop Trials. Trials were performed between 2012 and 2016 during three cropping seasons in Ethiopia, five cropping seasons in Nicaragua, and four cropping seasons in India (*SI Appendix*, Table S1). Trial design followed the tricot citizen science approach (14, 15). Sets of varieties were allocated randomly to farms as incomplete blocks (7), maintaining spatial balance by assigning roughly equal frequencies of the varieties to each area. In Nicaragua and India, incomplete blocks contained three varieties. In Ethiopia, we used a modified approach that included four varieties per farm. Plots were small to facilitate farmer participation but in all cases, large enough to avoid strong edge effects. Farmers indicated the relative performance of varieties through ranking. Ranking is a robust data collection approach that avoids observer drift (32) and allows for aggregation across disparate datasets (33).

The trials required three moments of contact with the farmers: (*i*) explaining the experiment and distributing the seeds, (*ii*) collecting evaluation data, and (*iii*) returning the results. Data were initially collected using paper forms and in subsequent seasons, through electronic formats linked to a purpose-built digital platform, https://climmob.net. In the trials presented here, field agents collected the data through visits (phone calls are also feasible).

Data Analysis. All analyses were done in R (34). For the analysis of the variety-ranking data generated by farmers, we used the Plackett–Luce model (35, 36). The Plackett–Luce model estimates for each variety the probability that it wins, beating all other varieties in the set. The model determines the values of positive-valued parameters α_i (worth) associated with each variety i. These parameters α are related to the probability that variety i wins against all other n varieties in the following way:

$$P\left(i \succ \{j, \ldots, n\}\right) = \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_n}. \qquad [1]$$

The probability that variety i beats another variety j is calculated in a similar way.

$$P(i \succ j) = \frac{\alpha_i}{\alpha_i + \alpha_j}. \qquad [2]$$

Eq. **2** also serves to calculate the reliability of a variety—its probability of beating a check variety (37). These equations follow from Luce's Choice Axiom, which states that the probability that one item beats another is independent from the presence or absence of any other items in the set (36). We report worth values that sum to one. This makes each worth value α_i equal to the probability of variety i outperforming all other varieties:

$$P\left(i \succ \{j, \ldots, n\}\right) = \frac{\alpha_i}{\alpha_1 + \cdots + \alpha_n} = \frac{\alpha_i}{1} = \alpha_i. \qquad [3]$$

In the trials, we used rankings of three varieties ($i \succ j \succ k$), which have the following probability of occurring according to the Plackett–Luce model:

$$P(i \succ j \succ k) = P(i \succ \{j, k\}) \cdot P(j \succ k). \qquad [4]$$

The log likelihood for a ranking $i \succ j \succ k$ follows from Eqs. **1**, **2**, and **4** and takes the following form (38):

$$\begin{aligned} \ell(\alpha) &= \ln(P(i \succ \{j, k\})) + \ln(P(j \succ k)) \\ &= \ln(\alpha_i) - \ln(\alpha_i + \alpha_j + \alpha_k) + \ln(\alpha_j) - \ln(\alpha_j + \alpha_k). \end{aligned} \qquad [5]$$

The log likelihood is then the sum of the log-likelihood $\ell(\alpha)$ values across all rankings. Using an iterative algorithm, the log likelihood is maximized to identify the α values that make the observed rankings most probable. We also generated quasi-SEs for α (39). To take into account covariates, we created PLTs through recursive partitioning (40). Additional details are given in *SI Appendix*.

Data and Code Availability. Full data are available through Dataverse (41). Code is available in *SI Appendix*.

ACKNOWLEDGMENTS. We thank all farmers who evaluated varieties in Nicaragua, Ethiopia, and India. We also thank Vincent Johnson and Olga Spellman for editorial support and Heather Turner for support on scientific programing. Part of this research was supported by Cooperative Agreement AID-OAA-F-14-00035, which was made possible by the generous support of the American people through the US Agency for International Development. The research received financial support from McKnight Foundation Grant CCRP 16–098, German Federal Ministry for Economic Cooperation and Development Contract 81194988, and the Indian Council of

AGRICULTURAL SCIENCES

SUSTAINABILITY SCIENCE

Agricultural Research Annual Workplan. This work was implemented as part of the CGIAR Research Program on Climate Change, Agriculture and Food Security, which is carried out with support from the CGIAR Trust Fund and through bilateral funding agreements (details are at https://ccafs.cgiar.org/donors). The views expressed in this document cannot be taken to reflect the official opinions of these organizations.

1. Porter JR, et al. (2014) Food security and food production systems in climate change 2014: Impacts, adaptation, and vulnerability. *Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds Field CB, et al. (Cambridge Univ Press, Cambridge, UK), pp 485–533.
2. Atlin GN, Cairns JE, Das B (2017) Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change. *Glob Food Secur* 12:31–37.
3. Challinor AJ, Koehler AK, Ramirez-Villegas J, Whitfield S, Das B (2016) Current warming will reduce yields unless maize breeding and seed systems adapt immediately. *Nat Clim Change* 6:954–958.
4. Bellon MR, Hodson D, Hellin J (2011) Assessing the vulnerability of traditional maize seed systems in Mexico to climate change. *Proc Natl Acad Sci USA* 108:13432–13437.
5. Annicchiarico P (2002) *Genotype x Environment Interactions: Challenges and Opportunities for Plant Breeding and Cultivar Recommendations* (Food & Agriculture Organization, Rome).
6. Dawson JC, Murphy KM, Jones SS (2008) Decentralized selection and participatory approaches in plant breeding for low-input systems. *Euphytica* 160:143–154.
7. Atlin G, Cooper M, Bjørnstad Å (2001) A comparison of formal and participatory breeding approaches using selection theory. *Euphytica* 122:463–475.
8. Ceccarelli S (2015) Efficiency of plant breeding. *Crop Sci* 55:87–97.
9. Nelson R, Coe R, Haussmann BI (June 3, 2016) Farmer research networks as a strategy for matching diverse options and contexts in smallholder agriculture. *Exp Agric*, 10.1017/S0014479716000454.
10. Bonney R, et al. (2009) Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience* 59:977–984.
11. Minet J, et al. (2017) Crowdsourcing for agricultural applications: A review of uses and opportunities for a farmsourcing approach. *Comput Electronics Agric* 142:126–138.
12. Ryan S, et al. (2018) The role of citizen science in addressing grand challenges in food and agriculture research. *Proc R Soc B* 285:20181977.
13. Cooper CB, Shirk J, Zuckerberg B (2014) The invisible prevalence of citizen science in global research: Migratory birds and climate change. *PLoS One* 9:e106508.
14. van Etten J (2011) Crowdsourcing crop improvement in sub-saharan Africa: A proposal for a scalable and inclusive approach to food security. *IDS Bull* 42:102–110.
15. van Etten J, et al. (October 17, 2016) First experiences with a novel farmer citizen science approach: Crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Exp Agric*, 10.1017/S0014479716000739.
16. Steinke J, van Etten J, Zelan PM (2017) The accuracy of farmer-generated data in an agricultural citizen science methodology. *Agron Sustainable Dev* 37:32.
17. Beza E, et al. (2017) What are the prospects for citizen science in agriculture? Evidence from three continents on motivation and mobile telephone use of resource-poor farmers. *PLoS One* 12:e0175700.
18. Rainey KM, Griffiths PD (2005) Inheritance of heat tolerance during reproductive development in snap bean (Phaseolus vulgaris L.). *J Am Soc Hortic Sci* 130:700–706.
19. Fowler DB (2008) Cold acclimation threshold induction temperatures in cereals. *Crop Sci* 48:1147–1154.
20. Rao BB, Chowdary PS, Sandeep V, Pramod V, Rao V (2015) Spatial analysis of the sensitivity of wheat yields to temperature in India. *Agric For Meteorol.* 200:192–202.
21. Savage LJ (1951) The theory of statistical decision. *J Am Stat Assoc* 46:55–67.
22. Sukcharoen K, Leatham D (2016) Mean-variance versus mean-expected shortfall models: An application to wheat variety selection. *J Agric Appl Econ* 48:148–172.
23. López Silva E, Munguía Sandoval J (2014) *Catálogo de cultivares de granos básicos registrados y autorizados para la producción y comercialización* (Departamento de Semillas MAGFOR-DGPSA and INTA, Managua, Nicaragua).
24. CIMMYT *Wheat Atlas*. Available at beta.wheatatlas.org. Accessed July 15, 2018.
25. Annicchiarico P, Bellah F, Chiari T (2006) Repeatable genotype × location interaction and its exploitation by conventional and GIS-based cultivar recommendation for durum wheat in Algeria. *Eur J Agron* 24:70–81.
26. Mengistu DK, Kidane YG, Fadda C, Pè ME (2018) Genetic diversity in Ethiopian durum wheat (Triticum turgidum var durum) inferred from phenotypic variations. *Plant Genet Resour* 16:39–49.
27. ICAR-IIWBR (2017) *Progress Report of AICRP on Wheat and Barley 2016–17, Social Sciences* (ICAR–Indian Institute of Wheat and Barley Research, Karnal, Haryana, India).
28. Klemm T, McPherson RA (2017) The development of seasonal climate forecasting for agricultural producers. *Agric For Meteorol* 232:384–399.
29. Hammond J, et al. (2017) The Rural Household Multi-Indicator Survey (RHoMIS) for rapid characterisation of households to inform climate smart agriculture interventions: Description and applications in East Africa and Central America. *Agric Syst* 151:225–233.
30. Kidane YG, et al. (2017) Genome wide association study to identify the genetic base of smallholder farmer preferences of durum wheat traits. *Front Plant Sci* 8:1230.
31. Hewitt CD, Stone RC, Tait AB (2017) Improving the use of climate information in decision-making. *Nat Clim Change* 7:614–616.
32. Halekoh U, Kristensen K (2008) Evaluation of treatment effects by ranking. *J Agric Sci* 146:471–481.
33. Simko I, Piepho HP (2011) Combining phenotypic data from ordinal rating scales in multiple plant experiments. *Trends Plant Sci* 16:235–237.
34. R Core Team (2017) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna), Version 3.4.3.
35. Plackett RL (1975) The analysis of permutations. *J R Stat Soc Ser C Appl Stat* 24:193–202.
36. Luce RD (1959) *Individual Choice Behavior: A Theoretical Analysis* (Wiley, New York).
37. Eskridge K, Mumm R (1992) Choosing plant cultivars based on the probability of outperforming a check. *Theor Appl Genet* 84:494–500.
38. Hunter DR (2004) MM algorithms for generalized Bradley-Terry models. *Ann Stat* 32:384–406.
39. Turner HL, van Etten J, Firth D, Kosmidis I (2018) Modelling rankings in R: The PlackettLuce package. arXiv:1810.12068.
40. Strobl C, Wickelmaier F, Zeileis A (2011) Accounting for individual differences in Bradley-Terry models by means of recursive partitioning. *J Educ Behav Stat* 36:135–153.
41. van Etten J, et al. (2018) Replication data for "Crop variety management for climate adaptation supported by citizen science." Harvard Dataverse. Available at https://doi.org/10.7910/DVN/4ICF6W. Deposited July 27, 2018.

4

Durum wheat in a participatory variety selection plot in Tigray, Ethiopia. Credit: M. Dell'Acqua

Data-driven decentralised breeding increases genetic gain in a challenging crop production environment

Kauê de Sousa[1,2], Jacob van Etten[2], Jesse Poland[3], Carlo Fadda[2], Jean-Luc Jannink[4,5], Yosef Gebrehawaryat Kidane[6,7], Basazen Fantahun Lakew[6,8], Dejene Kassahun Mengistu[6,7], Mario Enrico Pè[7,9], Svein Øivind Solberg[1], Matteo Dell'Acqua[7,9][*]

[1] Department of Agricultural Sciences, Inland Norway University of Applied Sciences, 2318 Hamar, Norway
[2] Bioversity International, 00054 Maccarese, Rome, Italy
[3] Department of Plant Pathology, Kansas State University, KS 66506 Manhattan, USA
[4] College of Agriculture and Life Sciences, Cornell University, NY 14853 Ithaca, USA
[5] Agricultural Research Service, United States Department of Agriculture, NY 14853 Ithaca, USA
[6] Bioversity International, Ethiopia Office, 1000 Addis Ababa, Ethiopia
[7] Institute of Life Sciences, Scuola Superiore Sant'Anna, 56124 Pisa, Italy
[8] Ethiopian Biodiversity Institute, 1000 Addis Ababa, Ethiopia
[9] Center for Climate Change Studies and Sustainable Actions (3CSA), Pisa, Italy
[*]Correspondence should be addressed to: m.dellacqua@santannapisa.it

Crop breeding must embrace the broad diversity of smallholder agricultural systems to ensure food security to the hundreds of millions of people living in marginal production environments. This challenge can be addressed by combining genomics, farmers' knowledge, and environmental analysis into a data-driven decentralised approach (3D-breeding). We tested this idea as a proof-of-concept by comparing a durum wheat (*Triticum durum* Desf.) decentralised trial distributed as incomplete blocks in 1,165 farmer-managed plots across the Ethiopian highlands with a benchmark representing genomic selection applied to conventional breeding. We found that 3D-breeding could double the accuracy of the benchmark. 3D-breeding could identify genotypes with enhanced local adaptation providing consistent yield advantages across seasons and locations. We propose this decentralised approach to leverage the diversity in farmers' fields and complement conventional plant breeding to enhance local adaptation in challenging crop production environments.

