# RTB Working Paper

# The tricot citizen science approach applied to on-farm variety evaluation: methodological progress and perspectives

Jacob van Etten, Erna Abidin, Elizabeth Arnaud, David Brown, Edward Carey, Marie-Angélique Laporte, Isabel López-Noriega, Brandon Madriz, Rhys Manners, Berta Ortiz-Crespo, Carlos Quirós, Kauê de Sousa, Béla Teeken, Hale Ann Tufan, Jacob Ulzen, José F Valle-Soto

DECEMBER 2020

RESEARCH PROGRAM ON Roots, Tubers and Bananas

**RTB Working Paper**

## The tricot citizen science approach applied to on-farm variety evaluation: methodological progress and perspectives

**Published by the CGIAR Research Program on Roots, Tubers and Bananas**

The CGIAR Research Program on Roots, Tubers and Bananas (RTB) is a partnership collaboration led by the International Potato Center implemented jointly with the Alliance of Bioversity International and the International Center for Tropical Agriculture (CIAT), the International Institute of Tropical Agriculture (IITA), and the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), that includes a growing number of research and development partners. RTB brings together research on its mandate crops: bananas and plantains, cassava, potato, sweetpotato, yams, and minor roots and tubers, to improve nutrition and food security and foster greater gender equity especially among some of the world's poorest and most vulnerable populations.
www.rtb.cgiar.org/

**Contact**:
RTB Program Management Unit
International Potato Center (CIP)
Apartado 1558, Lima 12, Peru
rtb@cgiar.org • www.rtb.cgiar.org

ISSN: 2309-6586
DOI: 10.4160/23096586RTBWP20212

**Correct citation:**
van Etten, J., Abidin, E., Arnaud, D., Brown, E., Carey, E., Laporte, M.-L., López-Noriega, I., Madriz, B., Manners, R., Ortiz-Crespo, B., Quirós, C., de Sousa, K, Teeken, B., Tufan, H.A., Ulzen, J., & Valle-Soto, J. (2020). *The tricot citizen science approach applied to on-farm variety evaluation: methodological progress and perspectives.* Lima, Peru: CGIAR Research Program on Roots, Tubers and Bananas (RTB). RTB Working Paper. Nº. 2021-2. www.rtb.cgiar.org

**December 2020**

# Contents

# Abstract

Tricot (triadic comparisons of technologies) is a citizen science approach for testing technology options in their use environments, which is being applied to on-farm testing of crop varieties. Over the last years, important progress has been made on the tricot methodology of which an overview is given. Trial dimensions depend on several factors but tricot implies that plot size is as small as possible to include farmers with small plots (yet avoiding excessive interplot competition) while many locations are included to ensure representativeness of trials. Gender and socio-economic work is focused on better household characterization and recruitment strategies that move beyond sex-aggregation to address aspects of intersectionality. Ethics, privacy and traditional knowledge aspects will be addressed through expanding digital support in this direction. Genetic gain estimates need to be addressed by yield measurements, which can be generated by farmers themselves. There is conceptual clarity about the needs for documentation of trials and publishing data but this aspect requires further digital development. Much progress has been made on the ClimMob digital platform already, which is user friendly and supports trials in the main steps and includes open-source data analytics packages. Further improvements need to be made to ensure better integration with other tools. A next step will be the development of scaling strategies that involve business development. An important input into these strategies are economic studies, which are ongoing.

# Acknowledgments

This research was undertaken as part of, and funded by, the CGIAR Research Program on Roots, Tubers and Bananas (RTB) and supported by CGIAR Trust Fund contributors.

# The tricot citizen science approach applied to on-farm variety evaluation: methodological progress and perspectives

## 1 INTRODUCTION

*Tricot* (triadic comparisons of technologies, pronounced "try-cot") is a citizen science approach for testing technology options in their use environments, originally conceived in 2011 (van Etten 2011). The Oxford English Dictionary defines citizen science as "the collection and analysis of data relating to the natural world by members of the general public, typically as part of a collaborative project with professional scientists". Different definitions are given by others, but our use of it is not far from this one. As a citizen science approach, tricot actively involves non-scientists in experimental data generation and interpretation. This follows a broader movement of applying citizen science and crowdsourcing methods in research on food and agriculture, providing a fresh lease of life to participatory agricultural research (van de Gevel et al., 2020; Ryan et al., 2018; Minet et al., 2017).

Tricot addresses important challenges that have plagued on-farm testing, which is an often-underrated activity in the agricultural sciences (Kool, Andersson, and Giller, 2020). Also, the approach is increasingly used for areas closely related to on-farm testing of varieties, such as fertilizer testing (AKILIMO scaling project by IITA and CIP in Rwanda) and food product testing for sweetpotato implemented by CIP (Moyo et al., 2020, 2021). Testing technologies in their use environments is important for *external validity* of experiments, the degree to which the findings have application outside of the experimental setting. To overcome common issues in user testing, the tricot approach streamlines the approach through digital support throughout the experimental cycle, simplifies the experimentation format to make user participation easy, and enhances data analysis by enriching it with data about the user context.

The method was first implemented and tested in the period between 2013 and 2016 for on-farm testing of varieties, and an earlier article reported about methodological progress in this period (van Etten et al., 2019). Much of this work was part of the Seeds for Needs initiative, aiming at broadening the range of varietal diversity to farmers to adapt to climate change (Fadda et al., 2020). These projects were focused on cereals and grain legumes. Since then, the tricot approach has been used for other trials, by different organizations (including private sector) and for different applications (food products, fertilizers, etc.), and for clonal crops (cassava, sweetpotato, potato), vegetables, and a perennial crop (cocoa). The present article reports on 1) methodological progress; 2) discusses important considerations that implementers of the tricot approach need to consider; and 3) areas open for future research.

# 2 DESCRIPTION OF THE TRICOT APPROACH

The tricot approach has been described in several publications (van Etten et al., 2019; van Etten, 2011; van Etten et al., 2020; Steinke et al., 2017; van Etten et al., 2019; Fadda et al., 2020; Beza et al., 2017). Here, a short synthesis will be provided.

The word 'tricot' is derived from triadic comparisons of technology options. 'Triadic' refers to the sets of three technology options that are compared by each participant. Tricot enables many citizen scientists doing a small experiment while contributing to answering a larger question. Researchers and citizen science participants are supported throughout the experiment cycle by digital tools to design, execute, monitor and analyze the trials. As many citizen scientists contribute and do experiments in their typical use environments using their usual practices, it becomes possible to start to understand how variation in environments and practices affects the results.

The particular way in which tricot works makes these steps possible. The following aspects are key to tricot:

1. the use of incomplete blocks of three items – to make the threshold of participation low in terms of farm size, and reduce resource needs and training required;
2. the use of ranking as the main way to report observations -- to facilitate digital data collection and to make it possible to evaluate a tricot plot with very little training (in contrast, scoring requires calibration and absolute yield measurements require training);
3. the limited control of experimental conditions – following common local technology use practices to maximize external validity;
4. the use of a streamlined digital process from trial design to analysis – to make it manageable, executable with many participants, to reduce errors, to reduce costs, and to quickly deliver feedback to achieve high motivation and impact on subsequent decisions;
5. early feedback of the results to the participants -- to provide ownership to and stimulate engagement of participating "citizen scientists" and to validate results.

Tricot builds on existing participatory research formats that have been used in the past, as documented by (Van Etten et al., 2019). The novelty of the format is the combination of the different elements in a standardized, widely used approach supported by a corresponding digital platform, ClimMob (https://climmob.net).

Another innovation behind the tricot approach is the use of the Plackett-Luce model (Luce, 1959; Plackett, 1975). This statistical model is also not new, but an appropriate software implementation was not available. In previous analyses of on-farm data, data were converted to pairwise comparisons, after which the Bradley-Terry model was used (Coe, 2002; van Etten et al., 2019; Steinke et al., 2019; Dittrich et al., 2000). However, this leads to anti-conservative statistical error estimates and the conversion from rankings to pairwise comparisons implies information loss. This was the reason to implement the Plackett-Luce model in R (Turner et al., 2020). Also, a number of other R packages were created to support data management and analysis. These are described in section 12 below.

The approach is supported by the ClimMob digital platform (https://ClimMob.net). The platform will be described in detail in a forthcoming paper (Quirós et al., forthcoming). It supports the user in designing a trial, randomizing the entries, creating electronic questionnaires, collecting the data, monitoring trial progress, and generating reports.

There are several other elements that support the users. The different steps of the tricot approach are described in a manual (van Etten et al., 2020). Also, there are online guides and videos available from https://ClimMob.net.

