Open Agriculture. 2018; 3: 180–189 DE GRUYTER

Research article

Dieudonné Harahagazwe*, Bruno Condori, Carolina Barreda, Astère Bararyenya, Arinaitwe Abel Byarugaba, Danbaba Anthony Kude, Charles Lung'aho, Carolino Martinho, Daniel Mbiri, Bouwe Nasona, Bruce Ochieng, John Onditi, Jean Marc Randrianaivoarivony, Christopher M. Tankou, Alemu Worku, Elmar Schulte-Geldermann, Victor Mares, Felipe de Mendiburu, Roberto Quiroz*

How big is the potato (*Solanum tuberosum* L.) yield gap in Sub-Saharan Africa and why? A participatory approach

https://doi.org/10.1515/opag-2018-0019
received February 20, 2018; accepted June 4, 2018

Abstract: According to potato experts from ten Sub-Saharan Africa (SSA) countries working together in a community of practice (CoP) over a 3-years period, potato farmers across SSA can increase their current annual production of 10.8 million metric tons by 140% if they had access to high quality seed along with improved management practices. This paper describes this innovative new methodology tested on potato for the first time, combining modelling and a comprehensive online survey through a CoP. The intent was to overcome the paucity of experimental information required for crop modelling. Researchers, whose data contributed to estimating model parameters, participated in the study using Solanum, a crop model developed by the International Potato Center (CIP). The first finding was that model parameters estimated through participatory modelling using experts' knowledge were good approximations of those obtained experimentally. The estimated yield gap was 58 Mg ha^{-1}, of which 35 corresponded to a research gap (potential yield minus research yield) and 24 to farmers' gap (research yield minus farmer's yield). Over a 6-month period, SurveyMonkey, a Web-based platform was used to assess yield gap drivers. The survey revealed that poor quality seed and bacterial wilt were the main yield gap drivers as perceived by survey respondents.

Keywords: yield gap drivers, participatory modelling, community of practice, crop modelling

1 Introduction

Potato (*Solanum tuberosum* L.), the third most important food crop after rice and wheat, is consumed by over a billion people (Devaux et al. 2014; Haverkort and Struik 2015). In 2005, global potato production reached 325

***Corresponding author: Roberto Quiroz,** International Potato Center (CIP), P.O. Box 1558, Lima, Peru, E-mail: raquirozguerra@gmail.com
Dieudonné Harahagazwe, Daniel Mbiri, Bruce Ochieng, Elmar Schulte-Geldermann, International Potato Center (CIP), P.O Box 25171-00603, Nairobi, Kenya
Bruno Condori, United States Department of Agriculture - Agricultural Research Service (USDA-ARS), Crop Systems and Global Change Lab. 10300 Baltimore Avenue. Beltsville, MD 20705-2350, Maryland, USA
Carolina Barreda, Victor Mares, International Potato Center (CIP), P.O. Box 1558, Lima, Peru
Astère Bararyenya, National Agricultural Research Institute (ISABU), P.O Box 795, Gisozi, Burundi
Arinaitwe Abel Byarugaba, Kachwekano Zonal Agricultural Research and Development Institute (KAZARDI), Kabale, Uganda
Danbaba Anthony Kude, National Root Crops Research Institute, P.M 04, Vom, Plateau State, Nigeria
Charles Lung'aho, International Potato Center (CIP), Chimoio, Mozambique
Carolino Martinho, Mozambique Institute of Agricultural Research (IIAM). P.O. Box 238, Lichinga, Mozambique
Bouwe Nasona, Institut National pour l'Etude et la Recherche Agronomiques (INERA), Mulungu, Democratic Republic of Congo
John Onditi, Kenya Agricultural and Livestock Research Organization (KALRO), P. O. Box 338, Tigoni, Kenya
Jean Marc Randrianaivoarivony, FIFAMANOR, Antsirabe, Madagascar
Christopher M. Tankou, Department of Crop Science, University of Dschang, P.O. 222 Dschang, Cameroon
Alemu Worku, Ethiopian Institute of Agricultural Research (EIAR), P.O. Box 08, Bahir Dar, Ethiopia
Felipe de Mendiburu, Universidad Nacional Agraria La Molina, Peru