Introduction

The big data revolution in genomic tools has transformed plant breeding with inexpensive sequencing methods, enabling greatly accelerated variety development[1–3]. At present, plant breeders use data-driven methods, including genomic selection, to increase selection intensity while reducing the time of the breeding cycle and deriving greater genetic gain[4]. Most conventional breeding programs still rely on a centralised scheme aimed at maximizing genetic diversity (G) in the early stages of selection and then identifying superior germplasm on the basis of phenotypic observations made in a limited number of research stations with explicit environmental (E) and management (M) conditions. In this setting, genomic selection may be used to predict the performance of untested new genotypes but is bound to the $G \times E \times M$ interactions captured by the research stations that are used to train the selection models[5]. This limitation of centralised breeding approaches may result in sub-optimal development and deployment of crop varieties for use by farmers seeking local adaptation in challenging environments[6].

To respond to local needs impacted by climate change, breeders need to find new ways to accelerate variety development while directly addressing $G \times E \times M$ interactions to the fullest[3,7,8]. Mobilizing farmers' traditional knowledge of crop varieties and local adaptation can address this challenge[6,9,10] in a coherent, decentralised breeding program relying on farmer-participatory selection[11–13]. A crowdsourced citizen science approach has demonstrated the feasibility of a data-driven decentralised variety evaluation[14] that enables on-farm variety testing in a digitally supported and cost-efficient way[15]. Predictive accuracy of farmer selection criteria may outperform breeder evaluations even in a context of modern agriculture[16].

Crowdsourced citizen science further integrates the E and M components into breeding by performing selection directly in target environments and using environmental data to analyse genotypic responses. Thus, the citizen science approach scales E and M data collection to generate a volume of data that matches the big data dimension of G. Combining genomic selection with citizen science opens the possibility of simultaneously capturing the three dimensions of crop performance, G, E, and M, in a data-driven way. Here, we describe and demonstrate potential benefits of this approach that we call *data-driven decentralised breeding*, or 3D-breeding, for short. Potentially, 3D-breeding could benefit the ~500 million smallholder farmers around the world who often produce in marginal, low-input environments and work with diverse cropping and farming systems and respond to local consumption preferences[17].

We tested the 3D-breeding approach in the Ethiopian highlands, where many smallholder farmers grow durum wheat (*Triticum durum* Desf.) and select landraces following criteria related to environmental adaptation, food culture, and market demand[18,19]. Rich local wheat diversity has co-evolved with local cultures and landscapes over millennia. Consequently, Ethiopian farmers still often select and cultivate local landraces, which under local conditions tend to outperform modern varieties produced by centralised breeding[20]. In this context, 3D-breeding can leverage local wheat diversity and knowledge, and bring breeding closer to the target environments cutting through the complexity of $G \times E \times M$.

We compared 3D-breeding with a centralised breeding approach. In a centralised trial, we collected phenotypic data and farmer evaluations on a panel of 400 fully genotyped durum wheat lines derived from genebank accessions[18] in two managed fields commonly used for varietal selection in the Ethiopian highlands in 2012 and 2013. In a decentralised trial, we distributed a subset of 41 genotypes (Fig. S1) as packaged sets containing incomplete blocks of three genotypes, plus one commercial variety to each farmer, following the *tricot* citizen science approach[15]. A total of 1,165 farmers planted these packages on their farms across three administrative regions of Ethiopia (Fig. S2). We use the data from the centralised and decentralised trials to evaluate whether 3D-breeding could complement genomics assisted breeding by increasing prediction accuracy in challenging environments.

Results

Benchmark: centralised breeding enhanced by genomic selection and farmer participation

We established a benchmark that represents a centralised breeding approach that is a competitive alternative to 3D-breeding. We focused on grain yield (GY) and farmers' overall appreciation of genotypes (OA), which were both recorded in centralised (station) and decentralised (farm) trials. Centralised stations and farmer fields belong to the same agroecological zones of Ethiopia (Fig. S3). To establish the benchmark, we used a genomic selection model trained on $GY_{STATION}$ and $OA_{STATION}$ to predict, respectively, GY_{FARM} and OA_{FARM} in 1,165 farmer fields located in the same breeding mega-environment (Fig. 1A). The benchmark represents a centralised breeding approach using farmer on-station selection. Its scope and size reflect a regional variety trial, an advanced stage in breeding that focuses on a set of genetic materials and target environments with the aim of selecting the best genotypes for varietal release and recommendation. The stations are commonly used as breeding field trials for Amhara and Tigray regions of Ethiopia, and differ in altitude, temperature, rainfall, and soil[20]. Additional multilocation trials would typically occur in earlier stages of the breeding cycle. On-station involvement of farmers is not common practice but is increasingly conducted in association with breeding[12,16] and makes the benchmark more competitive. 3D-breeding expands on the centralised genotype characterisation by moving the selection to farmer fields (Fig. 1B).

Heritability (H^2), the proportion of phenotypic variance explained by genotypic variance, was 0.55 and 0.42 for $GY_{STATION}$ across locations for 2012 and 2013 respectively (Table S1). Men

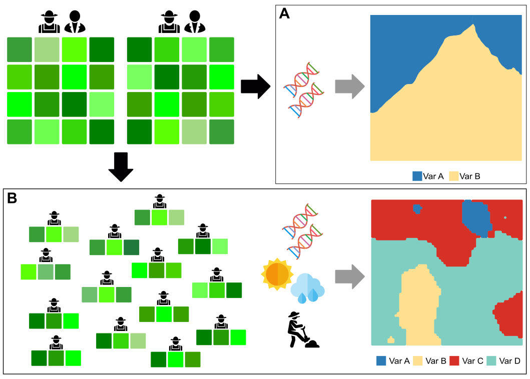

Fig. 1. Centralised breeding (**A**) derives recommendations from breeders' evaluation and possibly participatory assessments in a limited set of stations, using genomics to accelerate the production of varieties that are eventually recommended with coarse spatial resolution. This system may become more efficient if complemented by 3D-breeding (**B**), a decentralised approach where the best candidate genotypes are tested by farmers in small, blinded and randomized sets. 3D-breeding produces scalable solutions that can be linked to genomics, farmers' knowledge and environmental data, to enhance the local adaptation of the resulting varieties and tailor their recommendation to the landscape.

and women farmer provided consistent evaluations, although with differences across locations (Fig. S4). In order to capture farmers' traditional knowledge regardless of gender, farmer scores were combined across men and women respondents, the H^2 of $OA_{STATION}$ was 0.78 across locations (Table S2). We validated the centralised prospect by predicting on-station performance from one season to the next, focusing on a subset of 41 top-performing genotypes in managed trials that were later distributed in decentralised farmer fields. This led to accuracies up to $\tau = 0.248$ in predicting $GY_{STATION}$ in the following season (Fig. S5). We found that farmers' evaluations ($OA_{STATION}$) were a better predictor than $GY_{STATION}$ to capture both $OA_{STATION}$ and $GY_{STATION}$, including when disaggregated by gender (Fig. S6). GY and OA collected in stations showed poor correlations with on-farm performance (Fig. S7).

The benchmark had a low prediction accuracy when using $GY_{STATION}$ to predict GY_{FARM} in individual seasons, with an average of $\tau = 0.046$. When using $OA_{STATION}$ to predict OA_{FARM}, the average was $\tau = 0.141$ (Table 1). Accuracy remained low when $GY_{STATION}$ was used to predict measures of GY_{FARM} and OA_{FARM} combined across seasons. However, $OA_{STATION}$ had consistent positive accuracy in predicting GY_{FARM} and OA_{FARM} (Fig. S8). This confirmed that genomic selection can be enhanced by farmers' traditional knowledge whereas selection based only on GY could result in reduced appreciation by farmers (Fig. S9).

$GY_{STATION}$ provided a more accurate prediction of GY_{FARM} when restricting the model to cold-tolerant genotypes (Fig. S10). This was likely due to the partial representation of the climatic variation that can be provided by a centralised approach with a handful of stations (Fig. S11), as farms could experience lower temperatures than stations (Fig. S12). Still, centralised predictions of increasingly distant farm environments shown an erratic pattern, showing that variation at the farming sites goes beyond that captured by temperature variation (Fig. S13). Regardless the fact that both stations and farms were located in the same agroecological zone (Fig. S3), the benchmark failed to predict performance under production conditions, showing that the small-scale variation in climate and management may hamper the success of centralised breeding decisions.

3D-breeding provides higher prediction accuracy than the benchmark

3D-breeding uses GY_{FARM} and OA_{FARM} to generate a model that affords predictive extrapolation across space and time. We determined the accuracy of 3D-breeding with cross-validated Plackett-Luce trees[21] considering seasons as bins. Environmental indices were derived from seasonal time-series climatic conditions observed in each plot. In this case, the model selected the minimum night temperature during vegetative growth and maximum night temperature during reproductive growth as the most critical indices in determining the performance of genotypes (Fig. S14). Genotypes' DNA markers were included in the model as an additive matrix in a Bayesian framework. 3D-breeding consistently provided higher accuracy than the benchmark for GY_{FARM} and OA_{FARM} with $\tau = 0.109$ and $\tau = 0.251$ (Table 1). The prediction accuracy of the 3D-breeding approach was not biased towards specific environmental conditions, suggesting that it could capture the environmental diversity of test sites better than the benchmark (Fig. S15).

Table 1: Performance of the 3D-breeding compared with the benchmark of a centralised genomic selection. 3D-breeding provides higher across-season goodness-of-fit (Kendall τ) than centralised genomic selection on overall appreciation (OA) and grain yield (GY) derived from farmer rankings on decentralised fields.

Approach	OA	GY
Centralised GS		
Season 1 (n=179)	0.134	-0.012
Season 2 (n=651)	0.105	0.076
Season 3 (n=335)	0.183	0.073
	0.141 (±0.03)	**0.046 (±0.04)**
3D-breeding		
Season 1 (n=179)	0.270	0.160
Season 2 (n=651)	0.276	0.078
Season 3 (n=335)	0.203	0.119
	0.251 (±0.04)	**0.109 (±0.04)**

Overall appreciation of genotypes in 3D-breeding resulted in higher accuracies than GY_{FARM} in all farmers' fields (Fig. S16). Previous studies showed that farmer evaluations are able to capture agronomic performance of genotypes in untested locations[16,19], as confirmed by the high H^2 observed for $OA_{STATION}$ (Table S2). Farmers provided OA according to their own experience and preferences, and it presumably depended on a combination of traits of which GY represented only one dimension[20]. By eliciting traditional knowledge of men and women farmers at cropping sites, 3D-breeding successfully predicted varietal performance under local growing conditions (Fig. S8). GY_{FARM} is objectively and independently measured at each plot and therefore it could not be biased by OA_{FARM}. It is possible that $GY_{STATION}$ and GY_{FARM} failed to capture secondary traits with high heritability (Table S1) that were observed by farmers and that were correlated to the GY_{FARM} of genotypes under marginal conditions[19,20]. As OA_{FARM} is directly related to the probability of variety adoption it is an important complement to GY in driving varietal development for marginal environments.

3D-breeding provides consistent recommendations across seasons

Next, we extrapolated the 3D-breeding model predictions to assess the probability that the genotypes selected by 3D-breeding on the basis of OA will outperform current recommendations as per the Wheat Atlas[22]. We found that the best three genotypes in each terminal node of the 3D-breeding model splits had a genetic background markedly separated from that of varieties currently recommended for the region, and consistently higher *worth* (Fig. 2A). Indeed, the model selected genotypes derived from landraces over improved varieties. We estimated the probability that the model recommendation exceeds the current recommendation in terms of OA_{FARM}. In this assessment, predictions from 3D-breeding outperformed the current varietal recommendations[22] in most of the farmers' fields, with consistently higher probabilities (0.83-0.91), including in marginal areas for which the centralised breeding approach could not provide accurate predictions (Fig. 2B). To provide an agronomic measure, we also predicted the increase in GY_{FARM} and tested to see if the yield advantage could be maintained by selecting the best three genotypes indicated by 3D-breeding under 15 different growing seasons simulated on target farms. We found that 3D-breeding ensured consistent recommendations over years with expected increases in yield of about 20% (Fig. 2C). Thus, 3D-breeding accurately identified the best performing genotypes to be advanced in breeding efforts targeting local growing conditions, to be developed into suitable new varieties, and to be promoted with environmental-specific recommendations.