# 3 TRIAL DIMENSIONS

A recurrent issue is deciding about trial dimensions: plot size and the number of plots (replications). In the literature about crop trials, there are several methods to guide decision making. At the same time, this is not just a question of statistics, but also biological considerations are important. For tricot trials, there are a number of additional considerations related to farmers' capacities.

## 3.1 Plot size

Having larger plots and more replications reduces the variance within entries and increases the accuracy of the value estimates for each entry, while it increases the costs. The tricot approach is driven by external validity, the ability to replicate the findings in target environments. For external validity of on-farm trials, it is important to represent the diverse growing conditions in the target environment in the trial as well as the gender and social heterogeneity of end users. This can be done best by having a large number of farms, and to generate data about the conditions of the use environment and the users that can be entered into the analysis as covariates. Small plots can be easily accommodated on both small and large farms, avoiding a bias towards the latter. Also, small plots help to reduce the quantities of planting materials that are needed for the trial, which is often a limiting factor.

In training courses on tricot, we have repeatedly noticed that agricultural scientists have strong views about trial plot size. Course participants argued that results from small plots are not reproducible on large plots, which tend to give lower yields in all cases, and concluded that large plots are needed for on-farm work. This perception is perhaps at least partly due to the *regression fallacy*, the mistaken expectation that selecting entries with a high mean performance in a trial will reproduce the same mean performance in a subsequent season. In reality a 'regression to the mean' effect is what is to be expected, as not all the entries with a high yield in the first trial are truly superior; an important part of the differences in yield between entries are due to non-genetic factors (error) (Galton, 1886). So systematically lower yields on larger plots is not a valid concern.

A valid concern is that plot size biases performance results via neighbor effects, and edge or border effects. When the plots are so small that the borders become more important, the results can be biased by the competition between varieties or the resource advantage (or sometimes disadvantage) of border rows. Much depends on the differences in competitive ability between the varieties and their ability to take advantage of the extra resources and light on the border. Rebetzke et al. (2014) provide a review on plot size, focusing mainly on wheat in Australia. Differences in height between cultivars caused neighbor effects. Plots with four rows of wheat showed a 10% bias in one example, reversing the ranking of varieties. They discuss how under drought more competitive varieties outyielded more drought-tolerant varieties. They recommend plots with at least 6 rows for accurate yield assessment in the case of wheat. Omitting border rows can further help to reduce biases arising from border effects. Border effects due to plant height are less accentuated at lower latitudes as shadowing is less important. Also, competition is less important in areas where resources for water or nutrients are not the main yield limitation, but heat or cold stress, or pests and diseases. Omitting border rows is often a standard practice in on-farm trials.

Biological understanding of differences between varieties is needed to make judicious decisions about plot size. This may generally be based on the experience of breeders and the literature. For example, for potato competition effects between plots consisting of single ridges seems unimportant for yield, as stolons rarely extend beyond ridges (Connolly et al., 1993). For cassava, interplot competition effects have been found to extend beyond the first row (Elias et al., 2018). For sweetpotato, interplot interaction is thought to be substantial due to above-ground competition (Grüneberg et al., 2019). For sweetpotato trials, 30 $m^2$ plots with 100 vine cuttings have been recommended for on-farm trials (Grüneberg et al., 2019). For a tricot trial in Ghana, however, much smaller 6 $m^2$ plots with 20 vine cuttings were used, which is two thirds of the recommended plot size of preliminary (on-station) trials.

Statistically, the neighboring effect can be partially dealt with by considering the ranking order effect (the middle position in each incomplete block has two neighbors, whereas the first and third position have one neighbor only). The order effect is not yet available in the PlackettLuce R package yet (Turner et al., 2020). This enhancement is planned for 2021.

Plot size is closely related to the number of seeds that is provided to farmers. In grain crops, breeders often provide seeds based on the average weight needed for a unit of land. However, this can be problematic when there are seed size differences between varieties. As was pointed out to us by bean breeder Juan Carlos Hernandez (INTA Costa Rica, personal communication), this means that a small-seeded variety would be represented by more seeds. Consequently, the farmer may decide to increase the plant density of a small-seeded variety or add more planting positions. This could bias yield estimates to favor small-seeded varieties. It is therefore recommendable not to provide the same weight of seed for each variety, but the same number of seeds.

For clonally propagated crops, weight biases may be less of a concern as usually seed quantities are determined in terms of the number of units (cuttings, seedlings, etc.), rather than weight. However, clonal crops have another set of issues, especially related to the perishability of planting materials, which are often also bulky. In our experience, it is important to account for possible losses of planting materials during transport and distribution.

Food products provide another set of constraints. For a tricot evaluation in which processing and culinary aspects are evaluated at the same time as agronomic aspects, the plot size would also need to be sufficient to produce the minimum quantity of product necessary for food processing. For example, cassava is elaborated into many food products using batch processing techniques that require fair amounts of product (e.g., 50 kg of product in Nigeria). The minimal quantity needed that can be processed by participants using local customs, expertise and processing equipment should be considered to determine the plot size (Teeken et al., 2020).

## 3.2 Number of blocks

Another decision that needs to be taken is the number of incomplete blocks, which is equal to the number of farmers in tricot for on-farm evaluation. The numbers that are needed depend on farmers' accuracy in observing differences between varieties, as well as the expected size of the differences. Often, the numbers to do power calculations are lacking as no previous trials have been done. (Steinke et al., 2017) estimated the accuracy of farmers for a bean trial in Central America and provided some calculations to guide trial size decisions. The results suggest that for a trial with around 12 entries (varieties, lines, etc.) typically 100-200 farms would provide solid results to make recommendations. This is the same order of magnitude that was found in previous on-farm trial work with cereals (Atlin, personal communication, 2020). If the trial covers more agro-ecological environments (to which the set of varieties is expected to respond in different ways), the number of farmers should be proportionally higher. Future studies should provide better guidance regarding optimal trial dimensions.

# 4 MAKING TRICOT INCLUSIVE: GENDER AND SOCIAL HETEROGENEITY

External validity of tricot trials has an important social science aspect. As has been indicated above, tricot trials imply sampling a representative range of use contexts, which are characterized not only by environmental variation, but also by gender and social heterogeneity, which will have an effect on variety preferences through various proximate causal factors. Firstly, crop management tends to reflect cultural and socio-economic conditions and identities (Adekambi et al., 2020). For example, the ability to purchase fertilizers or spend sufficient labor on weeding will influence how the trial plots are managed and will influence perceptions of variety performance. Another example is that farmers and processors might favor a particular variety because of its suitability for preparing a food product that is locally important or consumed by a particular social segment of the population. For example, farmers' orientation towards market production and household consumption can influence how they perceive traits related to marketability, cooking or taste (Adekambi et al., 2020). Thirdly, the degree to which farmers that participate in tricot trials have adequate knowledge of a different aspect of variety performance will depend on their involvement in different agronomic, processing and culinary activities (Teeken et al., 2020).

Gender is important in all three of these aspects (Weltzien et al., 2019). There may be differences in socio-economic status between men and women, as well as gender-based labor division for crop-related tasks. In the past, many trials have therefore addressed issues of gender by including sex of the participant as an important covariate. However, so far no tricot data analyses have shown that there are statistically detectable differences between men and women. This contrasts with the finding that trait prioritization exercises often end up with different traits mentioned by men and women, reflecting their tasks and final use of the product (Weltzien et al., 2019). This contrast may have different explanations.

First of all, tricot data and analysis did not include other social identities that can strongly intersect with gender or gender-related constraints on access to resources, knowledge and opportunities. Statistical interactions between these other social variables and gender could be revealed in aggregate datasets. This will only be possible when such data becomes available (see below). Gendered norms and roles do often not follow generalized stereotypes and can change over time, for example when outmigration of men leads to a feminization of agriculture (Abidin, 2004). Certain tasks are executed by both men and women. Gender and social heterogeneity in study areas may lead to aggregate tricot results in which general variation overwhelms any differences between men and women.

On the other hand, existing studies prior to tricot may have some limitations as well. Few studies ask participants to rank the importance of traits directly (Weltzien et al., 2019). Most studies rely on free-listing exercises, in which participants mention all the traits that occur to them. Free-listing has methodological limitations if it is used as a comparative approach. If free-listing is done in focus group discussions, they may be influenced by leadership effects (which make more senior members more influential in the results) (Richards, 2005). Also, free-listing exercises measure perceptual saliency and importance in local discourse, which may not always translate to relative importance in a realistic decision-making context in which tacit knowledge comes into play. Relative weights are often difficult to elicit through deliberation. Another possible factor is the loss of information in translation during data interpretation (for example, overzealous lumping of local concepts into more general categories) and translation from local languages.