million metric tons, and has increased faster than that of any other major crop in developing countries (FAO 2009). According to FAO, potato production in Africa tripled from 1994 through 2011, from 8 to 24 million metric tons, largely due to the increase of cropping area. Same FAO data shows that the total production in Africa which was only 4% of global supply increased to 9% ten years later. However, food demand is increasing along with global population and average income (Lobell et al. 2009; Monfreda et al. 2008). This trend will be accentuated in Sub-Saharan Africa (SSA) as the region is expected to account for half of the world population increment by 2050, compared to one fifth of the increment in 1999 (Alexandratos 1999). As yields of cereals such as rice and wheat are likely to level off or even decline in many regions of the world over the next decades because maximum achieved yields are closer to the crop potential yields (Licker et al. 2010; Lobell et al. 2009), potatoes are likely to play a major role in achieving the Sustainable Development Goal of Zero hunger in SSA. Current and future cereal imports may not reverse that trend. However, this will require lots of efforts to close the huge potato yield gaps found in many developing countries, particularly in SSA where limiting factors (water, nutrients and biotic) are yet to be fully controlled. Although yield gaps for most food crops are extensively documented, what is often lacking in those studies is how stakeholders perceive the yield gaps along with their major causes, called drivers in this paper.

Yield gap (Y_g) is a simple concept: quantitative differences between a base-line yield (generally: average farmers' yield) and either attainable (generally: experiment-based yield) or potential yield (Y_p) over some specified spatial and temporal scale (Sadras et al. 2015). However, the conceptual framework for its calculation is complex (Licker et al. 2010; Lobell et al. 2009; van Ittersum et al. 2013). The most difficult task is estimating potential yield, which is defined as the yield of a cultivar when grown in an environment to which it is adapted, with optimal amounts of water and nutrients, and all biotic stresses effectively controlled. Potential yield is relevant to crops and environments where irrigation, the amount and distribution of rainfall, or a combination of irrigation and rainfall ensure that water deficits do not constrain yield (Haverkort and Struik 2015; Licker et al. 2010; van Ittersum and Rabbinge 1997). In case of rainfed systems, where non-supplemented water deficits occur, the Y_p is substituted by the water-limited potential yield (Y_w) (Lobell et al. 2009; van Ittersum et al. 2013). Since determining the Y_p (or Y_w) depends on various biophysical variables that are not precisely measured and controlled in the field, these yields are more a construct based on a number of assumptions rather than a measurable property. Therefore, the best assessment of Y_p (or Y_w) requires the integration of methods such as remote sensing, geospatial analysis and simulation models combined with field experiments and on-farm validation (Lobell et al. 2009). Three techniques are normally used to estimate potential yields (Lobell et al. 2009): (i) model simulations, (ii) field experiments, and (iii) yield contests and maximum farmer yields. Among these techniques, modelling is the most reliable (Hochman et al. 2013; Lobell 2013; van Ittersum et al. 2013) and thus the approach adopted in this study. In the literature, potential yield from simulations is defined as the 90th percentile yield achieved for a given climate/cropping season. Nevertheless, the task is not that easy in most developing countries where historical field data is limited or no experiments with sequential sampling to estimate model parameters exist. Thus, innovative approaches to overcome the problem of absent information are required.

Yield gap analysis measures untapped food production capacity (Grassini et al. 2015; Lobell et al. 2009; van Wart et al. 2013) but most Y_g analyses have been conducted on cereals (Grassini et al. 2015; van Wart et al. 2013) with limited information on other crops like potatoes. In this study we express Y_g (in Mg ha^{-1}) as the difference between the Y_p and a given base-line yield, which can either be an experiment-based or research yield (Y_r) or the average farmers' yield (Y_f), under both rainfed (water limited) and irrigated conditions. Thus, several yield gap types can be discerned. In this study, the research yield gap ($Y_{g(r)}$) is defined as the difference between the research yield (based on relevant experiments: Y_r) and Y_p, whereas the absolute yield gap (Y_g) is derived from the difference between the average farmer's yield (in a particular location: Y_f) and the Y_p. The difference between Y_f and Y_r is named farmer's yield gap ($Y_{g(f)}$). The absolute yield gap is therefore a sum of two components as shown by **Equation 1**:

$$\text{Absolute yield gap } (Y_g) = \text{research yield gap } (Y_{g(r)}) + \text{farmer's yield gap } (Y_{g(f)}) \quad (1)$$

This study is an attempt to develop an innovative consultation approach for yield gap assessment in the SSA region, based on synergies between modelling techniques and historical non-published data of potato experiments provided by potato experts organized in a community of practice (CoP). The participatory methodology used involves key local actors to reveal and analyse yield gap drivers. This methodology is novel and unique as it combines modelling and a comprehensive online survey through a CoP, and its target crop is not a cereal, but the potato. Only one global survey involving potato stakeholders in developing countries has been published to date (Fuglie 2007), but only eight responses out of fifty-five collected pertained to SSA and it was not designed for cross-constraint analysis.