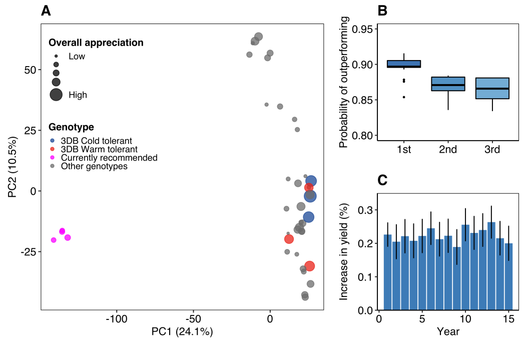

Fig. 2. Selection of durum wheat genotypes based on 3D-breeding. (**A**) Principal component coordinates of the genetic diversity of tested genotypes. Pink dots represent the varieties currently recommended for the area of study. 3DB Cold tolerant (blue) represents the top 3 genotypes selected by 3D-breeding in cold areas (minimum night temperature < 11.5 °C). 3DB Warm tolerant (red) represents the top 3 genotypes selected by 3D-breeding in warm areas (minimum night temperature > 11.5 °C). Size of dots represents the performance of genotypes in farmer fields as overall appreciation (OA). (**B**) Probability of outperforming improved varieties currently recommended by using genotype selection generated by 3D-breeding with OA. The panel shows the probability of the top 3 genotypes in a given location in outperforming the improved variety recommended for that location. (**C**) Expected increase in yield across 15 consecutive growing seasons (2001 to 2015) for genotype selection from 3D-breeding.

Discussion

Genomic selection is a well-known approach to accelerate breeding programs, but current implementations in plant breeding have not yet reached their full potential. The earliest and most successful implementations of genomic selection have arguably occurred in dairy cattle breeding[23]. The accelerated evaluation of bull net merit was key to this[24], but that success also depended on the fact that breeders had access to phenotyping data from a broad range of environments in the form of milking records, which farmers record for their own management benefit. In conventional crop breeding, all of the phenotyping costs fall on the breeding program and limit the number of target environments that can be represented in the selection process. 3D-breeding seeks to complement and expand the flow of information from a few centralised locations to the whole mega-environment where results from numerous decentralised observations and farmer knowledge may converge to inform breeding decisions.

In centralised breeding, the environmental variation of target environments is factored through experimental control or indirectly as an average response across breeding stations as in our benchmark. This makes extrapolation to real farming conditions challenging. $G \times E$ affects yield and its components[25,26] and calls for selection models to explicitly account for it[27]. These models, however, are bound to the observations that can be made in resource-intensive breeding trials. Data from crowdsourced citizen science, like 3D-breeding, may further our understanding of the $G \times E$ interactions that are observed in farmer fields and allow the integration of increasingly accurate seasonal prediction models[28] in breeding and germplasm recommendation pipelines.

The 3D-breeding approach addresses the low correlation between performance in selection environments and production environments, while taking a step forward to fully data-driven breeding. In this, 3D-breeding is a promising approach that could add to conventional breeding increasing varietal performance in smallholder agriculture, which accounts for more than 80% of all global farms[17]. Here, the adoption rate of current breeding innovation may be suboptimal due to socioeconomic and environmental factors[20,29–31]. Climate change is pushing these farming systems to the edge of their adaptation capacity with increasing pressure from pest and diseases[32,33], threats of yield loss[7,34] and increased seasonal climatic variability[35,36], calling for tailored solutions. 3D-breeding may speed up the turnover of varietal release to address these challenges. As farmers are at the centre of the experimental design, varieties deriving from 3D-breeding are more likely to be adopted and suited to local cultivation[10,37], increasing the effectiveness of breeding efforts. Indeed, we found that farmers' OA was a better predictor than GY in predicting yield realized both in centralised and decentralised trials (Table 1). Likewise, varieties derived from landraces consistently outranked the performance of improved varieties (Fig. 2A) derived from centralised breeding[18]. Beyond varietal recommendations, 3D-breeding can direct the choice of parents to crosses aiming at the production of recombinant lines to provide higher and more stable yields in local agriculture.

3D-breeding is useful beyond smallholder farming agriculture, and the citizen science approach on which it relies has already been applied to several crops to enhance the selection of climate-adapted varieties[14]. Its general scheme may also be useful in high-input, yield maximizing agriculture to enhance local adaptation and support sustainability and food security, where the usefulness of farmers' evaluations in a genomic setting was already demonstrated[16]. There are a number of open questions in relation to decentralised crop breeding, including how to best motivate new farmers to participate in the evaluation of materials, how much planting material each farmer needs, the logistics of providing farmers with the genetic material, and how to share benefits deriving from the utilisation of farmers' knowledge to produce new varieties.

3D-breeding may be most effective as a complement to a centralised breeding system providing a high-throughput evaluation of correlated traits to support earlier varietal selection to be tested in farmer fields[38]. Accuracy is just one among the factors controlling genetic gain[39], thus our findings should be integrated in the broader picture of modern breeding. Multi-trait models may increase prediction accuracy by measuring correlated traits with higher heritability[38,40,41]. These models could be employed in centralised stations and used to narrow down the set of varieties to be distributed to farmers in the 3D-breeding approach aiming to fine-tune local adaptation. Moreover, our findings support the need to further explore the challenge to model farmers' appreciation at the genomic level to improve the effectiveness of genotypes evaluation trials[16].

The advantages provided by the approach are clear: phenotyping costs would be divided in much smaller packets, supporting the modular expansion of the breeding effort towards new genetic materials or new locations. In return, each generated datapoint would be a better representation of the true farming conditions to which varieties are directed. Previous research found that the involvement of farmers in selection experiments has negligible effects on costs[42]. In 3D breeding the costs are shared by farmers, who would in exchange obtain access to the best materials for their farm. Farmer preference would be collected directly on farms rather than derived from correlated metrics that come from on-station evaluations in centralised breeding. In terms of absolute costs, an implementation of 3D-breeding based on OA would only require seed amplification, seed distribution and telecommunications to obtain feedback from farmers. Genotyping costs are negligible thanks to ever increasing sequencing capabilities[1]. Indeed, a tricot experiment conducted in Nicaragua[14] resulted in a lower cost per datapoint in decentralised evaluations than in stations experiments.

The data-driven focus of 3D-breeding enables embracing the complexity of real-world $G \times E$ for the benefit of breeding. Such a multidimensional, collaborative approach calls for best practices in data management and sharing[43]. 3D-breeding is based on a documented set of methods, from experimental design[15] to data curation and analysis[21,44]. While our demonstration of these methods relied on a large dataset, we believe that much larger field sample designs and genomic variant datasets are quite feasible and will provide additional power, as is also much in evidence in livestock genetics. The expansion of the design with the addition of further testing seasons and local management conditions may allow to highlight drivers of local performance of genotypes beyond temperature[45]. The crowdsourced citizen science approach associated with open-source digital tools makes it possible for breeders and farmers to apply 3D-breeding in new contexts and crops, dependent only on creativity in identifying untested production niches, potentiating a culturally-driven coevolution between farming systems and data-driven breeding to complement traditional breeding.

Materials and Methods

Genotypes sampling and DNA extraction

We selected 400 durum wheat (*Triticum durum* Desf.) genotypes from a representative collection of accessions from the Ethiopian Biodiversity Institute. Genomic DNA was extracted from fresh leaves pooled from five seedlings for each of the accessions with the *GenElute*^{TM} Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, St Louis, USA) following manufacturer's instructions in the Molecular and Biotechnology Laboratory at Mekelle University, Ethiopia. Genomic DNA was checked for quantity and quality by electrophoresis on 1% agarose gel and NanodropTM 2000 (Thermo Fisher Scientific Inc., Waltham, USA). Genotyping was performed on the Infinium 90k wheat chip at TraitGenetics GmbH (Gatersleben, Germany). Single nucleotide polymorphisms (SNPs) were called using the tetraploid wheat pipeline in GenomeStudio V11 (Illumina, Inc., San Diego, CA, USA). SNP calls were cleaned for quality by filtering positions and samples with failure rate above 80% and heterozygosity above 50%. Full details on the genotyping are given by Mengistu et al. (2016)[18]. The SNP calls for the genotypes included in this study and the details on the provenance of genotypes tested are given as part of the full dataset on Dataverse[46].

Evaluation of genotypes in centralised trials

Centralised trials were performed in 2012 and 2013 in the districts of Geregera (Amhara) and Hagreselam (Tigray) (Fig. S1). The experimental stations were chosen to represent the highland agroecology of Ethiopia, and are often used as varietal testing sites for local agriculture. The trial was laid out in a replicated alpha lattice design, and field managements were conducted as per local guidelines with manual weeding. Accessions were sown in four rows 2.5 m long, at a seeding rate of 100 $kg \cdot ha^{-1}$. At sowing, 100 $kg \cdot ha^{-1}$ diammonium phosphate and 50 $kg \cdot ha^{-1}$ urea were applied, with additional 50 $kg \cdot ha^{-1}$ urea at tillering. In 2012, thirty experienced smallholder farmers growing durum wheat (15 men and 15 women) were invited to participate in the trial evaluations at the station plots, held concurrently after flowering stage. The farmers had no previous knowledge of the genotypes included in this study to prevent bias in the evaluations. The participants provided appraisal with Likert[47] scales (1 to 5, worse to best) given to genotypes for overall appreciation (OA)[19,20]. Farmers did not use half-values in order to streamline the evaluation effort. Research technicians measured grain yield (GY) as grams of grain produced per plot, then converted into $t \cdot ha^{-1}$. Absolute values of GY and OA measured in centralised trials were converted into ordinal rankings.

Evaluation of genotypes in decentralised trials

A total of 1,165 decentralised plots were stablished between 2013 and 2015 during three growing seasons across the regions of Amhara (471), Oromia (399) and Tigray (295) (Fig. S1) using a subset of the 41 best genotypes identified through farmer evaluation in centralised trials[20]. The farms were sampled in the same agroecological zones of the centralised fields (Fig. S3). Season 1 (2013) comprised 179 fields, Season 2 (2014) comprised 651 fields, and Season 3 (2015) comprised 335 field. Trials (farmer-managed plots) followed the triadic comparison of technologies (*tricot*) approach[15]. Sets of three local genotypes plus an improved variety (Asassa in Tigray and Amhara, and Hitosa and Ude in Oromia) were allocated randomly to farmers as incomplete blocks, maintaining spatial balance by assigning roughly equal frequencies of the genotypes. Trial size ranged from 0.4 m^2 to 1.6 m^2 depending on season and location. Farmers set, managed and evaluated their own experiments indicating the OA of genotypes through ranking the four entries that they received from best to worst, using pre-defined answer forms at the end of the growing season. Differently from the centralised trials, the OA was derived from the relative rankings of genotypes as each farmer evaluated a different set of materials. Research technicians collected GY measures in farmers' plots after harvesting.

Centralised trait data analysis

All analyses were done in R[48]. $GY_{STATION}$ and $OA_{STATION}$ measured in centralised trials were used to derive best linear unbiased prediction (BLUP) values using the R package ASReml[49], treating locations as a fixed factor and all other factors as random. Full model details are reported in Supplementary Note 1. For the central comparison between benchmark and 3D-breeding, we used measures of $GY_{STATION}$ combined across seasons and locations. Similarly, $OA_{STATION}$ in the central comparison represents OA values combined across genders and locations. When relevant, $GY_{STATION}$ and $OA_{STATION}$ measures are split by location, season or gender. Agreement between farmer gender groups in evaluating centralised station data was derived from a linear model fit. Spearman correlations between location specific BLUP values and farm performance were also computed.

Decentralised trait data analysis

For the analysis of the decentralised data we used the Plackett-Luce model[21,50,51]. Plackett-Luce is a rank-based model that estimates the *worth* parameter. These parameters α are related to the probability (P) that one genotype i wins against all other n genotypes in set, and are obtained using the following equation:

Equation [1]

$$P(i \succ \{j, ..., n\}) = \frac{a_i}{a_1 + ... + a_n}$$

Each farmer ranking in the input OA_{FARM} data was converted to a *dense* ranking, which ranked genotypes from 1 (first place) to n_r (last place). Genotypes not ranked in a given farmer plot had a rank of 0. For GY_{FARM}, we converted each observation into ordinal rankings, by assigning values from 1 (highest yield) to n_r (lowest yield), and then converted into a *dense* ranking. Altogether, OA_{FARM} and GY_{FARM} represented the probably of winning of any given genotype in any set of testing sites by either farmer choice (OA) or production (GY). In order to detail specific aspects of the dataset, the probability of winning was computed restricting farms to those belonging to any of specific subset (*e.g.* year, environmental distance class, region). Turner et al 2020[21] introduces the Plackett-Luce model and its applications with partial rankings, which was the case in this research. The implementation of Plackett-Luce model to analyse data from decentralised crop variety trials is demonstrated by van Etten et al. 2019[14].