Specific elicitation exercises to put weights on traits and segment user groups have become more prevalent recently as a result of methodological simplification, providing viable alternatives to the usual approaches from economics (conjoint analysis) which were somewhat burdensome (Byrne et al., 2012; Steinke and van Etten, 2017). This could provide important opportunities to avoid the limitations of free-listing. These new approaches use pairwise comparisons and are therefore methodologically very similar to the ranking approach used in the tricot approach. Our comments on the specific limitations of free-listing should not be interpreted as a diatribe against free-listing *per se* or qualitative methods in general, just as a caveat against the possible overinterpretation of qualitative results in uncontrolled and unrepeated comparisons. We advocate for judicious combinations of different qualitative and quantitative methods.

Sex disaggregation used in isolation will tend to overlook other issues that may correlate but also intersect with gender, such as income, occupation, marital status, ethnicity, age, or social status. Sex disaggregation alone as a basis for gender analysis will therefore not capture the high heterogeneity within the two resulting segments and give limited insights in causal relationships. This means we need to move to more subtle approaches that address intersectionality. This will require innovating on methods of analysis to analyze social differences and how they come to bear on trait and varietal choices. Innovation in two directions is ongoing.

The first innovation direction involves the use of RHoMIS (Hammond et al., 2017; van Wijk et al., 2020). This is a standardized household survey method that includes questions about the gendered execution and control of activities and control over the income derived from them. Also, the survey covers questions about household composition, farming system, nutrition, poverty and other indicators. For tricot, a selection of questions and indicators has been made to reduce the length of the questionnaire to the bare minimum to reduce respondent fatigue. The resulting data will be used to analyze the farmer-generated tricot data to determine how gender and socio-economic factors affect trial management, variety performance, and farmer variety preferences. A publication of this “layering” of RHoMIS onto tricot trial data for cassava is forthcoming. The promise of RHoMIS is that it could combine with tricot to a standardized approach that will enable comparisons across studies in variety evaluation. This does not preclude that the precise RHoMIS format as applied in combination with tricot may still need further methodological evolution.

The second innovation direction is to get a better grip on participant recruitment. Again, often fairly simplistic methods are used to address social/gender inclusion, generally quota recruitment to arrive at balanced numbers of men and women as participants. This was done in tricot trials in India, for example (van Etten et al., 2019). In a way, this puts a small set of variables upfront as explanatory factors, ignoring the importance of intersectionality or the possibility that non-identified variables may be more important to differentiate locally important social segments. For example, differences between people who are long-term residents and recent immigrants in the village may be more important than overall gender differences and can constitute important gender differences, for example, where women immigrants are in a very different position than autochthonous women (Forsythe et al., 2016). This would be impossible to capture through sex-based quota sampling, which may miss out migrants entirely. Also, during recruitment, there may be a bias towards more outspoken, talkative individuals who may not always have the best observation and judgement skills for variety evaluation. Random recruitment from the membership base of collaborating organizations has been used. This can suffice if the resulting participants represent the target population and a widely grown crop is targeted, but often local social segments remain invisible and can therefore be under or over-represented Also, in the analysis a reweighting can be done if recruitment is not representative, however excluding participants reduces statistical power and increases the relative costs of studies. However, for RTB crops generally the volume of planting materials is an important limitation. Also, not all farmers may grow relevant quantities of the target crop. Both these cases call for a better-informed sampling strategy.

IITA has implemented a purposive sampling strategy with a gender dimension for cassava trials. This sampling strategy starts with qualitative work in communities to define locally relevant social groups. Participants are then selected making sure that each local social group, and gender within them, in which cassava growing and processing expertise is present is proportionally represented. To achieve this potential participants (cassava farmer/processors) in each group are randomly interviewed and evaluated on thorough experience in cassava farming and processing (using enumerators equally having experience in this domain to assure a good check) to also capture feedback from processors that are important additional stakeholders in addition to farmers and are often also marketers and very much informed by market demand and related traits (see determination of stakeholder/value chain actors section below). This approach then makes it possible to perform a better-informed gender analysis by comparing men and women's preferences with regard to the same expertise and across different relevant social identities. This is even better facilitated as all participants are interviewed using a RHoMIS questionnaire assuring the availability of standard demographic information next to the locally determined social grouping based on the qualitative research in the communities. This approach therefore focuses on the participation of task groups/segments (Maat, 2018; Richards, 2000; McFeat, 1974). These are segments/groups that are organized around a task (for example, processing cassava into gari) and are internally relatively homogeneous in their work culture (but groups doing the same task may have other differences between them). Task groups develop a focused skilled practice which tends to generate shared language and thinking. Tapping into the expertise of these task groups is therefore an appropriate way to organize participation in order to assure that each participant is skilled and experienced which is an important condition if we want to know about crop related user preferences. It mobilizes participants around a skill set and professional identity in which they tend to take pride. A focus on task groups may also help to avoid micropolitical considerations, make the process transparent, and be more inclusive to less outspoken professionals. Task groups can be identified by tracing who does what task in the crop value chain trajectory from seed to stomach. This is done by considering local identities, including gender, but also other potential factors (e.g., age). If gender is the overriding factor in the constitution of task groups, it would accentuate the need for a nuanced gender analysis that takes into account intersectional identities beyond only sex-disaggregation of data. By using ethnographic observation methods (interviews, transect walks, market visits, etc.) the information to identify these groups can be gathered. IITA has prepared a draft guide to implement this approach (Teeken et al., in preparation).

# 5 TRICOT VARIETY EVALUATION BEYOND FARMERS

Variety evaluation not only affects farmers, but also other stakeholders, including processors, traders, and consumers. Tricot has already been used for sweetpotato product testing in Ghana and Uganda (Moyo et al., 2021). In both countries, tests were done at participants' homes as well as at central locations: markets in Ghana, and community meetings in Uganda. With tricot, it was possible to generate a large amount of data in a relatively smooth way and in little time through the centralized testing. The home-based testing entailed more difficulty in data retrieval. Overall agreement between centralized testing and home testing was high in Uganda, but moderate in Ghana, even though the top varieties coincided between the two types of testing in both countries. Performing tricot trials in markets in Ghana allowed for better statistical discrimination between varieties whereas home testing did worse. In Uganda, a tighter protocol for home testing was followed, leading to similar information between the two types of testing. If between-household variation in preparation methods is not strongly affecting consumer test results, the centralized testing will be more effective and more easily realized. Centralized testing will need at least a basic characterization of participants to allow for segmentation and reweighting to be able to generalize findings.

Doing farmer and consumer testing separately will be appropriate in many contexts, as farmers are only one subset of consumers. Separate consumer testing comes with the methodological challenge of bringing together different aspects into a single assessment of variety performance, supported by a single variety performance score. Different approaches are possible to generate a single score from the farmers' and consumers' scores. We briefly consider options to illustrate the type of decisions involved in Box 1.

**Box 1**. Mathematically combining farmer and consumer variety scores

Imagine we have two scores for a set of variety *i*, a farmer score (*a*) and a consumer score (*b*), in the form of the probability of beating all other varieties (Turner et al., 2020; van Etten et al., 2019). To convert these two scores per variety into a single score (*S*), we have to select a method. Some possibilities are discussed in what follows.

*Average of scores*. $S_i = (a_i+b_i)/2$.

This gives equal weight to the two groups of stakeholders. It is not clear which weights would be appropriate if a weighted average were used. A low farmers' score can be compensated by a high consumers' score. This is not a desirable characteristic of this score. It can lead to relatively high scores for varieties that are not acceptable to one group of stakeholders.

*Minimum score*. $S_i = \min(a_i, b_i)$.

This is a simple minimum threshold for each of the stakeholders, in which no compensation occurs. It may be a bit extreme in the sense that it may rule out varieties that do compensate somewhat. For example, farmers may be willing to cultivate a variety with a suboptimal agronomic performance if they know it has very good consumer acceptance.

*Worst regret*. $S_i = \max(a_{max} - a_i, b_{max} - b_i)$.

The largest difference with the respective best score. This is different from the previous measures in that the score contains some information about the spread of the scores, because it depends on the respective best scores for each stakeholder group. This may be interesting if there are strong differences between farmers and consumers in the *spread* of the variety performance values. This happens, for example, if farmers give very similar scores to all varieties but consumers give some varieties very low and other varieties very high scores. Otherwise, it is similar to the minimum score method. Regret scores are dependent on the value of the highest score of each stakeholder group, however, which can be a weakness as a single variety may have an extreme value. To increase the stability of the method, the distance to the average of the top-*k* varieties could be calculated instead. (This also applies to the next measure.)