2 Materials and Methods

The participatory assessment of Yg and its drivers consisted of the following three major components: (1) development of site-specific crop modelling tools and acquisition of agricultural statistics, (2) regional workshops for modelling and yield gap assessment, and (3) online survey on yield gap drivers. We simultaneously carried out the two first steps through a CoP that we established in the target area in thirteen SSA countries: Nigeria in West Africa; Burundi, Rwanda, Kenya, Uganda, Tanzania, Democratic Republic of Congo and Ethiopia in Eastern and Central Africa; Cameroon in Central Africa; and Angola, Malawi, Madagascar and Mozambique in Southern Africa. A priority setting survey based on local importance of potato crop and a composite indicator of livelihood that was commissioned by the International Potato Center (CIP) for the purpose of corporate strategic planning (Thiele et al. 2010; Devaux et al. 2014) was used to identify the target countries. The study was carried out from March 2013 to December 2016.

2.1 Framework for participatory yield gap assessment

To overcome the paucity of experimental information required for crop modelling in developing countries as reported in literature (Grassini et al. 2015; Hochman et al. 2013), we established a three-stage protocol. These steps are defined as follows: (1) develop a routine within a crop model capable of translating expert knowledge on the crop into model parameters; (2) field experiments to validate parameters estimated and (3) participatory modelling with experts to estimate potential yield and various yield gaps as defined above.

2.1.1 Parameter estimation and validation

To estimate the parameters, a routine known as Parameter Estimator was developed within the potato-specific crop model Solanum. Solanum, a crop model developed by CIP, is normally used to estimate Y_p and Y_w and is well documented in literature. It simulates tuber dry mass (DM) assimilation and partitioning for different potato species (*Solanum* sp.), varieties and hybrids, following principles of crop physiology (Condori et al. 2014, 2010; Fleisher et al. 2017). Based on the light interception and utilization (LINTUL) framework extensively described in the literature (Condori et al. 2014, 2010; Harahagazwe et al. 2012; Haverkort and Struik 2015; Kooman and Haverkort 1995; Svubure et al. 2015; van Ittersum et al. 2013) the model estimates tuber yield under non-limited, water-limited, and frost-limited growing conditions.

During two consecutive workshops, knowledge and data provided by convened experts was sorted into the Parameter Estimator. Based on allometric and heuristic methods, the tool used the relationship between aerial and tuber partitioning crop growth functions to estimate crop parameters. Parameter Estimator was built upon three principles. First, mathematical functions describing either canopy cover or the portion of DM partitioned to the tubers can be generic throughput the crop (i.e. will not change with different varieties or environmental conditions), however, the equations - specific parameters for the generic function - might be different for each variety. Second, parameters were estimated through numerical methods by forcing the function to fit a minimum number of data points. Lastly, for all the varieties included in the analysis, expert knowledge provided the pre-defined minimum number of data points needed to fit the function as well as all input needed to estimate canopy cover and partition to the tuber. The input data provided by experts to estimate sets of site-specific parameters of the model comprised the following variables: planting and harvest dates, days to reach 1% canopy cover (corresponding to the emergence day), days at maximum canopy cover, maximum canopy cover index, days at physiological maturity and, optionally, days at tuber initiation. Participants also provided daily temperature and solar radiation data needed to model the yield potential. Workshop participants discussed the estimated parameters in groups until they reached a consensus.

In this study, we used the Beta function for canopy cover evolution (**Equation 2** from Yin et al. 2003) and the Gompertz function for tuber partition over time (**Equation 3** from Winsor 1932). We estimated the onset of tuber initiation as the minimum of the partition function (second derivative=0).

Beta function:

$$W = W_{max} (1+(t_e-t)/(t_e-t_m))(t/t_e)^{(te/(te-tm))} \quad \text{with } 0 \le t_m \le t_e \qquad (2)$$

Where W_{max} = maximum canopy cover value, t_m = thermal time at maximum canopy cover growth rate, t_e = thermal time at maximum canopy cover value.

Gompertz function:

$$Y = A * \exp(-\exp(-(t-T_u)/b)) \qquad (3)$$

Where A = maximum harvest index, T_u = thermal time at maximum tuber partition rate, b = thermal time before tuber initiation process.