Implementation of the genomic selection benchmark

The benchmark representing a centralised breeding system was conducted using genomic selection models and marker-based genetic relationship matrices computed on BLUP data with the package rrBLUP[52]. To measure accuracy of genomic selection predictions, we calculated the Kendall's tau coefficient (τ), a measure of similarity of rankings[53], between predicted values and observed values. The use of the τ metric, uncommon in breeding, allowed to compare accuracies with the 3D-breeding approach. A Pearson's correlation, the standard metric for genomic selection accuracy, was also computed but did not show any relevant difference with the Kendall τ.

The following genomic selection scenarios were considered: (i) using $GY_{STATION}$ and $OA_{STATION}$ to predict $GY_{STATION}$ measured in the same locations in the following season; (ii) using $GY_{STATION}$ and $OA_{STATION}$ to predict GY_{FARM} and OA_{FARM}. In the first scenario, the training population was made of all genotypes measured in stations in 2012 and the test population was the subset of 41 genotypes included in the decentralised trials and measured in stations in 2013. In the second scenario, the training population resulted from the combined measures of $GY_{STATION}$ and $OA_{STATION}$ across seasons and locations and the test population was the subset of 41 genotypes measured for GY_{FARM} and OA_{FARM} independently for each season. In a siding analysis stations were used to predict increasingly different farms based on quantiles of environmental distances according to the distances derived from climatic data. Note that in both scenarios the use of a training sample overlapping the test sample was meant to allow a fair comparison with the 3D-breeding that also uses the entire dataset to train the model.

Mirroring the approach used in the 3D-breeding, the accuracy of genomic selection in the second scenario was derived from a cross-validation approach averaging Kendall τ specific for Season 1, Season 2, and Season 3 using the square root of the sample size as weights[54].

Alternative genomic selection scenarios were also performed but showed consistent accuracy values. Most notably: (i) without overlap between training and test samples, (ii) restricting training and test samples to the subset distributed to decentralised fields, (iii) using rankings derived from BLUP values to train the GY and OA model. In none of the cases above, the benchmark provided accuracies with a noticeable difference from the second scenario reported above and used for the central comparison with 3D-breeding.

3D-breeding implementation

The model representing the 3D-breeding approach was built with the data generated by the citizen science decentralised trials using the Plackett-Luce model. We used two variants of the model, one using OA_{FARM} and the other using GY_{FARM} to check which of these metrics had the higher model precision, as shown in Table 1. DNA data from SNPs was added into the model as a prior using an additive matrix. Climatic variables were linked to the rankings using the Model-Based Recursive Partitioning approach[55]. Daily temperature and precipitation data were obtained from the NASA LaRC POWER Project (https://power.larc.nasa.gov/), using the R package nasapower[56]. A set of climatic covariates were extracted for the vegetative, reproductive and grain filling phases and the whole growth period (from planting date to harvesting as measured on-site) in each observation point using the R package climatrends[57]. This resulted in 110 covariates.

To create a model that provides generalizable predictions across seasons with few covariates, we used blocked cross-validation (with seasons as blocks) combined with a forward selection[58]. We used the deviance values of each validation season to calculate an Akaike weight, which is the probability that a given variable combination represents the best model[59]. We performed forward selection, using this combined Akaike weight as our selection criterion. The PLT models had a cut-off value of $\alpha = 0.01$ and a minimal group size of 20 percent of the total dataset partitioning selection. The variables selected under this procedure were the maximum night temperature (°C) during reproductive growth and the minimum night temperature (°C) during the vegetative growth. To compare the accuracy of the model representing 3D-breeding with the benchmark, we calculated the Kendall τ between observed rankings and predicted coefficients.

Generalisation of the 3D-breeding

We then evaluated if the model obtained with the variable selection procedure retained predictive power across seasons. We simulated untested future seasonal climate with representative seasonal scenarios of past climate conditions by extracting the last 15 years of daily climate data derived from NASA POWER (2001-2015). We determined three windows for sowing dates in each growing season as the midpoints of equiprobable quantile intervals estimated from the observed planting dates in the data set. We predicted genotype performance for 15 seasons $\times$ 3 sowing dates (45 seasonal scenarios) for 1,200 random points generated across an alpha hull area within the range of the trials' coordinates. We averaged genotype probability of winning across these scenarios for each planting date interval, excluding the seasons used as testing data.

We calculated the reliability (Fig. 2B), the probability of outperforming a check variety[60]. We used the *worth* parameters from Plackett-Luce to determine the values of positive-valued parameters α_i associated with each genotype i, by comparing the *worth* from the check variety (Asassa, Hitosa and Ude, currently recommended for the mega-environment[22]) with the *worth* of the selected genotypes from 3D-breeding. These parameters (α) are related to the probability (P) that genotype i wins against all other n genotypes in set following the Luce's Choice Axiom, which states that the probability that one item beats another is independent from the presence or absence of any other items in the set. Reliability was calculated following Equation 2:

Equation [2]

$$P(i \succ \{j, ..., n\}) = \frac{a_i}{a_1 + ... + a_n} = \frac{a_i}{1} = a_i$$

Environmental characterisation of test sites and genotypes

The agroecological zonation of Ethiopia was obtained by the Ethiopian Institute of Agricultural Research (EIAR)[61]. GPS coordinates of centralised stations and decentralised farmer fields were used to retrieve climatic data from NASA POWER. Temperature indices for covariates used in the PL model were retrieved for the growing seasons object of the study in the time span from sowing date and flowering dates as measured on-site. Climatic variables considered were the maximum night temperature (°C) during reproductive growth and the minimum night temperature (°C) during the vegetative growth, which showed to be the most relevant for the sampled data. A principal component analysis (PCA) was used to summarise and depict variation at test sites. Climatic distance of test sites was derived from a multidimensional scaling (MDS) of the multivariate climate dataset. For each of the two stations, climatic distance was computed with all farm sites. Wheat genotypes were split in cold adapted and warm adapted according to the altitude of their original sampling site with a one-tailed, unequal-variance t-test.

Supporting software

Organising the datasets relied on R packages data.table[62], caret[63], gosset[64], janitor[65], magrittr[66] and tidyverse[67]. Climatic variables were obtained using the packages climatrends[57] and nasapower[56]. Statistical analysis was performed using packages PlackettLuce[21], gosset[64] and qvcalc[68]. Spatial visualisation was performed with the packages dismo[69], raster[70], sf[71] and smoothr[72]. Charts were produced using packages corrplot[73], ggplot2[74] and patchwork[75].

Data availability

Full data and code are available through Dataverse[46].

References

1. Poland, J. Breeding-assisted genomics. *Current Opinion in Plant Biology* **24**, 119–124 (2015).

2. Hickey, J. M., Chiurugwi, T., Mackay, I., Powell, W. & Implementing Genomic Selection in CGIAR Breeding Programs Workshop Participants. Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature Genetics* **49**, 1297–1303 (2017).

3. Eshed, Y. & Lippman, Z. B. Revolutions in agriculture chart a course for targeted breeding of old and new crops. *Science* **366**, eaax0025 (2019).

4. Hickey, L. T. *et al.* Breeding crops to feed 10 billion. *Nature Biotechnology* 1 (2019). doi:10.1038/s41587-019-0152-9

5. Heffner, E. L., Sorrells, M. E. & Jannink, J. L. Genomic selection for crop improvement. *Crop Science* **49**, 1–12 (2009).

6. Ceccarelli, S. & Grando, S. Decentralized-participatory plant breeding: An example of demand driven research. *Euphytica* **155**, 349–360 (2007).

7. Challinor, A. J., Koehler, A.-K., Ramirez-Villegas, J., Whitfield, S. & Das, B. Current warming will reduce yields unless maize breeding and seed systems adapt immediately. *Nature Climate Change* **6**, 954–958 (2016).

8. Tester, M. & Langridge, P. Breeding technologies to increase crop production in a changing world. *Science* **327**, 818–822 (2010).

9. Witcombe, J. R., Joshi, A., Joshi, K. D. & Sthapit, B. R. Farmer participatory crop improvement. I. Varietal selection and breeding methods and their impact on biodiversity. *Experimental Agriculture* **32**, 445–460 (1996).

10. Rhoades, R. E. & Booth, R. H. Farmer-back-to-farmer: A model for generating acceptable agricultural technology. *Agricultural Administration* **11**, 127–137 (1982).

11. Ceccarelli, S. Efficiency of Plant Breeding. *Crop Science* **55**, 87 (2015).

12. Ceccarelli, S. & Grando, S. Participatory plant breeding: Who did it, who does it and where? *Experimental Agriculture* **2019**, 1–11 (2019).

13. van Eeuwijk, F. A., Cooper, M., DeLacy, I. H., Ceccarelli, S. & Grando, S. Some vocabulary and grammar for the analysis of multi-environment trials, as applied to the analysis of FPB and PPB trials. *Euphytica* **122**, 477–490 (2001).

14. van Etten, J. *et al.* Crop variety management for climate adaptation supported by citizen

science. *Proceedings of the National Academy of Sciences* **116**, 4194–4199 (2019).

15. van Etten, J. *et al.* First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Experimental Agriculture* **55**, 275–296 (2019).

16. Annicchiarico, P., Russi, L., Romani, M., Pecetti, L. & Nazzicari, N. Farmer-participatory vs. conventional market-oriented breeding of inbred crops using phenotypic and genome-enabled approaches: A pea case study. *Field Crops Research* **232**, 30–39 (2019).

17. Lowder, S. K., Skoet, J. & Raney, T. The Number, Size, and Distribution of Farms, Smallholder Farms, and Family Farms Worldwide. *World Development* **87**, 16–29 (2016).

18. Mengistu, D. K. *et al.* High-density molecular characterization and association mapping in Ethiopian durum wheat landraces reveals high diversity and potential for wheat breeding. *Plant Biotechnology Journal* **14**, 1800–1812 (2016).

19. Kidane, Y. G. *et al.* Genome Wide Association Study to Identify the Genetic Base of Smallholder Farmer Preferences of Durum Wheat Traits. *Frontiers in Plant Science* **8**, 1230 (2017).

20. Mancini, C. *et al.* Joining smallholder farmers' traditional knowledge with metric traits to select better varieties of Ethiopian wheat. *Scientific Reports* **7**, 9120 (2017).

21. Turner, H. L., van Etten, J., Firth, D. & Kosmidis, I. Modelling rankings in R: the PlackettLuce package. *Computational Statistics* (2020). doi:10.1007/s00180-020-00959-3

22. CIMMYT. Wheat Atlas. Available at: http://beta.wheatatlas.org/.

23. Meuwissen, T. H. E., Hayes, B. J. & Goddard, M. E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* **157**, 1819–1829 (2001).

24. Hayes, B. J. & Daetwyler, H. D. 1000 Bull Genomes Project to Map Simple and Complex Genetic Traits in Cattle: Applications and Outcomes. *Annual Review of Animal Biosciences* **7**, 89–102 (2019).

25. Quintero, A., Molero, G., Reynolds, M. P. & Calderini, D. F. Trade-off between grain weight and grain number in wheat depends on GxE interaction: A case study of an elite CIMMYT panel (CIMCOG). *European Journal of Agronomy* **92**, 17–29 (2018).

26. Ceccarelli, S., Erskine, W., Hamblin, J. & Grando, S. Genotype by environment interaction and international breeding programmes. *Experimental Agriculture* **30**, 177–187 (1994).

27. Eeuwijk, F. A. van *et al.* Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. *Plant Science* **282**, 23–39 (2019).

28. Bauer, P., Thorpe, A. & Brunet, G. The quiet revolution of numerical weather prediction. *Nature* **525**, 47–55 (2015).

29. Dixon, J. *et al.* Adoption and economic impact of improved wheat varieties in the developing world. **144**, 489–502 (2006).

30. Jalleta, T. Participatory evaluation of the performance of some improved bread wheat (Triticum aestivum) varieties in the Jijiga plains of eastern Ethiopia. *Experimental Agriculture* **40**, 89–97 (2004).

31. Tesfaye, S., Bedada, B. & Mesay, Y. Impact of improved wheat technology adoption on productivity and income in Ethiopia. *African Crop Science Journal* **24**, 127–135 (2016).