*Minimum relative regret*. $S_i = \text{sqrt}( ((a_{max} - a_i)/a_{max})^2 + ((b_{max} - b_i)/b_{max})^2 )$

This is the most complex method but it overcomes some of the shortcomings of the previous scores. It involves rescaling all the variety scores by dividing by the highest score, separately for farmers and consumers. Then the Euclidean distance from each variety to the (1,1) point is calculated. It allows for some compensation between stakeholders (but the trade-off is not linear). Since the regret score is rescaled, it deals to some extent with the differences in spread.

None of the presented methods is perfect and picking a method involves balancing different stakeholder interests. Minimum relative regret is a promising compromise as a heuristic for decision making. It could be supplemented by a visualization of the data to facilitate discussions among decision makers.

# 6 ETHICS, PRIVACY AND RIGHTS ON TRADITIONAL KNOWLEDGE

Tricot involves human subjects and must therefore observe certain research ethics standards. In general terms, the application of tricot must minimize the possible risks, discomfort, nuisances and costs for participants while maximizing the benefits that they and other farmers may obtain (directly or indirectly) from the trial data obtained through tricot.

Tricot is also subject to privacy issues, and data management needs to conform to General Data Protection Regulation as the Alliance of Bioversity International and CIAT is headquartered in the European Union (Italy).

In general, this will mean the following for tricot trials:

- Research ethics clearing is obtained from the relevant Institutional Review Board (IRB).
- Research ethics clearance may be also necessary from a national organization. For this purpose, Tricot users must take national laws and guidelines into account.
- Prior informed consent is obtained from all participants, which would allow for data publication after anonymization.
- Participants are given the right to withdraw from the study while it is executed.
- Participants are given the right to withdraw their data from the study while it is in the course of being executed.
- Participants can indicate if they want to be recognized with their name in the publications based on the data. This does not compromise privacy (names cannot be linked to personal identifiable information such as addresses, telephone numbers or coordinates).

In practice, this means the following for the further development of the tricot approach and the ClimMob platform:

- ClimMob should provide features to make it easy for trial designers to follow the principles and procedures indicated above:
  - Automatically generated document to request IRB clearance;
  - Standardized, short prior informed consent forms and practical ways to implement paper-based signature + photograph of the document, electronic signature, or spoken approval (audio);
  - Names of participants that want to be named in the research publication exported by the platform.
  - Anonymization of data before exporting. This can be automatized through automatic detection of potential personal identifiable information (see https://dataverse.scio.systems:9443/).
- Throughout the design of an experiment, ClimMob should provide cues to prompt users to consider research ethics, privacy and traditional knowledge rights in the design of tricot trials;
- ClimMob needs to be GDPR-compliant to users (cookie policy, explicit notice about usage of data). The version available at the moment of writing already has this implemented.

A more complex topic that deserves separate discussion is that tricot may be affected by national laws on the *access to genetic resources and associated traditional knowledge and the sharing of benefits arising from their use* (ABS, for short). There are two aspects in which tricot is affected by ABS, via the use of traditional varieties and via the use of traditional knowledge held by participants. We consider both aspects.

Firstly, tricot may need to observe ABS rules when using traditional varieties. Tricot is usually applied to test the performance of new, improved varieties. However, in some cases, genetic materials of traditional varieties are included in trials as ‘check varieties’, to make comparisons. Although the utilization of the check varieties does not fall within the activities that are usually subject to ABS requirements in most countries, whether or not ABS obligations apply will depend on the definition of utilization adopted by the country of provenance of the variety (i.e., the country where the research is implemented). Therefore, tricot users will need to analyze the applicable access rules in the country where they are operating, obtain the access permits and negotiate mutually agreed terms when necessary. If the country where the traditional varieties come from is a party to the International Treaty on Plant Genetic Resources for Food and Agriculture (Plant Treaty), the acquisition of the traditional varieties for their use in tricot may be subject to the terms and conditions of the Plant Treaty’s multilateral system of access and benefit-sharing. In this case, access to the samples would be facilitated by the Standard Material Transfer Agreement. Since the purpose is not to breed the traditional varieties or incorporate them in new, improved lines, the multilateral system’s mandatory monetary benefit-sharing conditions would not apply, and thus the tricot users would not have any benefit-sharing obligation. However, they would have the obligation to transfer the varieties they have obtained with the SMTA under the same terms and conditions as those of the multilateral system, whenever the recipients of such material are going to use it for conservation, research, training and breeding.

Secondly, tricot may be exposed to ABS laws when using traditional knowledge. Farmers’ ability to perceive crop characteristics is often considered to be part of traditional knowledge related to genetic resources (Mancini et al., 2017). In tricot trials, farmers use their skills to produce new knowledge, which would usually not fall under national ABS laws, but whose use may be anyway subject to rules and protocols related to the interaction with indigenous and local communities, the access to their knowledge and their natural resources. Even if the country has not yet enacted ABS legislation in relation to genetic resources and/or traditional knowledge, or even if the existing laws and regulation do not apply to tricot trials in a particular context, it is wise to observe, the CBD and the Nagoya Protocol principles in the management of farmers’ varieties and knowledge in tricot trials, as ‘best practice’, as recommended by the Guidelines on the Nagoya Protocol for CGIAR Research Centers. This means, among other things, sharing non-monetary benefits back with the participants, in the form of informational results, best performing varieties and other types of technologies.

# 7 GENETIC GAIN: ON-FARM YIELD ESTIMATION

One of the goals of on-farm testing is to get insights into genetic gain achieved by breeding programmes. Some aspects of genetic gain are related to traits that are highly heritable so that on-farm performance is not different from on-station performance. For example, the color of the product may not be affected by genotype by environment interactions. An aspect of genetic gain that is important as a goal shared by most breeding programmes is the yield. As tricot is based mainly on rankings, generally yield estimations have been provided in that form. This provides an insight into the yield-based *reliability*, the probability that a new variety will outperform the current market leader, an important indicator for breeders and product managers to make decisions (Eskridge and Mumm, 1992). The CGIAR Excellence in Breeding strategy focuses on product profiles that emphasize cumulative gains towards product replacement, taking over market share from existing varieties (Cobb et al., 2019). Tricot is well suited to address the challenge of providing early indicators of the probability that product replacement happens.

In many cases, however, breeders need to have absolute estimates of yield levels, for example because this is a requirement for a variety release procedure. In one case, a subset of the fields has been visited to obtain yield estimates (NextGen Cassava), in other cases, all fields were visited for yield measurements (de Sousa et al., 2020). This 'undermines' the tricot approach to some degree in the sense that the field visits become an important cost driver. This leads to the question whether farmers themselves can provide reliable yield estimates.

Ochieng, Ojime, and Otieno (2019) have addressed this question by comparing yield estimates by researchers (taking into account grain moisture) and by farmers (volumetric, using 250 ml tins). They set up an experiment with common bean (*P. vulgaris*) in Kenya. They obtained a high correlation between the two types of measurements when all seasons and locations were aggregated (r = 0.98). When differences were smaller than 0.5 t/ha, the match between values provided by farmers and researchers decreased. We aim to replicate these studies in other contexts with other crops in order to get a better grip on the accuracy of farmers' measurements and to use these accuracy estimates in statistical analyses. On the other hand, these studies will provide insights in how to maximize farmers' accuracy.

It would be ideal to be able to combine yield ranking data and yield measurement data when the measurement data is only available for a part of the trial. It is possible to feed absolute measurements and ordinal (ranking) data into the same statistical model, directly (Böckenholt 2004) or through a Bayesian approach. This has not been implemented in software yet; this is a pending task.

# 8 DOCUMENTING TRIALS AND PUBLISHING DATA

Open access publication of the data should be a goal of the trial. Tricot has already published a number of sizable datasets from on-farm trials (van Etten et al., 2018; Moyo et al., 2020; de Sousa et al., 2020). These datasets could become important for other research that repurposes these datasets (see section 11 below). Kool et al. (2020) have provided an incisive critique of on-farm testing in agronomy, especially the limited replicability of many trials as authors fail to report contextual factors (crop management) and sampling of locations and participating farmers. Similarly, a study on PVS in RTB crops reveals that on-farm trials are often documented in a very deficient way and that data are hardly published at all (Jose Valle et al., forthcoming). Data publication could become more attractive if it is easy to do and has rewards (citations of datasets repurposed by others). Publishing all data from trials could prevent the so-called file-drawer problem, which means that only certain datasets (for example, novel analyses, striking results) are published, which then lead to biased statistics in meta-analyses.