The first version of Parameter Estimator was written in the program R, in which numerical solutions were implemented. For the Beta function, the bisection numerical method for analysis of nonlinear functions was used. For the tuber partition curve, algebraic analysis was most suitable to solve the unknown function.

In preparation to the first workshop held in Nairobi, Kenya on 24-26 June 2013, the method had been evaluated on 14 local varieties from four locations in Peru and Bolivia. Sequential harvest data from all experiments were used by the modelling team to estimate the parameters described above, through the fitting of the Beta and Gompertz functions, using nonlinear techniques. Agronomists responsible for the experiments were also requested to use their expert knowledge with the Parameter Estimator and the parameters thus estimated were statistically compared with those estimated from experimental data. All statistical metrics (described in next paragraph) showed that experts can reliably estimate growth parameters using the Parameter Estimator. Participants in the first workshop, after using the R-based Parameter Estimator, requested a user-friendlier version. The routine was therefore improved, re-programmed and included within the Solanum model. Participants in the second workshop held in Addis Ababa, Ethiopia on 14-18 October 2013, therefore used a new version of Solanum for both generating parameters and modelling. They used the values obtained and groups of researchers working with the same varieties agreed to model the productivity of “their” potato.

To validate values generated by the Parameter Estimator, field experiments in four countries (Cameroon, Democratic Republic of Congo, Kenya and Uganda) using a standardized protocol for the set up and data collection were later carried out. Participating scientists conducted the experiment from January to August 2014, following the agricultural calendar in each country. After harvest, results were brought for discussion and analysis in a third workshop organized in Entebbe, Uganda on 15-19 December 2014. To assess the accuracy of estimated parameters, they were compared with experimental data using three statistical metrics: (i) the Pearson correlation coefficient (r); (ii) the relative root mean square error (RRMSE) and (iii) the relative mean absolute error (RMAE), which are commonly used in model evaluation studies (Chai and Draxler 2014; Willmott and Matsuura 2005).

2.1.2 Yield gap assessment

Over twenty-five experienced breeders and field researchers from ten SSA countries, who provided data and contributed to the estimation of the required crop parameters for modelling, participated in the yield gap assessment. These countries (and the locations within countries) were Burundi (Rwegura), Cameroon (Fongo-Tongo), Democratic Republic of Congo (Mulungu), Ethiopia (Adet), Kenya (Tigoni, Kabuku, and Kabete), Madagascar (Mimosa), Mozambique (Sussundenga), Nigeria (Kuru), Uganda (Kalengyere) and Malawi (Bembeke). Sites were geo-referenced using the coordinates given by participants who then validated the exact position in Google Earth, making adjustments when necessary.

Participants produced their own yield gap results and then validated these against their own field data and knowledge. Participants simulated the potential yield of a total of twelve potato genotypes used in the respective research programs: CIP381381.20 also known as Asante or Victoria, Dosa, Guassa (CIP384321.9), Gudene (CIP386423.13), Kenya Mpya (CIP393371.58), Unica (CIP392797.22), Meva (CIP377957.5), CIP381381.13 also called Lulimile or Tigoni, Diamant, CIP395112.9, CIP396038.107 and CIP396036.201. Most of the genotypes belong to the CIP germplasm except two, Dosa grown in Cameroon and Diamant grown in Nigeria. Local farmers grow the first nine genotypes listed here. Participants provided average farmers’ yields from the neighbourhood of their experimental sites/research stations. Sources for average farmers’ yields varied but the major sources cited were the Ministries of Agriculture, FAO, own surveys, scientific papers and other reports. It is important to note that a regional CoP on potato yield gap was established, as the work required many interactions throughout the study.

2.2 Yield gap drivers: data acquisition and analysis

During the Entebbe workshop, major yield gap drivers were discussed and a preliminary list of over 40 challenges was generated. This list was then shared with other potato experts for enrichment through the virtual CoP to form the basis of the survey. In the end, thirty yield gap drivers were used in the survey. In designing this survey, possible linkages between drivers like seed quality and diseases were considered though not explicitly mentioned in the survey questionnaire, to avoid confusing the respondents. The CoP suggested keeping the questionnaire as simple as possible without grouping drivers.