32. Deutsch, C. A. *et al.* Increase in crop losses to insect pests in a warming climate. *Science* **361**, 916–919 (2018).

33. Carvajal-Yepes, M. *et al.* A global surveillance system for crop diseases. *Science* **364**, 1237–1239 (2019).

34. Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C. & Foley, J. A. Recent patterns of crop yield growth and stagnation. *Nature Communications* **3**, (2012).

35. Cai, W. *et al.* Increasing frequency of extreme El Niño events due to greenhouse warming. *Nature Climate Change* **5**, 1–6 (2014).

36. Holmgren, M., Hirota, M., van Nes, E. H. & Scheffer, M. Effects of interannual climate variability on tropical tree cover. *Nature Climate Change* **3**, 755–758 (2013).

37. Ceccarelli, S. *Plant Breeding with Farmers: A technical manual.* 125 (International Center for Agricultural Research in the Dry Areas (ICARDA), 2012). doi:https://hdl.handle.net/20.500.11766/7745

38. Lado, B. *et al.* Resource allocation optimization with multi-trait genomic prediction for bread wheat (Triticum aestivum L.) baking quality. *Theoretical and Applied Genetics* **131**, 2719–2731 (2018).

39. Heffner, E. L., Lorenz, A. J., Jannink, J.-L. & Sorrells, M. E. Plant Breeding with Genomic Selection: Gain per Unit Time and Cost. *Crop Science* **50**, 1681–1690 (2010).

40. Runcie, D. & Cheng, H. Pitfalls and remedies for cross validation with multi-trait genomic prediction methods. *G3: Genes, Genomes, Genetics* **9**, 3727–3741 (2019).

41. Ibba, M. I. *et al.* Genome-based prediction of multiple wheat quality traits in multiple years. *The Plant Genome* (2020). doi:10.1002/tpg2.20034

42. Mangione, D., Senni, S., Puccioni, M., Grando, S. & Ceccarelli, S. The cost of participatory barley breeding. *Euphytica* **150**, 289–306 (2006).

43. Leonelli, S., Davey, R. P., Arnaud, E., Parry, G. & Bastow, R. Data management and best practice for plant science. *Nature Plants* **3**, 1–4 (2017).

44. van Etten, J. *et al.* ClimMob: Software to support experimental citizen science in agriculture. version 3.1. (2020). Available at: https://climmob.net.

45. Kehel, Z., Crossa, J. & Reynolds, M. Identifying Climate Patterns during the Crop-Growing Cycle from 30 Years of CIMMYT Elite Spring Wheat International Yield Trials. in *Applied mathematics and omics to assess crop genetic resources for climate change adaptive traits* (eds. Bari, A., Damania, A. B., Mackay, M. & Dayanandan, S.) 151–174 (CRC Press, 2016).

46. de Sousa, K. *et al.* Replication data for: "Data-driven decentralized breeding increases genetic gain in a challenging crop production environment. (2020). doi:10.7910/DVN/OEZGVP

47. Likert, R. A technique for the measurement of attitudes. *Archives of Psychology* **22**, 55 (1932).

48. R Core Team. R: A language and environment for statistical computing. version 4.0.2. (2020).

49. Gilmour, A. R., Gogel, B. J., Cullis, B. R., Welham, S. J. & Thompson, R. *ASReml User Guide Release 4.1.* (Structural Specification, VSN International Ltd, Hemel Hempstead, HP1 1ES, 2015).

50. Plackett, R. L. The Analysis of Permutations. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **24**, 193–202 (1975).

51. Luce, R. D. *Individual Choice Behavior.* 153 (Courier Corporation, 1959).

52. Endelman, J. B. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. *The Plant Genome Journal* **4**, 250 (2011).

53. Kendall, M. G. A new measure of ranking correlation. *Biometrika* **30**, 81–93 (1938).

54. Whitlock, M. C. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *Journal of Evolutionary Biology* **18**, 1368–1373 (2005).

55. Zeileis, A., Hothorn, T. & Hornik, K. Model-based recursive partitioning. *Journal of Computational and Graphical Statistics* **17**, 492–514 (2008).

56. Sparks, A. H. nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and

Climatology Data Client for R. *Journal of Open Source Software* **3**, 1035 (2018).

57. de Sousa, K., van Etten, J. & Solberg, S. Ø. climatrends: Climate Variability Indices for Ecological Modelling. R package version 0.1.6. (2020). Available at: https://CRAN.R-project.org/package=climatrends.

58. Meyer, H., Reudenbach, C., Hengl, T., Katurji, M. & Nauss, T. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Environmental Modelling and Software* **101**, 1–9 (2018).

59. Wagenmakers, E. J. & Farrell, S. AIC model selection using Akaike weights. *Psychonomic Bulletin and Review* **11**, 192–196 (2004).

60. Eskridge, K. M. & Mumm, R. F. Choosing plant cultivars based on the probability of outperforming a check. *Theoretical and Applied Genetics* **84-84**, 494–500 (1992).

61. Ministry of Agriculture of Ethiopia. *Agroecological Zonations of Ethiopia.* (2020).

62. Dowle, M. & Srinivasan, A. data.table: Extension of ‘data.frame‘. R package version 1.12.8. (2019). Available at: https://CRAN.R-project.org/package=data.table.

63. Kuhn, M. caret: Classification and Regression Training. R package version 6.0-85. (2020). Available at: https://CRAN.R-project.org/package=caret.

64. de Sousa, K., van Etten, J., Dumble, S., Greliche, N. & Steinke, J. gosset: Modelling Metadata and Crowdsourced Citizen Science. R package version 0.2.1. Available at: https://agrobioinfoservices.github.io/gosset/.

65. Firke, S. janitor: Simple Tools for Examining and Cleaning Dirty Data. R package version 1.2.1. (2020). Available at: https://CRAN.R-project.org/package=janitor.

66. Bache, S. M. & Wickham, H. magrittr: A Forward-Pipe Operator for R. R package version 1.5. (2014). Available at: https://CRAN.R-project.org/package=magrittr.

67. Wickham, H. *et al.* Welcome to the Tidyverse. *Journal of Open Source Software* **4**, 1686 (2019).

68. Firth, D. qvcalc: Quasi Variances for Factor Effects in Statistical Models. R package version 1.0.1. (2019). Available at: https://CRAN.R-project.org/package=qvcalc.

69. Hijmans, R. J., Phillips, S., Leathwick, J. & Elith, J. dismo: Species Distribution Modeling. R package version 1.1-4. (2017). Available at: https://CRAN.R-project.org/package=dismo.

70. Hijmans, R. J. *et al.* raster: Geographic Data Analysis and Modeling. R package version 2.5-8. (2015). Available at: https://cran.r-project.org/package=raster.

71. Pebesma, E. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* **10**, 439–446 (2018).

72. Strimas-Mackey, M. smoothr: Smooth and Tidy Spatial Features. R package version 0.1.2. (2020). Available at: https://CRAN.R-project.org/package=smoothr.

73. Wei, T. & Simko, V. *R package "corrplot": Visualization of a correlation matrix.* (2017).

74. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis.* (Springer-Verlag New York, 2016).

75. Pedersen, T. L. patchwork: The Composer of Plots. R package version 1.0.0. (2019). Available at: https://CRAN.R-project.org/package=patchwork.

5

The R package `chirps` hex logo

chirps: API Client for the CHIRPS Precipitation Data in R

Kauê de Sousa[1, 2], **Adam H. Sparks**[3], **William Ashmall**[4], **Jacob van Etten**[2], and **Svein Ø. Solberg**[1]

1 Department of Agricultural Sciences, Inland Norway University of Applied Sciences, Hamar, Norway **2** Bioversity International, Rome, Italy **3** Centre for Crop Health, University of Southern Queensland, Toowoomba, Australia **4** Universities Space Research Association, National Aeronautics and Space Administration (NASA), Huntsville, USA

DOI: 10.21105/joss.02419

Software

- Review
- Repository
- Archive

Editor: Daniel S. Katz

Reviewers:

- @danielskatz

Submitted: 01 July 2020
Published: 01 July 2020

Summary

The *chirps* package provides functionalities for reproducible analysis in R (R Core Team, 2020) using the CHIRPS (Funk et al., 2015) data. CHIRPS is daily precipitation data set developed by the Climate Hazards Group (Funk et al., 2015) for high resolution precipitation gridded data. Spanning 50° S to 50° N (and all longitudes) and ranging from 1981 to near-present (normally with a 45 day lag), CHIRPS incorporates 0.05 arc-degree resolution satellite imagery, and in-situ station data to create gridded precipitation time series for trend analysis and seasonal drought monitoring (Funk et al., 2015). Additionally, the package provides the API client for the IMERG (Huffman et al., 2014) and ESI (SERVIR Global, 2019a) data. The Integrated Multi-satelliE Retrievals for GPM (IMERG) data provides near-real time global observations of rainfall at 0.5 arc-degree resolution, which can be used to estimate total rainfall accumulation from storm systems and quantify the intensity of rainfall and flood impacts from tropical cyclones and other storm systems. IMERG is a daily precipitation dataset available from 2015 to near-present. The evaporative stress index (ESI) data describes temporal anomalies in evapotranspiration produced weekly at 0.25 arc-degree resolution for the entire globe (Anderson et al., 2011). The ESI data is based on satellite observations of land surface temperature, which are used to estimate water loss due to evapotranspiration (the sum of evaporation and plant transpiration from the Earth's land and ocean surface to the atmosphere). The ESI data is available from 2001 to near-present. When using these data sets in publications please cite Funk et al. (2015) for CHIRPS, Huffman et al. (2014) for IMERG and SERVIR Global (2019a) for ESI.

Implementation

Four main functions are provided, `get_chirps()`, `get_imerg()`, `get_esi()` and `precip_indices()`. The `get_chirps()` function provides access to CHIRPS data via the ClimateSERV API Client (SERVIR Global, 2019b) with methods to handle objects of class 'data.frame', 'geojson' and 'sf' via the package *methods* (R Core Team, 2020). To accept the query, ClimateSERV requires a geojson object of type 'Polygon' (one single polygon per request). Using the package *sf* (Pebesma, 2018) internally, the input provided in `get_chirps()` is transformed into a list of polygons with a small buffer area (0.0001 arc-sec by default) around the point and transformed into a list of geojson strings. *chirps* uses *crul* (Chamberlain, 2019) to interface with ClimateSERV API. The query returns a JSON object parsed to *jsonlite* (Ooms, 2014) to obtain the data frame for the time series required. `get_chirps()` returns a data.frame, which also inherits the classes 'chirps' and 'chirps_df', where each id represents the index for the rows in the in-putted 'object'. The function `get_imerg()` returns the precipitation data from the IMERG data set. The function works with the same parameters described for `get_chirps()` and also inherits the classes 'chirps' and 'chirps_df'. The function `get_esi()` returns the evaporative stress index (ESI) data (Anderson et al., 2011), and works similarly to `get_chirps()` returning a data.frame which inherit the class 'chirps_df'. Users providing objects of class 'sf' and 'geojson' in `get_chirps()`, `get_imerg()` and `get_esi()` can also choose to return an object with the same class as the object provided using the arguments 'as.sf = TRUE' or 'as.geojson = TRUE'. With the function `precip_indices()` users can assess how the precipitation changes across the requested time series using precipitation variability indices (Aguilar et al., 2005), computed using *stats* (R Core Team, 2020), the main input is an object of class 'chirps'. Extended documentation is provided with examples on how to increase the buffer area and draw quadrants for the geojson polygon using *sf* (Pebesma, 2018).

Application: a case study in the Tapajós National Forest

The *Tapajós* National Forest is a protected area in the Brazilian Amazon. Located within the coordinates -55.4° and -54.8° E and -4.1° and -2.7° S with ~ 527,400 ha of multiple Amazonian ecosystems. We take twenty random points across its area to get the precipitation from Jan-2008 to Dec-2018 using `get_chirps()`. We use an object of class 'sf' which is passed to the method `get_chirps.sf()`. Then, we compute the precipitation indices for the time series with intervals of 30 days using `precip_indices()`.

```
library("chirps")
library("sf")

data("tapajos", package = "chirps")
set.seed(1234)
tp <- st_sample(tapajos, 20)
tp <- st_as_sf(tp)

dt <- get_chirps(tp, dates = c("2008-01-01","2018-01-31"))

p_ind <- precip_indices(dt, timeseries = TRUE, intervals = 30)
```

We selected four indices for the visualization using *tidyverse* (Wickham et al., 2019). Plots were ensembled together using *gridExtra* (Auguie, 2017). Here we see how these indices are changing across the time series (Figure 1). In this quick assessment, we note an increasing extent of consecutive dry days (MLDS) across the time series, with also a decrease in the number of consecutive rainy days (MLWS), which stays above the historical average for MLDS and bellow the historical average for MLWS. The trends also show a decrease in the total rainfall in the 30-days intervals, staying below the average after 2014. Finally, we note a decrease in maximum consecutive 5-days precipitation, which also stays bellow the historical average.

