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

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.

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, includ-

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).

ing 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.

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
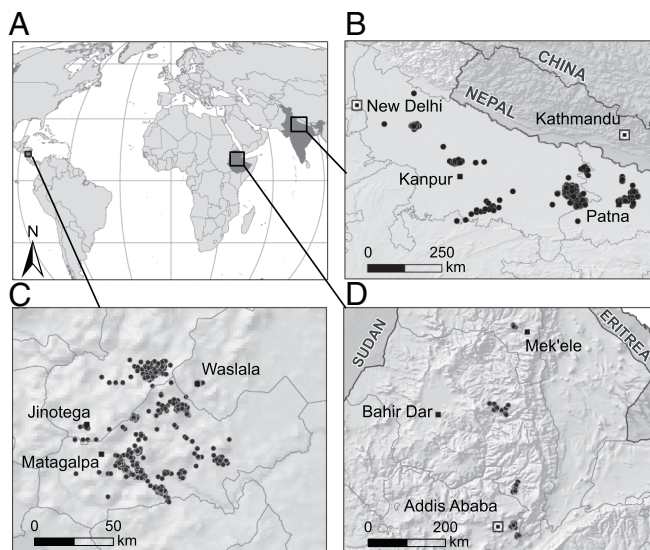

Fig. 1. Research sites: (*A*) overview, (*B*) India, (*C*) Nicaragua, and (*D*) Ethiopia. Farms included in the trials are indicated as dots.
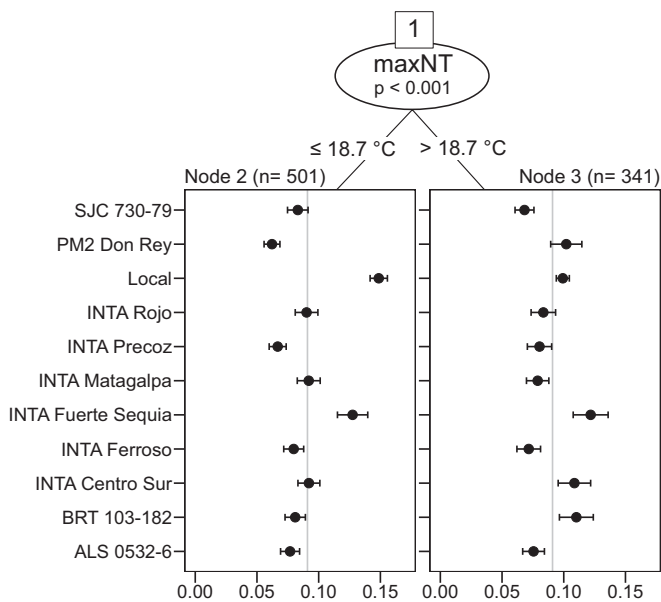
Fig. 2. Plackett–Luce trees of tricot trial data and associated climatic data for common bean in Nicaragua. The horizontal axis of each panel is the probability of winning of varieties. Error bars show quasi-SEs. The gray vertical lines indicate the average probability of winning (1/number of varieties). 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. 3A). 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. 3B). 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. 3B. 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
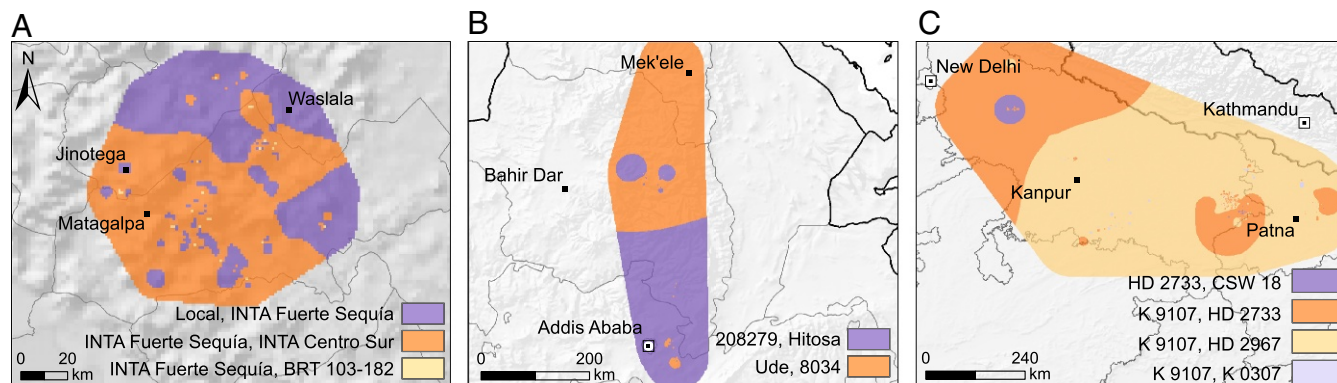
Fig. 3. Variety recommendations based on average season predictions from PLTs using climatic variables for (A) common bean in Nicaragua (Apante season), (B) durum wheat in Ethiopia (Meher season), and (C) bread wheat in India (Rabi season). Map categories show the top two varieties for each area according 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:

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

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

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

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:

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

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:

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

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

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

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*.

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.