The tricot approach should address this issue by facilitating and standardizing the way in which on-farm trials are documented and published. Standardization should be done using the insights of the studies cited above. Specifically, meta-data on the trials could be standardized and some elements on the trial context could become recommended elements that are easily available from within the software. For example, it is becoming more and more clear that plot use histories and fertilization in preceding seasons of plots are highly influential on yields (Njoroge et al., 2019; Zingore et al., 2007). For this, an existing metadata schema for phenotypic experiments could be adapted (Papoutsoglou et al., 2020). Also, the data publication process should be automatized, including the anonymization procedure (removing personal identifiable information such as names, addresses and telephone numbers as well as aggregating geographic data to a sufficient level to prevent identification).

# 9 DIGITAL DEVELOPMENT - RECENT PROGRESS

The tricot experimental cycle is supported in all its steps by a digital platform called ClimMob (Figure 1). A paper describing ClimMob is nearly finished and will be submitted in early 2021. ClimMob is a digital platform built with existing open software, including Open Data Kit, ODK Tools, Formshare, OData. The main platform is implemented in Python, a language which is appropriate for both web design and data handling. Data analysis is mostly done in R (R Core Team, 2017). The R code makes uses of several R packages, including ClimMobTools to do the randomization (see Table 1), rmarkdown to generate the reports (Allaire et al., 2020) and Plackett-Luce to analyze the data (Turner et al., 2020).

The platform was built based on much user testing leading to iterative improvements. It was built using professional software development principles, is highly modular and can be easily be further improved and expanded.

An overview of the functionality of ClimMob is provided in Figure 1. The project is designed on the platform as a first step. This includes the design of electronic questionnaires for participant registration and to record their observations, identifying field agents, the varieties to be tested. ClimMob will generate a randomization of the packages. The experimental design is balanced so that all the varieties are represented with the same frequency. Balancing is sequential, so that by handing out packages in consecutive order, also balance in different villages will be ensured.


**Figure 1**. ClimMob overview: steps, digital outputs and underlying open-source software

Some recently added features include:

- A 'What You See Is What You Get' (WYSIWYG) module to design the questionnaires, which visualize a preview of the format as it will appear in ODK Collect on Android screens (Figure 2).
- The possibility to register participants and immediately evaluate the tested items. This facilitates food product testing.
- QR codes connect the Android phones to the right project and to scan test packages when they are distributed to farmers. This makes these procedures easier and less error-prone.
- Detailed progress monitoring of data collection is possible through the general dashboard and specific reports that are generated by the platform.
- Easy correction of conflicting data entry is possible on the ClimMob platform.
- A complete revamped reporting module that generates automatic reports based on analyses in R (R Core Team, 2017). The report includes a description of the trial, a map of the farm locations, as well as detailed analyses of all traits.
- Most interactions with ClimMob can also be done by machines through APIs. These APIs have been documented so that other software developers can develop complementary applications that retrieve data from ClimMob and send data and instructions to ClimMob.

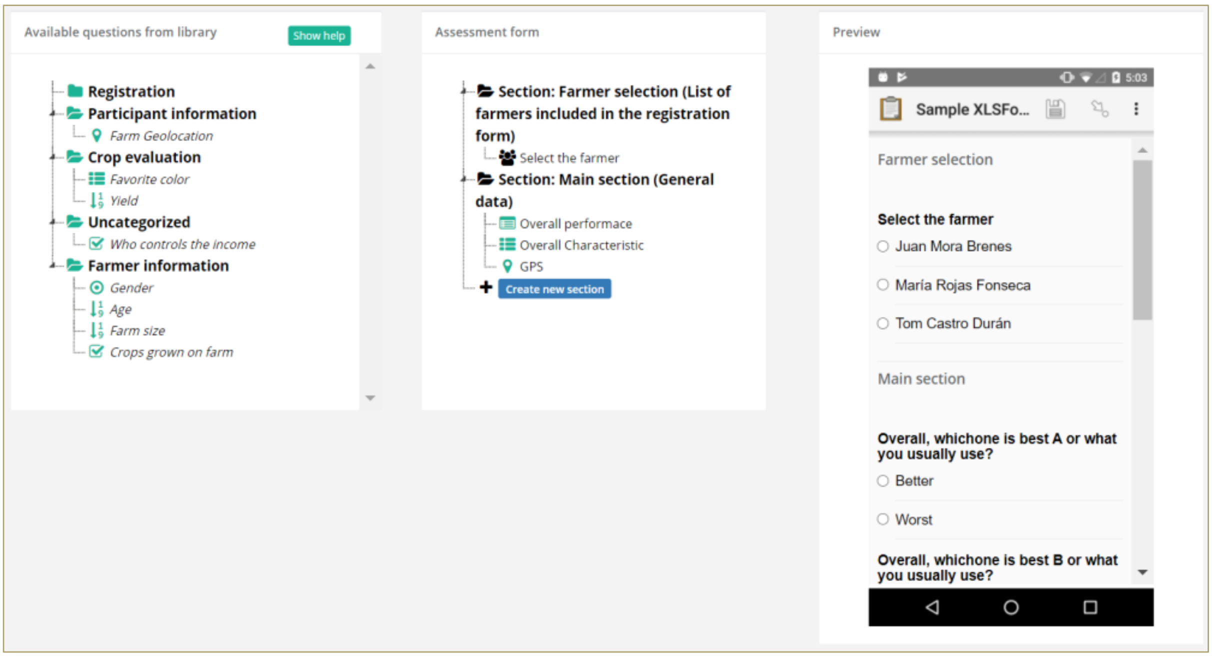


**Figure 2**. To generate questionnaires in ClimMob a preview is available directly.

# 10 DIGITAL DEVELOPMENT - FURTHER NEEDS

In the coming period, work on ClimMob will concentrate on the following issues.

- *Multi-language implementation*. The interface is currently available in English. The interface can be made available in other languages - the software has this already implemented but translations need to be prepared. Not only the interface, but also the library with questions (traits/variables) will be structured to be multilingual. It will be possible to add new versions of questions that already exist in other languages and conceptually link them. This will make it possible to have questionnaires in multiple local languages and will facilitate data analysis by scientists who may not know the local language or deal with data collected across different areas.
- *Alternative communication channels*. Possibilities to collect data via WhatsApp, Interactive Voice Response or SMS are planned or ongoing. This will make data collection more versatile. Currently data can only be entered via ODK Collect, a robust Android app, and directly online.

These features involve advanced digital development but no new scientific conceptualization. *Data standardization and ontologies* and *data analysis* also involve concomitant digital development work, but these aspects are discussed separately in sections 11 and 12, respectively.

# 11 DATA STANDARDIZATION AND ONTOLOGIES

ClimMob will connect with relevant ontologies for agriculture, thus securing its compliance with the Breeding API (BrAPI), for enabling the interoperability of the tricot data with breeding data. This way, ClimMob will extract defined traits and variables for the creation of project-specific questionnaires and storage in the database.

The Crop Ontology used by breeding databases provides descriptions, URIs (unique identifiers) and relationships of agronomic, morphological, physiological, quality, and stress traits. It follows a conceptual model that defines a phenotypic variable as a combination of a trait, a method and a scale (Shrestha et al, 2012, Arnaud et al, 2020) (Shrestha et al., 2012; Arnaud et al., 2020). Therefore, the tricot ranking method needs to be included into the Crop Ontology for all traits that are relevant to the tricot trials. CO contains today 4,456 traits and 6,292 variables with methods and scales for 31 plant species. The Agronomy Ontology provides descriptions of field management practices (Devare et al, 2016). The conceptual model is centered on the plot or the entire field. It describes planned and unplanned time-bound processes occurring in the plot (e.g., fertilizer application), along with 'participants' to the event that can be a tool, a chemical component (e.g., manure spray, limestone).

To support the connection of breeding product profiles to multiple sources of trait information, the ontology work is being extended to traits described by the social groups or market segments (e.g., sensory traits linked to the food products qualities, food product processing techniques for local processors) and will be completed by an ontology of the social groups and their roles in the value chain. The newly created socio-economic ontology (SEONT; Arnaud et al, 2020) based on the mini version of RHoMIS will support the use of socio-economic data for ClimMob projects.

Ontologies can also support the management of multilingual trait lists by mapping the concepts across languages. The agricultural thesaurus called AGROVOC, maintained by FAO, contains 38,000 concepts in around 40 languages and will be an important resource for concept translation.