The Journal of Open Source Software

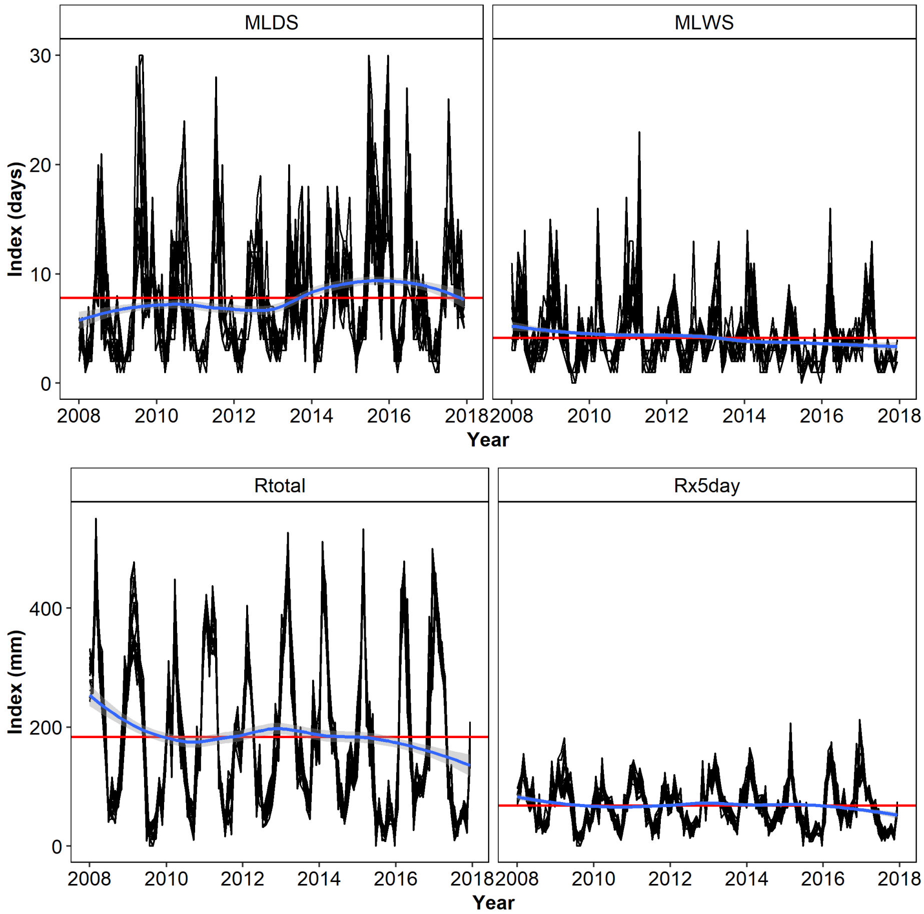

Figure 1: Trends in precipitation variability across the Tapajós National Forest, Brazil, for the period of 01-Jan-2010 to 31-Dec-2018 with four precipitation indices. MLDS, maximum length of consecutive dry days (days), MLWS, maximum length of consecutive wet days (days), Rtotal, total precipitation (mm), Rx5day, maximum consecutive 5-days precipitation (mm). Red lines indicates the historical mean of each index in the time series. Blue line indicates the smoothed trends in each index using the 'loess' method.

Other applications and conclusion

Deriving precipitation indices that can be obtained from CHIRPS proved to be an excellent approach to evaluate the climate variability using precipitation data (de Sousa et al., 2018) and the effects of climate change on a continental analysis (Aguilar et al., 2005). Additionally, these indices can be used to register specific effects of climate variability on crop varietal performance. In crop modelling, Kehel, Crossa, & Reynolds (2016) applied this to assess the interactions of wheat varieties with the environment, showing how severe drought, assessed with the maximum length of dry spell (MLDS), can affect the plant development and the yield. These indices can also be useful to improve variety recommendation for climate adaptation in marginal production environments (van Etten et al., 2019).

Overall, CHIRPS data can be used in many applications and currently has over 800 citations from studies using this tool. Many applications are the field of agriculture, hydrologic modelling and drought monitoring, but also some studies using this in disease control programs (e.g. Thomson et al. (2017), Horn et al. (2018)). The *chirps* package aims to make it possible for researchers in these fields to implement this tool into a replicable and reproducible workflow in R.

Acknowledgements

This work was supported by The Nordic Joint Committee for Agricultural and Food Research (grant num. 202100-2817). The idea for this package was conceived during the course "Analysing Spatial Data" at the Norwegian School of Economics (NHH), we thank Professor Roger Bivand for his insights.

References

Aguilar, E., Peterson, T. C., Obando, P. R., Frutos, R., Retana, J. A., Solera, M., Soley, J., et al. (2005). Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. *Journal of Geophysical Research*, *110*(D23), D23107. doi:10.1029/2005JD006119

Anderson, M. C., Hain, C., Wardlow, B., Pimstein, A., Mecikalski, J. R., & Kustas, W. P. (2011). Evaluation of Drought Indices Based on Thermal Remote Sensing of Evapotranspiration over the Continental United States. *Journal of Climate*, *24*(8), 2025–2044. doi:10.1175/2010JCLI3812.1

Auguie, B. (2017). *gridExtra: Miscellaneous Functions for "Grid" Graphics*. Retrieved from https://CRAN.R-project.org/package=gridExtra

Chamberlain, S. (2019). *crul: HTTP Client*. Retrieved from https://CRAN.R-project.org/package=crul

de Sousa, K., Casanoves, F., Sellare, J., Ospina, A., Suchini, J. G., Aguilar, A., & Mercado, L. (2018). How climate awareness influences farmers' adaptation decisions in Central America? *Journal of Rural Studies*, *64*, 11–19. doi:10.1016/j.jrurstud.2018.09.018

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., et al. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, *2*, 150066. doi:10.1038/sdata.2015.66

Horn, L. M., Hajat, A., Sheppard, L., Quinn, C., Colborn, J., Zermoglio, M. F., Gudo, E. S., et al. (2018). Association between precipitation and diarrheal disease in Mozambique. *International Journal of Environmental Research and Public Health*, *15*(4). doi:10.3390/ijerph15040709

Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., & Xie, P. (2014). Integrated Multi-satellitE Retrievals for GPM (IMERG), version 4.4. NASA's Precipitation Processing Center. Retrieved from ftp://arthurhou.pps.eosdis.nasa.gov/gpmdata/

Kehel, Z., Crossa, J., & Reynolds, M. (2016). Identifying Climate Patterns during the Crop-Growing Cycle from 30 Years of CIMMYT Elite Spring Wheat International Yield Trials. In A. Bari, A. B. Damania, M. Mackay, & S. Dayanandan (Eds.), *Applied Mathematics and Omics to Assess Crop Genetic Resources for Climate Change Adaptive Traits* (pp. 151–174). CRC Press. doi:10.1201/b19518-17

Ooms, J. (2014). The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects. *arXiv:1403.2805 [stat.CO]*. Retrieved from https://arxiv.org/abs/1403.2805

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439–446. doi:10.32614/RJ-2018-009

R Core Team. (2020). *R: A language and environment for statistical computing. Version 4.0.0.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

SERVIR Global. (2019a). *Evaporative stress index (ESI)*. National Aeronautics and Space Administration (NASA); United States Agency for International Development (USAID). Retrieved from https://climateserv.servirglobal.net/

SERVIR Global. (2019b). *ClimateSERV*. National Aeronautics and Space Administration (NASA); United States Agency for International Development (USAID). Retrieved from https://climateserv.servirglobal.net/

Thomson, M. C., Ukawuba, I., Hershey, C. L., Bennett, A., Ceccato, P., Lyon, B., & Dinku, T. (2017). Using rainfall and temperature data in the evaluation of national malaria control programs in Africa. *American Journal of Tropical Medicine and Hygiene*, *97*, 32–45. doi:10.4269/ajtmh.16-0696

van Etten, J., de Sousa, K., Aguilar, A., Barrios, M., Coto, A., Dell'Acqua, M., Fadda, C., et al. (2019). Crop variety management for climate adaptation supported by citizen science. *Proceedings of the National Academy of Sciences*, *116*(10), 4194–4199. doi:10.1073/pnas.1813720116

Wickham, H., Averick, M., Bryan, J., Chang, W., D', L., Mcgowan, A., François, R., et al. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. doi:10.21105/joss.01686

6

The R package `climatrends` hex logo

Climate variability indices for ecological and crop models in R: the `climatrends` package

A preprint

Kauê de Sousa[1,2][*], Jacob van Etten[2], Svein Øivind Solberg[1]
[1] Department of Agricultural Sciences, Inland Norway University of Applied Sciences, 2318 Hamar, Norway
[2] Bioversity International, 00054 Maccarese, Rome, Italy
[*] Correspondence should be addressed to: kaue.desousa@inn.no

Summary

Abiotic factors play an important role in most ecological and crop systems that depend on certain levels of temperature, light and precipitation (and their interplay) to initiate important physiological events (Schulze et al., 2019). In the walk of climate change, understanding how these factors drive the physiological processes is a key approach to provide recommendations for adaptation and biodiversity conservation. The package `climatrends` aims to provide the methods in R (R Core Team, 2020) to compute precipitation and temperature indices that serve as input for climate and crop models (Kehel, Crossa, & Reynolds, 2016; van Etten et al., 2019), trends in climate change (Aguilar et al., 2005; de Sousa et al., 2018) and applied ecology (Prentice et al., 1992; Liu & El-Kassaby, 2018).

Methods and features

Implementation

Six main functions are provided (Table 1), `crop_sensitive()`, `ETo()`, `GDD()`, `late_frost()`, `rainfall()` and `temperature()` with a default method for numeric 'vector' and additional methods implemented via the package `methods` (R Core Team, 2020) for classes 'matrix' (or array), 'data.frame', and 'sf' (of geometry POINT or POLYGON) (Pebesma, 2018). The last two methods are designed to fetch data from cloud sources, until now from the packages `nasapower` (Sparks, 2018) and `chirps` (de Sousa, Sparks, Ashmall, van Etten, & Solberg, 2020).

Table 1: Main functions available in climatrends.

Function	Definition
crop_sensitive()	Compute crop sensitive indices
ETo()	Reference evapotranspiration using the Blaney-Criddle method
GDD()	Compute growing degree-days
late_frost()	Compute the occurrence of late-spring frost
rainfall()	Precipitation indices
temperature()	Temperature indices

These functions started as a set of scripts to compute indices in citizen science trials. In these trials. Aiming to capture the environmental variation across different sites, which can differ since each data-point generally have a different starting day and duration, the arguments `day.one` and `span` are vectorised and may be variable across data-points. For time series analysis, where fixed periods are defined across many locations, the indices can be adjusted with the argument `last.day` linked to the argument `day.one`.

Temperature and precipitation indices

The package `climatrends` computes 12 temperature indices and 10 precipitation indices that were suggested by previous research on climatology and crop science (Aguilar et al., 2005; Kehel et al., 2016). The indices computed by the functions `temperature()` and `rainfall()` are described in Table 2.

Table 2: Temperature and precipitation indices available in climatrends.

Index	Definition	Unit
maxDT	Maximun day temperature	°C
minDT	Minimum day temperature	°C
maxNT	Maximun night temperature	°C
minNT	Minimum night temperature	°C
DTR	Diurnal temperature range (mean difference between DT and NT)	°C
SU	Summer days, number of days with maximum temperature > 30 °C	days
TR	Tropical nights, number of nights with maximum temperature > 25 °C	days
CFD	Consecutive frosty days, number of days with temperature < 0 °C	days
WSDI	Maximum warm spell duration, consecutive days with temperature > 90th percentile	days
CSDI	Maximum cold spell duration, consecutive nights with temperature < 10th percentile	days
T10p	The 10th percentile of night temperature	°C
T90p	The 90th percentile of day temperature	°C
MLDS	Maximum length of consecutive dry day, rain < 1 mm	days
MLWS	Maximum length of consecutive wet day, rain >= 1 mm	days
R10mm	Heavy precipitation days 10 >= rain < 20 mm	days
R20mm	Very heavy precipitation days rain >= 20	days
Rx1day	Maximum 1-day precipitation	mm
Rx5day	Maximum 5-day precipitation	mm
R95p	Total precipitation when rain > 95th percentile	mm
R99p	Total precipitation when rain > 99th percentile	mm
Rtotal	Total precipitation in wet days, rain >= 1 mm	mm
SDII	Simple daily intensity index, total precipitation divided by the number of wet days	mm/days

Growing degree-days

Growing degree-days (gdd) is an heuristic tool in phenology that measures heat accumulation and is used to predict plant and animal development rates (Prentice et al., 1992). Growing degree-days are calculated by taking the integral of warmth above a base temperature (T_0). The function `GDD()` applies by default the following equation.