Yield gap drivers in SSA were assessed over a 6-month period using the paid online survey platform SurveyMonkey (www.surveymonkey.com). This tool allows reaching out to a wider audience and getting real-time results (Parsa et al. 2014). The survey comprised 15 closed-ended questions allowing responders to provide their personal assessment of actual yield levels and rate the importance of the thirty yield gap drivers previously identified. Three dominant agro-ecologies where potato is grown were considered: tropical and sub-tropical highlands (over 1,800 masl), tropical and sub-tropical mid elevation (800-1,800 masl) and sub-tropical lowlands (winter potato found in Southern Africa).

Participants rated each of the thirty yield gap drivers using a Likert scale of 1 to 5 where 1 indicated *Not important* and 5 *Very important* (Parsa et al. 2014). To customize respondents experience over specific agro-ecologies, we enabled Question Skip Logic feature of SurveyMonkey, which allows respondents to skip a question, depending on the previous answer. The questionnaire comprised technical questions and non-identifying personal information, including current country base, gender, education degree, experience with the crop in SSA and area of expertise. The survey was uploaded in three languages spoken in the target countries: English, French and Portuguese.

The survey went live on 8 April 2015, after a test period of several weeks, and remained online for almost 6 months. We administered this survey through focal points in each target country as suggested by similar studies (Fuglie 2007; Parsa et al. 2014). In the tool, we set a password that we shared with potential respondents, as we wanted to prevent non-invited people accessing the platform. Initially we sent out invitations in the thirteen target countries. Later, we extended the survey to other people with long potato experience in SSA, even if they were no longer living in the region.

Prior to analysis, data accuracy and quality were checked and non-complying responses discarded. The consistency of the ordinal responses to the questions in the survey was assessed using Cronbach's Alpha index (Cronbach 1951). This index establishes the relationship between the variability of responses to each question and variability of the surveys. If all answers to a particular question were the same, they were deemed uniform and the variance equal to zero. In that case, the index would approach the unit, weighted by k/(k-1), where k would be the number of questions. If answers were very variable, the index would approach zero or less than zero in exceptional cases.

Likert scores were analysed using ordinal regression (Anderson 1984; McCullagh 1980) to link the categorical responses to a list of pre-defined factors known to limit potato productivity and determined those perceived as of high or low importance for each agro-ecology. Conditional probabilities for the categorical variable (scores) were estimated for each driver within agro-ecologies. To do so, an ordinal regression model of the logit or odd (i.e. the natural log of the ratio of the probability that the event occurs to the probability that the event does not occur) as a function of the explanatory factors (scores) to estimate the regression coefficients was run. These coefficients were then used to estimate the conditional probabilities as per McCullagh (1980). Five probabilities were computed for each driver in each agro-ecology. The probability associated to score = 3 or P (3) was considered to be neutral ($P_{neutral}$). The probability that a driver was perceived as not important P (1) or somewhat important P (2) were added to build an overall low importance probability (P_{low}). By the same logic, P (4) and P (5) represented the high importance probability (P_{high}). Note that the probability of equal outcomes is 1/3 or 0.33 and thus a probability > 0.33 will show the dominance of one of the three possibilities: P_{high}, $P_{neutral}$ and P_{low}. In order to minimize the chance of misinterpreting perceptions and to define whether a yield gap driver was perceived to have low, neutral or high probability , we assigned a very high cut off point of P=0.60.

Ethical approval: The conducted research is not related to either human or animals use.

3 Results and discussion

3.1 Model parameters

Results showed that the Parameter Estimator is a good tool to estimate model parameters based on expert knowledge of the crop (Table 1).

The distribution of the different parameters generated by the Parameter Estimator, calibrated with data from controlled experiments conducted by participating scientists is presented in Figure 1. Simulated and observed yield comparison showed an RMSE of 5.99 using parameters estimated with standard procedures, and an RMSE of 7.64, using Parameter Estimator tool.

3.2 Magnitude of potato yield gaps in SSA

If genotypes, seasons and sites are disregarded, the average yields calculated were: 66.35 (+/- 2.52), 31.15 (+/- 1.87) and 8.02 (+/- 0.71) Mg ha^{-1} for potential (Y_p), research (Y_r) and average farmers (Y_f), respectively. The absolute yield gap was therefore estimated to be 58.33 Mg ha^{-1}, as shown in Figure 2.