A closely related effort is to create consensus on the variety traits and the socio-economic variables that should be included in tricot trials. A high degree of consensus about the traits would benefit the combined analysis of different datasets (see section *Data analysis*), ensure that important traits or variables are not omitted, and reduce the time spent on debating the different options for the design of each project while permitting the flexibility to add traits and variables that are thought to be of importance to a particular trial.

Trait lists have been developed for a number of crops but not yet published. Table 1 gives an example for cassava varietal traits to be elicited at harvest. These trait lists and questions have been generated through iterated discussion between domain experts. These traits will be available in ClimMob for each crop. Also, drawings were made to illustrate each trait, which are used to develop printed materials for farmers. Figure 3 shows an example.

**Table 1.** Example of standardized list of traits: cassava at harvest

| Characteristic | Questions |
|---|---|
| Planting material | Which variety has the highest number of stems for planting?<br>Which variety has the lowest number of stems for planting? |
| Maturity | Which variety bulks earliest?<br>Which variety bulks latest? |
| Root yield | Which variety gives the highest yield?<br>Which variety gives the lowest yield? |
| Root size | Which variety has the biggest roots?<br>Which variety has the smallest roots? |
| Root shape | Which variety has the best shaped roots?<br>Which variety has the worst shaped roots? |
| CBSD root necrosis | Which variety has the least rotten roots?<br>Which variety has the most rotten roots? |

**Figure 3.** Drawings to illustrate the root shape trait for cassava. On the left, regularly shaped roots and on the right, irregularly shaped roots.

# 12 DATA ANALYSIS

Tricot data consist of rankings, an unusual data type in the agricultural sciences, in spite of some experience with it in participatory research (Coe, 2002). As indicated in section 2, the tricot approach relies much on the Plackett-Luce model. The Plackett-Luce model also allows for the inclusion of covariates using recursive partitioning, which uses binary splits (Strobl, Wickelmaier and Zeileis, 2011). At the moment, the Plackett-Luce package is being expanded to include Plackett-Luce regression, which uses linear covariates (Yildiz et al., 2020).

Van Etten et al., (2019) showed how the Plackett-Luce model can be used in combination with seasonal climate data and cross-validations to produce robust, locally-specific variety recommendations. Tricot data analysis has recently expanded into two directions.

Firstly, Brown et al., (2020) have described how the Plackett-Luce model can be used to synthesize trial data from across different trials. Ongoing work is doing this with tricot trials (and other trials, after converting absolute values to ranks to deal with highly heterogeneous data). This is a promising new direction, as it will show in the future that working with a standardized approach like tricot and sharing data openly has strong benefits for science in general and generates recognition for individual scientists who decide to publish their data. Also, it may stimulate data publication from trials that are not worth a peer-reviewed journal article on their own but gain value after being combined with data from other trials.

In a forthcoming paper, de Sousa et al. (2020) take the Plackett-Luce model in another direction, by adding genomic relatedness data to the model (as a covariance matrix). This increased the predictivity of the model in an important measure, showing that it may be feasible and relevant to use relatedness data to allow for more diverse sets of materials to be tested by farmers. This may require important changes in breeding approaches but it provides an interesting prospect.

The new Plackett-Luce regression approach will also allow the use of variety traits as covariates (Yildiz et al., 2020). This opens interesting possibilities to analyze the relative influence of known trait values on on-farm performance. It will also possibly provide avenues to link tricot results with trait prioritization exercises, discussed in section 4 above.

Data analysis has also been increasingly supported by implementing the existing code, which was generated to a large degree for the analyses presented in (van Etten et al., 2019), into R packages. The R packages that have been generated as a result from this research are listed in Table 2.

**Table 2.** R packages created to support the tricot approach

| Package | Function | Reference, URL |
| --- | --- | --- |
| PlackettLuce | Plackett-Luce model to analyze ranking data. | Turner et al. (2020)<br>https://cran.r-project.org/package=PlackettLuce |
| ClimMobTools | API to download data from ClimMob easily into R and start processing it. | https://cran.r-project.org/package=ClimMobTools |
| gosset | The gosset package provides the toolkit and a workflow to analyze metadata and experimental citizen science data, from collating data of different sources to model selection and visualization. | https://agrobioinfoservices.github.io/gosset/ |
| chirps | Obtain daily rainfall data in the public domain for any location in the tropics easily into R. | de Sousa et al. (2020)<br>https://cran.r-project.org/package=chirps |
| climatrends | Calculate agrometeorological indicators from weather data to use them in trial analysis. | https://cran.r-project.org/package=climatrends |

# 13 SCALING: COST REDUCTION AND NEW BUSINESS MODELS

An important benefit of tricot is the possible reduction of trial costs, which could drive its adoption across different organizations, apart from the improved insights that it produces. Precise cost comparisons are difficult as there is no gold standard to compare with. Both 'conventional' participatory variety selection and tricot can be implemented in different ways. If both are implemented in a very intensive way (organizing farmer groups, meetings, working with RTB seed materials), the cost reduction is estimated to be roughly 40%. At the other extreme, a tenfold cost reduction is possible in the US, where farmers receive seeds by mail and are connected by smartphones. One reason for cost reduction is that some costs are externalized to farmers who volunteer to execute mini-trials using their own labor, land and inputs. This raises the question whether farmers' motivation to participate is sufficiently enhanced by tricot to justify this investment. Previous studies showed that farmers' motivation to participate in tricot is mainly related to access to seeds and information (Beza et al., 2017). Cost and motivation analyses are underway and should be available for Rwanda and Ghana in 2021.

Further cost reductions are possible if farmer networks are maintained over time, if they are serviced through channels that are also used for other means (credit provision, for example), and if they can reach economies of scale by testing varieties and other options for multiple crops. Tricot would make it possible for breeders and agronomists to 'outsource' trials to farmer-facing organizations. Alternative business models have already been introduced in the US context by organizations such as the Farmer Business Network, FIRST (Farmers' Independent Research of Seed Technologies), and SeedLinked. The latter uses the tricot approach for its trials. The Alliance of Bioversity International and CIAT is exploring alternative business models following this trend focusing on the global South. Research is ongoing within RTB to determine the best scaling strategies for Rwanda and Ghana.

# 14 CONCLUSIONS

It has become clear that on-farm testing is complicated due to the need to coordinate multiple moving parts: the logistics of mobilizing farmers, getting seeds on time to them, designing a trial that generates data about crops, crop management, and socio-economic conditions, the need to address ethical, privacy and rights aspects, monitoring data collection, and analyzing the data and reporting on time. This can only be done with skilled practice by all those involved.

By streamlining the whole process of on-farm testing and standardizing some aspects, tricot hopes to deliver on the promise of richer insights from on-farm trials that can be obtained if all elements are well aligned. Tricot is rapidly evolving into a mature approach supported by mature tools that should transform on-farm testing across many applications.

As tricot has gained sufficient maturity to be scaled, efforts are underway to mainstream the approach into breeding and extension programmes, with support from the CGIAR research programmes RTB and GLDC, and specific funding from the Bill and Melinda Gates Foundation. These efforts are supported by very constructive collaborations between CGIAR centers, including the Alliance of Bioversity International and CIAT, CIP, IITA, ICRISAT, and CIMMYT. Also, numerous NARES partners, NGOs, and enterprises are implementing the tricot approach. For further scaling, a consistent investment in capacity building is needed to transform on-farm testing in breeding and extension. Sustainable scaling will also require the construction of a solid business model.

# 15 REFERENCES

Abidin, P. E. 2004. “Sweetpotato Breeding for Northeastern Uganda: Farmer Varieties, Farmer-Participatory Selection, and Stability of Performance.” PhD Thesis, Wageningen University. https://library.wur.nl/WebQuery/wurpubs/fulltext/193585.

Adekambi, S. A., J. J. Okello, S. Rajendran, K. Acheremu, E. E. Carey, J. Low, and P. E. Abidin. 2020. “Effect of Varietal Attributes on the Adoption of an Orange-Fleshed Sweetpotato Variety in Upper East and Northern Ghana.” *Outlook on Agriculture* 49 (4): 311–20. https://doi.org/10.1177/0030727020950324.

Allaire, J., Yihui Xie, Jonathan McPherson, Javier Luraschi, Kevin Ushey, Aron Atkins, Hadley Wickham, Joe Cheng, Winston Chang, and Richard Iannone. 2020. “Rmarkdown: Dynamic Documents for R, R Package Version 2.6.” 2020. https://github.com/rstudio/rmarkdown.