Equation [1]

$$GDD = \frac{T_{max} + T_{min}}{2} - T_0$$

where T_{max} is the maximum temperature in the given day, T_{min} is the minimum temperature in the given day and T_0 is the minimum temperature for growth (as per the physiology of the focal organism or ecosystem averages).

Additionally, the function `GDD()` offers three modified equations designed for cold environments and for tropical environments. For cold environments, where T_{min} may be lower than T_0, there are two modified equations that adjust either T_{mean} (variant a) or T_{min} (variant b). The variant a changes T_{mean} to T_0 if $T_{mean} < T_0$ and is expressed as follow.

Equation [2]

$$GDD = max\left(\frac{T_{max} + T_{min}}{2} - T_0,\ 0\right)$$

The variant b, is calculated using Equation 1, but adjusts T_{min} or T_{max} to T_0 if $T < T_0$, the equation is adjusted as follows.

Equation [3]

$$T < T_0\ \rightarrow\ T = T_0$$

where T may refer to T_{min} and/or T_{max} when the condition of being below T_0 applies.

For tropical areas, where the temperature may surpass a maximum threshold ($T_{0_{max}}$), resulting in limited development, the minimum temperature is adjusted using Equation 3 and the maximum temperature is adjusted to a maximum base temperature as follow.

Equation [4]

$$T_{max} > T_{0_{max}}\ \rightarrow\ T_{max} = T_{0_{max}}$$

where $T_{0_{max}}$ is the maximum base temperature for growth, defined in `GDD()` using the argument `tbase_max`.

These modified equations are defined as 'a', 'b' and 'c', respectively, and can be selected using the argument `equation`.

By default, the function returns the degree-days that is accumulated over the time series using Equation 1. Additionally, the function may return the daily values of degree-days or the number of days that a given organism required to reach a certain number of accumulated degree-days. These values are defined by 'acc', 'daily' or 'ndays' and can be adjusted using the argument `return.as`. The required accumulated gdd is defined with argument `degree.days`. For example, the Korean pine (*Pinus koraiensis*) requires 105 $°C$ accumulated gdd to onset the photosynthesis (Wu, Guan, Yuan, Wang, & Jin, 2013). In that case, `GDD()` will calculate the growing degree-days (gdd) and sum up the values until it reaches 105 $°C$ and return the number of days required in the given season (GDD_r), as follows.

Equation [5]

$$\parallel GDD_r \parallel = \; ggd_1 \; + \; ... \; + \; gdd_n$$

where GDD_r is the length of the vector with accumulated degree-days from day 1 to n.

Late-spring frost

Late-spring frost is a freezing event occurring after a substantial accumulation of warmth. Frost damage is a known issue in temperate and boreal regions, it is associated with the formation of extracellular ice crystals that cause damage in the membranes (Lambers, Chapin III, & Pons, 2008). Freezing occurring after an advanced phenological stage during spring may harm some plant species, resulting in lost of productivity in crop systems (Trnka et al., 2014) and important ecological impacts (Zohner et al., 2020).

The function `late_frost()` supports the computation of late-spring frost events. The function counts for the number of freezing days with minimum temperature below a certain threshold (argument `tfrost`). And returns the number of days spanned by frost events (temperature below `tfrost`), latency (event with no freezing temperature but also no accumulation of growing degree-days) and warming (when growing degree-days are accumulated enabling the development of the target organism). Additionally the function returns the first day of the events. The function calculates the growing degree-days applying the variant b (Eq. 3), which can be adjusted using the argument `equation` passed to `GDD()` as explained in the later section. The main inputs are a vector with maximum and minimum temperatures to compute the degree-days, a vector of dates (argument `date`), and, if needed, the `tbase` and `tfrost`, set by default to 4 and -2 $°C$.

Crop-related indices

Two functions in **climatrends** are mainly designed to capture the effects of climate on the development and stress of crop species, `crop_sensitive()` computes indices that aim to capture the changes in temperature extremes during key phenological stages (e.g. anthesis), and `ETo()` computes the reference evapotranspiration.

The crop sensitive indices available in **climatrends** are described in Table 3. These indices were previously used in crop models to project the impacts of climate change on crop yield (Trnka et al., 2014; Challinor, Koehler, Ramirez-Villegas, Whitfield, & Das, 2016). Each index has a default temperature threshold(s) which can be adjusted by using the arguments `*.threshold`. Where the `*` means the index. For example, to change the defaults for hts_max (high temperature stress), a vector with the temperature thresholds is passed through the argument `hts_max.thresholds`.

The reference evapotranspiration measures the influence of the climate on a given plant's water need (Brouwer & Heibloem, 1986). The function `ETo()` applies the Blaney-Criddle method, a general theoretical method used when only air-temperature is available locally. It should be noted that this method is not very accurate and aims to provide the order of magnitude of evapotranspitation. The reference evapotranspiration is calculated using the following equation.

Equation [6]

$$ETo = p \times \left(0.46 \times \frac{T_{max} + T_{min}}{2} + 8\right) \times K_c$$

Where p is the mean daily percentage of annual daytime hours, T_{max} is the maximum temperature, T_{min} is the minimum temperature, and K_c is the factor for organism water need.

The percentage of daytime hours (p) is calculated internally by the 'data.frame' and 'sf' methods in `ETo()` using the given latitude (taken from the inputted `object`) and date (taken from the inputted `day.one`). It matches the latitude and date with a table of daylight percentage derived from Brouwer and Heibloem (1986). The table can be verified using `climatrends:::daylight`.

Table 3: Crop sensitive indices computed by climatrends.

Index	Definition	Default thresholds
hts_mean	High temperature stress using daily mean temperature, and given as percentage number of days a certain threshold is exceeded	32, 35, 38 °C
hts_max	High temperature stress using daily max temperature, and given as percentage number of days a certain threshold is exceeded	36, 39, 42 °C
hse	Heat stress event, and given as percentage number of days a certain threshold is exceeded for at least two consecutive days	31 °C
hse_ms	Heat stress event, and given the maximum number of days a certain threshold is exceeded for at least two consecutive days	31 °C
cdi_mean	Crop duration index using daily mean temperature, and given as max(Tmean - threshold, 0)	22, 23, 24 °C
cdi_max	Crop duration index using daily max temperature, and given as max(Tmax - threshold, 0)	27, 28, 29 °C
lethal	Lethal temperatures, defined as percentage of days during the timeseries where daily mean temperature exceeds a given threshold	43, 46, 49 °C

Examples

Common bean

During five growing seasons (from 2015 to 2017) in Nicaragua, van Etten et al. (2019) conducted a crowdsourcing citizen-science experiment testing 11 common bean varieties (*Phaseolus vulgaris* L.) in 842 farmer-managed plots. Sets of three varieties were allocated randomly to farms as incomplete blocks. A Plackett–Luce model was used to analyse the data, this model estimates for each variety the probability that it wins, beating all other varieties in the set (Turner, van Etten, Firth, & Kosmidis, 2020). An earlier version of `climatrends` was used in this research to capture the seasonal climate variation, here we reproduce part of this analysis regarding calculation and application of the climate indices. The approach here is slightly different because it considers the growing-degree days from planting date to maturity (the earlier study used planting date to the end of reproductive stage) and add new indices to illustrate the package implementation.

The data used here is available as supplementary material as `cbean`. This contains a data.frame with a Plackett-Luce grouped rankings, the geographical coordinates of each sampled plot and the planting dates from where each farmer decided to start the experiment. The planting dates differ from each other in the same season. The temperature data used was the land surface temperature MODIS (MYD11A2) (Wan, Hook, & Hulley, 2015) and is storaged as an array with two layers (1st for the day and 2nd for the night temperatures). Each column corresponds to the dates (from 2015-09-10 to 2017-06-09) and the rows corresponds to the rows in the `cbean` data.frame.

Since the phenological stages were not available, we estimate these stages based on the amount of growing degree-days required to reach a given stage using the function `GDD()`. For common beans, we define 900 degree-days, from planting date to maturity (de Medeiros, Daniel, & Fengler, 2016). The input data is the array with the temperature data, the vector with planting dates (`cbean$planting_date`), the required amount of degree-days passed to the argument `degree.days` and the character string 'ndays' specifying that the function must return the values as number of days. `GDD()` calls internally the function `get_timeseries()` which will match the given dates in `day.one` with the column names in the array and concatenate the values for each row. Then `GDD()` computes the degree-days for the time series and return the length of the vector where the accumulated gdd reached the pre-defined threshold (900).

The degree-days spanned from 54 to 100 days as shown in Fig. 1a. For simplicity we take the average per season and use this vector to compute the temperature indices.

```
library("climatrends")
library("PlackettLuce")
library("tidyverse")

# compute the number of days required to accumulate
# gdd from planting date to maturity
gdd <- GDD(modis,
           day.one = cbean$planting_date,
           degree.days = 900,
           return.as = "ndays")

# add gdd to the cbean data and take the average
# of gdd per season
cbean %<>%
  mutate(gdd = gdd$gdd) %>%
  group_by(season) %>%
  mutate(gdds = as.integer(mean(gdd)))
```

To compute the temperature indices we use the array with tempeture data, the vector with planting dates, and the seasonal averaged degree-days passed as a vector using the argument `span`. `temperature()` concatenates the data from the given `day.one` to the given `span` and compute the indices for each row.

In van Etten (2019), a forward variable selection was applied to retain the most representative covariates based on the deviance reduction. This analysis retained the maximum night temperature (maxNT) as the most representative covariate. To illustrate how the Plackett-Luce trees can grow in complexity as we add more indices, we included the summer days (SU, number of days with maximum day temperature $> 30\ °C$) together with maxNT.

```
# compute the temperature indices from planting date to the
# number of days required to accumulate the gdd in each season
temp <- temperature(modis,
                    day.one = cbean$planting_date,
                    span = cbean$gdds)

# combine the indices with the main data
cbean <- cbind(cbean, temp)

# fit a Plackett-Luce tree
plt <- pltree(G ~ maxNT + SU, data = cbean, minsize = 50)
```

Across-season distribution of maxNT captured for each sample plot in this experiment is shown in Fig. 1b. The data has a bimodal distribution which is reflected in the splitting value ($18.7\ °C$) for the Plackett-Luce trees in Fig. 1c. The upper node splits with 49 summer days (SU). We can interpret these results as that differences in growing performance of common beans is led by a considerable amount of diurnal temperature above a warmer threshold of $30\ °C$ (in this case >70% of the growing days) and warmer nights ($> 18.7\ °C$).

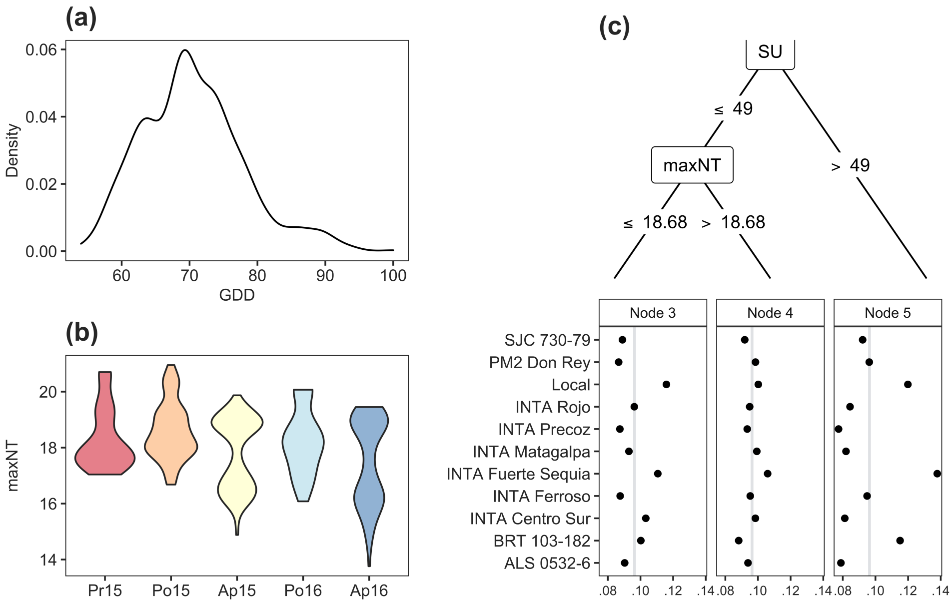

Fig. 1. Application of climatrends functions to support the analysis of a citizen-science data testing 11 common bean varieties in Nicaragua. (A) Days required to reach 900 growing-degree days from planting date calculated using the function GDD(). (B) Maximum night temperature (°C) distributed across seasons computed using the function temperature(). (C) Plackett-Luce Tree showing the probability of one common bean variety has to win against the others (axys X) in three different nodes splitted with the summer days (day temperature > 30 °C) and maximum night temperature (°C). Note: the first season (primera, Pr) spans from May to August, the second (postrera, Po) from September to October, and the third (apante, Ap) from November to January.