Arnaud, Elizabeth, Marie-Angélique Laporte, Soonho Kim, Céline Aubert, Sabina Leonelli, Berta Miro, Laurel Cooper, et al., 2020. “The Ontologies Community of Practice: A CGIAR Initiative for Big Data in Agrifood Systems.” *Patterns* 1 (7): 100105. https://doi.org/10.1016/j.patter.2020.100105.

Beza, Eskender, Jonathan Steinke, Jacob van Etten, Pytrik Reidsma, Carlo Fadda, Sarika Mittra, Prem Mathur, and Lammert Kooistra. 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.

Böckenholt, Ulf. 2004. “Comparative Judgments as an Alternative to Ratings: Identifying the Scale Origin.” *Psychological Methods* 9 (4): 453.

Brown, David, Inge Van den Bergh, Sytze de Bruin, Lewis Machida, and Jacob van Etten. 2020. “Data Synthesis for Crop Variety Evaluation. A Review.” *Agronomy for Sustainable Development* 40 (4): 25. https://doi.org/10.1007/s13593-020-00630-7.

Byrne, T. J., P. R. Amer, P. F. Fennessy, P. Hansen, and B. W. Wickham. 2012. “A Preference-Based Approach to Deriving Breeding Objectives: Applied to Sheep Breeding.” *Animal: An International Journal of Animal Bioscience* 6 (5): 778–88. https://doi.org/10.1017/S1751731111002060.

Cobb, Joshua N., Roselyne U. Juma, Partha S. Biswas, Juan D. Arbelaez, Jessica Rutkoski, Gary Atlin, Tom Hagen, Michael Quinn, and Eng Hwa Ng. 2019. “Enhancing the Rate of Genetic Gain in Public-Sector Plant Breeding Programs: Lessons from the Breeder’s Equation.” *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik* 132 (3): 627–45. https://doi.org/10.1007/s00122-019-03317-0.

Coe, Richard. 2002. “Analyzing Ranking and Rating Data from Participatory on-Farm Trials.” *Quantitative Analysis of Data from Participatory Methods in Plant Breeding*, 44–65. https://books.google.com/books?hl=en&lr=&id=iIGFJr4zz-EC&oi=fnd&pg=PA44&dq=coe+participatory&ots=HdfqVZtENj&sig=NRkvCthP_mb_Nk7WEtOXSEOX8aM.

Connolly, T., I. D. Currie, J. E. Bradshaw, and J. W. McNicol. 1993. “Inter-Plot Competition in Yield Trials of Potatoes (Solanum Tuberosum L.) with Single-Drill Plots.” *The Annals of Applied Biology* 123 (2): 367–77. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-7348.1993.tb04099.x.

Dittrich, R., W. Katzenbeisser, and H. Reisinger. 2000. “The Analysis of Rank Ordered Preference Data Based on Bradley-Terry Type Models.” *OR Spectrum* 22 (1): 117–34. https://doi.org/10.1007/s002910050008.

Elias, Ani A., Ismail Rabbi, Peter Kulakow, and Jean-Luc Jannink. 2018. “Improving Genomic Prediction in Cassava Field Experiments by Accounting for Interplot Competition.” *G3* 8 (3): 933–44. https://doi.org/10.1534/g3.117.300354.

Eskridge, K. M., and R. F. Mumm. 1992. “Choosing Plant Cultivars Based on the Probability of Outperforming a Check.” *Theoretical and Applied Genetics* 84 (3-4): 494–500.

Etten, Jacob van, Eskender Beza, Lluís Calderer, Kees Van Duijvendijk, Carlo Fadda, Basazen Fantahun, Yosef Gebrehawaryat Kidane, et al., 2019. “First Experiences with a Novel Farmer Citizen Science Approach: Crowdsourcing Participatory Variety Selection through on-Farm Triadic Comparison of Technologies (tricot).” Experimental Agriculture 55 (S1): 275–96. https://doi.org/10.1017/S0014479716000739.

Etten, Jacob van, Rhys Manners, Jonathan Steinke, Elsa Matthus, and Kauê de Sousa. 2020. “The Tricot Approach. Guide for Large-Scale Participatory Experiments.” https://cgspace.cgiar.org/handle/10568/109942.

Etten, Jacob van, Kauê de Sousa, Amílcar Aguilar, Mirna Barrios, Allan Coto, Matteo Dell’Acqua, Carlo Fadda, et al., 2019. “Crop Variety Management for Climate Adaptation Supported by Citizen Science.” *Proceedings of the National Academy of Sciences of the United States of America* 116 (10): 4194–99. https://doi.org/10.1073/pnas.1813720116.

Etten, J. van. 2011. “Crowdsourcing Crop Improvement in Sub-saharan Africa: A Proposal for a Scalable and Inclusive Approach to Food Security.” *IDS Bulletin*. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1759-5436.2011.00240.x.

Etten, J. van, K. De Sousa, A. Aguilar, M. Barrios, A. Coto, M. Dell’Acqua, C. Fadda, et al., 2018. “Replication Data for: ‘Crop Variety Management for Climate Adaptation Supported by Citizen Science.’” Harvard Dataverse. https://doi.org/10.7910/DVN/4ICF6W.

Fadda, Carlo, Dejene K. Mengistu, Yosef Gebrehawaryat Kidane, Matteo Dell’Acqua, Mario Enrico Pè, and Jacob van Etten. 2020. “Integrating Conventional and Participatory Crop Improvement for Smallholder Agriculture Using the Seeds for Needs Approach: A Review.” *Frontiers in Plant Science* 11: 1421.

Forsythe, Lora, Helena Posthumus, and Adrienne Martin. 2016. “A Crop of One’s Own? Women's Experiences of Cassava Commercialization in Nigeria and Malawi.” *Journal of Gender, Agriculture and Food Security* 1 (2): 110–28. http://gala.gre.ac.uk/id/eprint/14707/.

Galton, Francis. 1886. “Regression Towards Mediocrity in Hereditary Stature.” *The Journal of the Anthropological Institute of Great Britain and Ireland* 15: 246–63. https://doi.org/10.2307/2841583.

Gevel, Jeske van de, Jacob van Etten, and Sebastian Deterding. 2020. “Citizen Science Breathes New Life into Participatory Agricultural Research. A Review.” *Agronomy for Sustainable Development* 40 (5): 35. https://doi.org/10.1007/s13593-020-00636-1.

Grüneberg, W. J., R. Eyzaguirre, F. Diaz, B. de Boeck, J. Espinoza, J. Swanckaert, H. Dapaah, et al., 2019. “Procedures for the Evaluation of Sweetpotato Trials. Manual.” International Potato Center. https://cgspace.cgiar.org/handle/10568/105875.

Hammond, James, Simon Fraval, Jacob van Etten, Jose Gabriel Suchini, Leida Mercado, Tim Pagella, Romain Frelat, 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." *Agricultural Systems* 151: 225–33. https://www.sciencedirect.com/science/article/pii/S0308521X16301172.

Kool, Hanna, Jens A. Andersson, and Ken E. Giller. 2020. "Reproducibility and External Validity of on-Farm Experimental Research in Africa." *Experimental Agriculture* 56 (4): 587–607. https://doi.org/10.1017/S0014479720000174.

Luce, R. Duncan. 1959. *Individual Choice Behavior: A Theoretical Analysis*. Wiley.

Maat, Harro. 2018. "Group Compositions: The Politics of Technology Implemented in Smallholder Farming." *Journal of Political Ecology*. https://doi.org/10.2458/v25i1.22389.

Mancini, Chiara, Yosef G. Kidane, Dejene K. Mengistu, Mario Enrico Pè, Carlo Fadda, and Matteo Dell'Acqua. 2017. "Joining Smallholder Farmers' Traditional Knowledge with Metric Traits to Select Better Varieties of Ethiopian Wheat." *Scientific Reports* 7 (1): 9120.

McFeat, Tom. 1974. *Small-Group Cultures*. Pergamon. https://play.google.com/store/books/details?id=WR-0BQAAQBAJ.

Minet, Julien, Yannick Curnel, Anne Gobin, Jean-Pierre Goffart, François Mélard, Bernard Tychon, Joost Wellens, and Pierre Defourny. 2017. "Crowdsourcing for Agricultural Applications: A Review of Uses and Opportunities for a Farmsourcing Approach." *Computers and Electronics in Agriculture* 142: 126–38.

Moyo, Mukani, Sam Namanda, Reuben Ssali, Mariam Nakitto, Eric K. Dery, Daniel Akansake, Joseph Adjebeng-Danquah, et al., 2020. "Replication Data for: Consumer Preference Testing of Boiled Sweetpotato (Ipomoea Batatas (L.) Lam.) Using Crowdsourced Citizen Science in Ghana and Uganda." Zenodo. https://doi.org/10.5281/ZENODO.4117591.

Moyo, Mukani, Reuben Ssali, Sam Namanda, Mariam Nakitto, Eric K. Dery, Daniel Akansake, Joseph Adjebeng-Danquah, et al., 2021. "Consumer Preference Testing of Boiled Sweetpotato (Ipomoea Batatas (L.) Lam.) Using Crowdsourced Citizen Science in Ghana and Uganda." *Frontiers in Sustainable Food Systems*. https://doi.org/10.3389/fsufs.2021.620363.

Njoroge, Samuel, Antonius G. T. Schut, Ken E. Giller, and Shamie Zingore. 2019. "Learning from the Soil's Memory: Tailoring of Fertilizer Application Based on Past Manure Applications Increases Fertilizer Use Efficiency and Crop Productivity on Kenyan Smallholder Farms." *European Journal of Agronomy: The Journal of the European Society for Agronomy* 105: 52–61. https://www.sciencedirect.com/science/article/pii/S1161030118302557.

Ochieng, Hannington O., John Ojiem, and Joyce Otieno. 2019. "Farmer versus Researcher Data Collection Methodologies: Understanding Variations and Associated Trade-Offs." https://doi.org/10.31730/osf.io/ncw8a.

Papoutsoglou, Evangelia A., Daniel Faria, Daniel Arend, Elizabeth Arnaud, Ioannis N. Athanasiadis, Inês Chaves, Frederik Coppens, et al., 2020. "Enabling Reusability of Plant Phenomic Datasets with MIAPPE 1.1." *The New Phytologist* 227 (1): 260–73. https://doi.org/10.1111/nph.16544.

Plackett, Robin L. 1975. "The Analysis of Permutations." *Applied Statistics*, 193–202.

R Core Team. 2017. "R: A Language and Environment for Statistical Computing." Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rebetzke, Greg J., Ralph Tony A. Fischer, Anthony F. van Herwaarden, Dave G. Bonnett, Karine Chenu, Allan R. Rattey, and Neil A. Fettell. 2014. "Plot Size Matters: Interference from Intergenotypic Competition in Plant Phenotyping Studies." *Functional Plant Biology: FPB* 41 (2): 107–18. https://doi.org/10.1071/FP13177.

Richards, Paul. 2000. *Food Security, Safe Food: Biotechnology and Sustainable Development in Anthropological Perspective*. Wageningen University. https://www.cabdirect.org/cabdirect/abstract/20036795714.

———. 2005. "How Does Participation Work? Deliberation and Performance in African Food Security." *IDS Bulletin* 38 (5): 21–35. https://doi.org/10.1111/j.1759-5436.2005.tb00406.x.

Ryan, S. F., N. L. Adamson, A. Aktipis, L. K. Andersen, R. Austin, L. Barnes, M. R. Beasley, et al., 2018. "The Role of Citizen Science in Addressing Grand Challenges in Food and Agriculture Research." *Proceedings. Biological Sciences / The Royal Society* 285 (1891). https://doi.org/10.1098/rspb.2018.1977.

Shrestha, Rosemary, Luca Matteis, Milko Skofic, Arllet Portugal, Graham McLaren, Glenn Hyman, and Elizabeth Arnaud. 2012. "Bridging the Phenotypic and Genetic Data Useful for Integrated Breeding through a Data Annotation Using the Crop Ontology Developed by the Crop Communities of Practice." *Frontiers in Physiology* 3 (August): 326. https://doi.org/10.3389/fphys.2012.00326.

Sousa, Kauê de, Jacob van Etten, Jesse Poland, Carlo Fadda, Jean-Luc Jannink, Yosef Gebrehawaryat, Basazen Fantahun Lakew, et al., 2020. "Replication Data for: 'Data-Driven Decentralized Breeding Increases Genetic Gain in Challenging Crop Production Environments.'" Harvard Dataverse. https://doi.org/10.7910/DVN/OEZGVP.

Sousa, Kauê de, Adam H. Sparks, William Ashmall, Jacob van Etten, and Svein Ø. Solberg. 2020. "Chirps: API Client for the CHIRPS Precipitation Data in R." *Journal of Open Source Software* 5 (51): 2419.

Steinke, J., J. O. Achieng, J. Hammond, and S. S. Kebede. 2019. "Household-Specific Targeting of Agricultural Advice via Mobile Phones: Feasibility of a Minimum Data Approach for Smallholder Context." *Computers and Electronics in Agriculture*. https://www.sciencedirect.com/science/article/pii/S0168169918316806.

Steinke, J., and J. van Etten. 2017. "Gamification of Farmer-Participatory Priority Setting in Plant Breeding: Design and Validation of 'AgroDuos.'" *Journal of Crop Improvement*. https://www.tandfonline.com/doi/abs/10.1080/15427528.2017.1303801.

Steinke, Jonathan, Jacob van Etten, and Pablo Mejía Zelan. 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.

Strobl, Carolin, Florian Wickelmaier, and Achim Zeileis. 2011. "Accounting for Individual Differences in Bradley-Terry Models by Means of Recursive Partitioning." *Journal of Educational and Behavioral Statistics: A Quarterly Publication Sponsored by the American Educational Research Association and the American Statistical Association* 36 (2): 135–53.

Teeken, Béla, Afolabi Agbona, Abolore Bello, Olamide Olaosebikan, Emmanuel Alamu, Michael Adesokan, Wasiu Awoyale, et al., 2020. “Understanding Cassava Varietal Preferences through Pairwise Ranking of Gari-Eba and Fufu Prepared by Local Farmer--Processors.” *International Journal of Food Science & Technology*. https://doi.org/10.1111/ijfs.14862.

Turner, H. L., J. van Etten, D. Firth, and I. Kosmidis. 2020. “Modelling Rankings in R: The PlackettLuce Package.” *Computational Statistics* 35: 1027–57. https://link.springer.com/content/pdf/10.1007/s00180-020-00959-3.pdf.

Weltzien, Eva, Fred Rattunde, Anja Christinck, Krista Isaacs, and Jacqueline Ashby. 2019. “Gender and Farmer Preferences for Varietal Traits.” *Plant Breeding Reviews*, 243–78. https://doi.org/10.1002/9781119616801.ch7.

Wijk, Mark van, James Hammond, Leo Gorman, Sam Adams, Augustine Ayantunde, David Baines, Adrian Bolliger, et al., 2020. “The Rural Household Multiple Indicator Survey, Data from 13,310 Farm Households in 21 Countries.” *Scientific Data* 7 (1): 46. https://doi.org/10.1038/s41597-020-0388-8.

Yildiz, Ilkay, Jennifer Dy, Deniz Erdogmus, Jayashree Kalpathy-Cramer, Susan Ostmo, J. Peter Campbell, Michael F. Chiang, and Stratis Ioannidis. 2020. “Fast and Accurate Ranking Regression.” In , edited by Silvia Chiappa and Roberto Calandra, 108:77–88. Proceedings of Machine Learning Research. Online: PMLR. http://proceedings.mlr.press/v108/yildiz20a.html.

Zingore, S., H. K. Murwira, R. J. Delve, and K. E. Giller. 2007. “Soil Type, Management History and Current Resource Allocation: Three Dimensions Regulating Variability in Crop Productivity on African Smallholder Farms.” *Field Crops Research* 101 (3): 296–305. https://doi.org/10.1016/j.fcr.2006.12.006.

RESEARCH PROGRAM ON Roots, Tubers and Bananas

CGIAR

The CGIAR Research Program on Roots, Tubers and Bananas (RTB) is a partnership collaboration led by the International Potato Center implemented jointly with the Alliance of Bioversity International and the International Center for Tropical Agriculture (CIAT), the International Institute of Tropical Agriculture (IITA), and the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), that includes a growing number of research and development partners. RTB brings together research on its mandate crops: bananas and plantains, cassava, potato, sweetpotato, yams, and minor roots and tubers, to improve nutrition and food security and foster greater gender equity especially among some of the world's poorest and most vulnerable populations.

www.rtb.cgiar.org

Alliance

LED BY CIP INTERNATIONAL POTATO CENTER

Bioversity International

CIAT International Center for Tropical Agriculture Since 1967 Science to cultivate change

cirad AGRICULTURAL RESEARCH FOR DEVELOPMENT

IITA Transforming African Agriculture