Trends in climate variability in Norway and Sweden

We randomly selected 100 points in hexagonal within the coordinates 7° and 17° W, and 59 ° and 63 ° N, that comprises Norway and Sweden before the Arctic Circle. We compute the temperature indices from 2000-01-01 to 2019-12-31 using the function `temperature()` with the method for objects of class 'sf'. The temperature data is fetched from from the NASA Langley Research Center POWER Project funded through the NASA Earth Science Directorate Applied Science Program (https://power.larc.nasa.gov/), using the R package `nasapower` (Sparks, 2018).

```
library("climatrends")
library("sf")
library("nasapower")

# create a polygon within the coordinates 7, 17, 59, 63
e <- matrix(c(7, 59, 17, 59, 17, 63,
              7, 63, 7, 59),
            nrow = 5, ncol = 2, byrow = TRUE)

e <- st_polygon(list(e))

# sample 100 points in the hexagonal type
p <- st_sample(e, 100, type = "hexagonal")
p <- st_as_sf(p, crs = 4326)

# compute the temperature indices using the random points
temp <- temperature(p,
                    day.one = "2000-01-01",
                    last.day = "2019-12-31",
                    timeseries = TRUE,
                    intervals = 365)
```

We then select the indices CSDI (cold spell duration of night temperature), WSDI (warm spell duration of day temperature), and their associated indices the T10p (the 10th percentile of night temperature) and T90p (the 90th percentile of day temperature), in Figure 2. Plots are generated with `ggplot2` (Wickham, 2016) and `patchwork` (Pedersen, 2020).

The trends show a decrease in the cold spell duration (number of consecutive cold nights bellow the 10th percentile) and warm spell duration (number of consecutive warm days above the 90th percentile). However, the values of the percentiles show an increase over the time series. The T10p index shows a decrease around the year of 2010, but again rises up to the a value around the -10 °C, meaning that the could nights are becoming a bit warmer over the time. The T90p index also shows an increase in the temperature across the sampled area, with the average 90th percentile rising from ~ 16 °C to ~ 18 °C over the time series.

Further development

The package can support the integration with other datasets as they become available in `R` via API client packages. Also new indices related to the physiology of crops could be implemented.

Acknowledgements

This work was supported by The Nordic Joint Committee for Agricultural and Food Research (grant num. 202100-2817). We thank Julian Ramirez-Villegas and Marcel Schrijvers-Gonlag for the useful insights and discussion that helped in the development of this study.

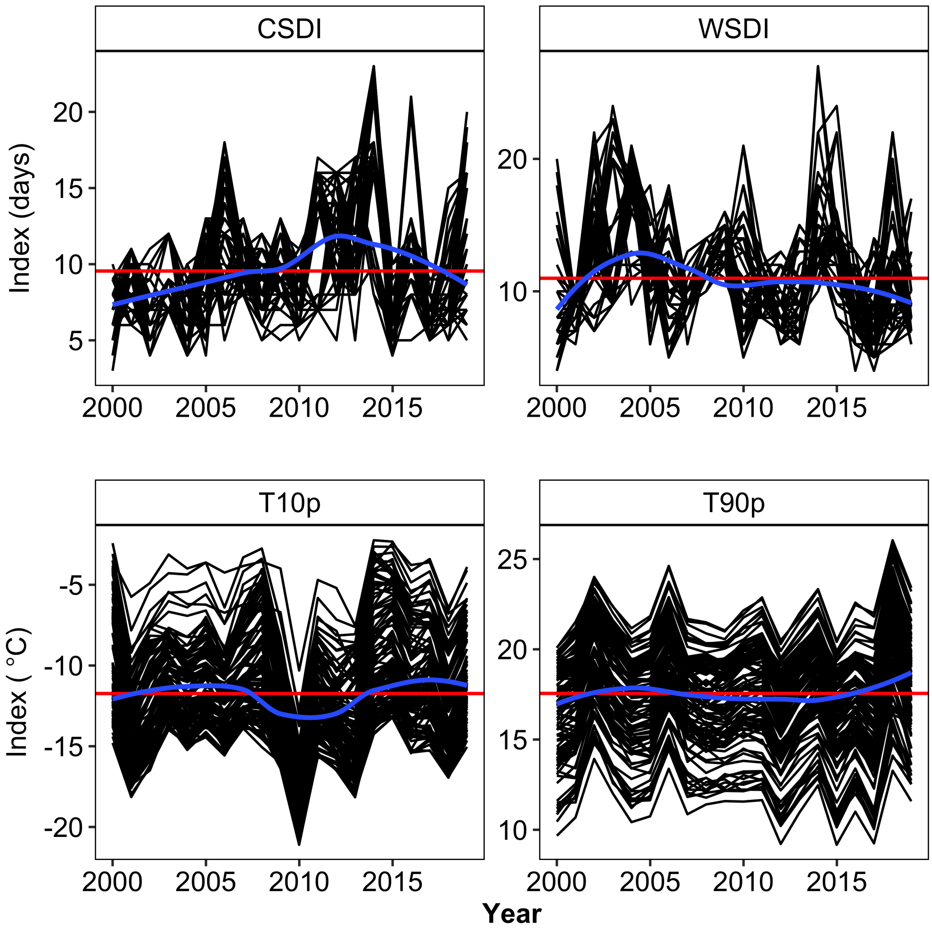

Fig. 2. Trends in temperature indices across Southern Norway and Sweden from 2000 to 2019. CSDI, maximum cold spell duration, consecutive nights with temperature < 10th percentile. WSDI, maximum warm spell duration, consecutive days with temperature > 90th percentile. T10p, the 10th percentile of night temperature. T90p, the 90th percentile of day temperature. Red line indicates the historical mean of each index in the time series. Blue line indicates the smoothed trends in each index using the 'loess' method.

Data availability statement

To explore the latest functionalities of `climatrends`, please check the package's updates at CRAN (https://cran.r-project.org/package=climatrends).

References

Aguilar, E., Peterson, T. C., Obando, P. R., Frutos, R., Retana, J. A., Solera, M., … Mayorga, R. (2005). Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. *Journal of Geophysical Research*, *110*(D23), D23107. doi:10.1029/2005JD006119

Brouwer, C., & Heibloem, M. (1986). *Irrigation water management: Irrigation water needs* (Training m). Rome, Italy: Food; Agriculture Organization of The United Nations. Retrieved from http://www.fao.org/3/S2022E/s2022e00.htm

Challinor, A. J., Koehler, A.-K., Ramirez-Villegas, J., Whitfield, S., & Das, B. (2016). Current warming will reduce yields unless maize breeding and seed systems adapt immediately. *Nature Climate Change*, *6*(10), 954–958. doi:10.1038/nclimate3061

de Medeiros, G. A., Daniel, L. A., & Fengler, F. H. (2016). Growth, Development, and Water Consumption of Irrigated Bean Crop Related to Growing Degree-Days on Different Soil Tillage Systems in Southeast Brazil. *International Journal of Agronomy*, *2016*, 8065985. doi:10.1155/2016/8065985

de Sousa, K., Casanoves, F., Sellare, J., Ospina, A., Suchini, J. G., Aguilar, A., & Mercado, L. (2018). How climate awareness influences farmers' adaptation decisions in Central America? *Journal of Rural Studies*, *64*, 11–19. doi:10.1016/j.jrurstud.2018.09.018

de Sousa, K., Sparks, A., Ashmall, W., van Etten, J., & Solberg, S. (2020). chirps: API Client for the CHIRPS Precipitation Data in R. *Journal of Open Source Software*, *5*(51), 2419. doi:10.21105/joss.02419

Kehel, Z., Crossa, J., & Reynolds, M. (2016). Identifying Climate Patterns during the Crop-Growing Cycle from 30 Years of CIMMYT Elite Spring Wheat International Yield Trials. In A. Bari, A. B. Damania, M. Mackay, & S. Dayanandan (Eds.), *Applied mathematics and omics to assess crop genetic resources for climate change adaptive traits* (pp. 151–174). CRC Press.

Lambers, H., Chapin III, F. S., & Pons, T. L. (2008). *Plant Physiological Ecology* (Second Edi, p. 620). doi:10.1007/978-0-387-78341-3

Liu, Y., & El-Kassaby, Y. A. (2018). Evapotranspiration and favorable growing degree-days are key to tree height growth and ecosystem functioning: Meta-analyses of Pacific Northwest historical data. *Scientific Reports*, *8228*. doi:10.1038/s41598-018-26681-1

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, *10*(1), 439–446. doi:10.32614/RJ-2018-009

Pedersen, T. L. (2020). *patchwork: The Composer of Plots*. Retrieved from https://CRAN.R-project.org/package=patchwork

Prentice, I. C., Cramer, W., Harrison, S. P., Leemans, R., Monserud, R. A., & Solomon, A. M. (1992). Special Paper: A Global Biome Model Based on Plant Physiology and Dominance, Soil Properties and Climate. *Journal of Biogeography*, *19*(2), 117. doi:10.2307/2845499

R Core Team. (2020). R: A language and environment for statistical computing. version 4.0.0. Vienna, Austria: CRAN R Project. Retrieved from https://www.r-project.org/

Schulze, E.-D., Beck, E., Buchmann, N., Clemens, S., Müller-Hohenstein, K., & Scherer-Lorenzen, M. (2019). *Plant Ecology* (Second Edi, p. 925). Berlin, Germany: Springer. doi:10.1007/978-3-662-56233-8

Sparks, A. H. (2018). nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R. *Journal of Open Source Software*, *3*(30), 1035. doi:10.21105/joss.01035

Trnka, M., Rötter, R. P., Ruiz-Ramos, M., Kersebaum, K. C., Olesen, J. E., Žalud, Z., & Semenov, M. A. (2014). Adverse weather conditions for European wheat production will become more frequent with climate change. *Nature Climate Change*, *4*(7), 637–643. doi:10.1038/nclimate2242

Turner, H. L., van Etten, J., Firth, D., & Kosmidis, I. (2020). Modelling rankings in R: the PlackettLuce package. *Computational Statistics*. doi:10.1007/s00180-020-00959-3

van Etten, J., de Sousa, K., Aguilar, A., Barrios, M., Coto, A., Dell'Acqua, M., … Steinke, J. (2019). Crop variety management for climate adaptation supported by citizen science. *Proceedings of the National Academy of Sciences*, *116*(10), 4194–4199. doi:10.1073/pnas.1813720116

Wan, Z., Hook, S., & Hulley, G. (2015). MYD11A1 MODIS/Aqua Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006 [Data set]. *NASA EOSDIS LP DAAC*. doi:10.5067/MODIS/MYD11A1.006

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wu, J., Guan, D., Yuan, F., Wang, A., & Jin, C. (2013). Soil Temperature Triggers the Onset of Photosynthesis in Korean Pine. *PLoS ONE*, *8*(6), e65401. doi:10.1371/journal.pone.0065401

Zohner, C. M., Mo, L., Renner, S. S., Svenning, J.-C., Vitasse, Y., Benito, B. M., … Crowther, T. W. (2020). Late-spring frost risk between 1959 and 2017 decreased in North America but increased in Europe and Asia. *Proceedings of the National Academy of Sciences*, 201920816. doi:10.1073/pnas.1920816117

Inland Norway
University of
Applied Sciences

Increased agrobiodiversity through farm diversification and varietal selection is an alternative to help farmers to cope with the negative effects of climate change while ensuring food security. However, such approach have been difficult to scale up, since we often lack information to understand the contexts that drive farmers' adaptation decisions and how to develop recommendations for adaptation.

This thesis presents results from three continents to improve this understanding, specifically in smallholder farming. It provide insights for the different biological levels: species, focusing on trees as slow grower organisms for interspecific diversification; varieties, looking for locally adapted phenotypes; and genotypes, focusing on genotype by environmental interactions.

The results show that farmers have a clear preference to a set of adaptation strategies, with agroforestry as the first choice. The most preferred trees in coffee and cocoa agroforestry are the most vulnerable, but farmers could re-think the agroforestry designs using a portfolio of underutilised species already present in low densities at the current systems. At the variety level, the results show that scaling agricultural experimentation with citizen science can support recommendations for crop variety management for climate adaptation. Also, that linking farmer-generated data to scientist-generated data can support breeding programs targeting challenging crop production environments. Overall, the results of this thesis should be seen as starting point to develop lines of research that support recommendations to adapt agricultural systems to a changing climate.