These results are consistent with relative yields (i.e. actual yields to simulated potential yields) that ranged from 8 to 35% found in Zimbabwe using LINTUL-POTATO Model (Svubure et al. 2015). Large yield gaps in the context of African smallholder farmers have been reported in the literature but most relate to cereals (Tittonell and Giller 2013). For example, global actual yields for maize are reported to be around 50% of the potential yield (Neumann et al. 2010) against 20% in Africa, which is due to biophysical and management conditions (Lobell et al. 2009). In the current study, the farmers' average yield gap, the one that the extension systems and other development agents strive to fill, was 23.54 Mg ha^{-1}. On the other hand, the average researcher gap, the difference between potential yield and yields obtained by researchers, was 35.20 Mg ha^{-1} (Figure 2 right). This is the yield gap that researchers try to reduce through introduction of genotypes and good agricultural practices.

The current average farmers' potato yield in SSA (5-year period, 2010-2014) is 10.3 Mg ha^{-1} (FAO, 2017). If farmers there could close the current yield gaps, SSA countries

Table 1. Comparative assessment of model parameters as estimated by sequential harvest or with the Parameters Estimator

Statistical metrics	Estimated model parameters					
	W_{max}	t_m	t_e	A	T_u	b
r	0.856	0.932	0.978	0.824	0.959	0.869
RRMSE	0.126	0.155	0.029	0.058	0.051	0.132
RMAE	0.094	0.171	0.02	0.046	0.044	0.108

r = Pearson correlation coefficient; RRMSE = Relative root mean square error; RMAE = Relative mean absolute error; W_{max} = maximum value of canopy cover; t_m = thermal time at maximum canopy cover rate; t_e = thermal time at the end of the growth period; A = maximum value of tuber partition; T_u = thermal time at maximum tuber partition rate; b = thermal time just before the tuber initiation process.

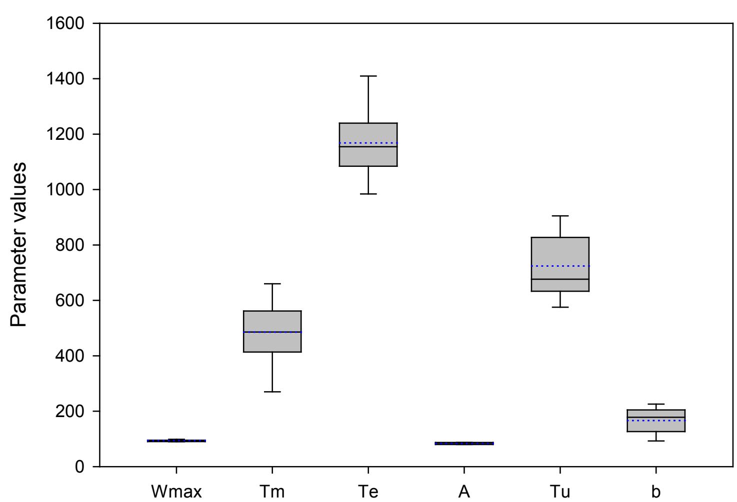

Figure 1: Distribution of generated parameters used for modelling in Solanum for 16 sites. W_{max} (%) = maximum value of canopy cover; t_m = thermal time at maximum canopy cover rate; t_e = thermal time at the end of the growth period; A (%) = maximum value of tuber partition; T_u = thermal time at maximum tuber partition rate; b = thermal time just before the tuber initiation process. The boxplots show the medians (solid line), the boxes and whisker represent 25th to 75th and 10th to 90th percentiles, respectively. Thermal times are expressed in °C days.

would easily increase current annual production of 10.8 million metric tons by 140%. This estimation is based on the assumption that the correction factor (loss) for extrapolating experimental plot yield to farm yield of 30% (J.-F. Ledent, Personal communication). For example, by raising farmers' yields to 60% of the potential yield (e.g. achieving a Y_f / Y_p ratio of 0.6 as done in the Netherlands and Unites States of America (Haverkort and Struik 2015)) the current annual total production in SSA could be 37 million metric tons, i.e. a more that threefold increase, without expanding production areas.

3.3 Perceived potato yield gap drivers in SSA

During the 5-months and 20-day run of the survey, 119 responses were collected, from 19 countries (13 from Africa and 6 from other continents). Only 10 responses were not useful for analysis. Out of the 109 valid respondents, only 12% were females, showing a gender imbalance in potato research and development in Africa. Respondents reported to work in research (71.4%), rural development/ government (29.4%), NGOs (17.6%), private sector (10.8%) and others (5.9%). Scientists responding had experience in agronomy (52.9%), seed production (69.0%), extension (37.9%), breeding (37.9%), phytopathology (34.5%), storage and processing (16.1%), crop modelling (10.3%) and socio-economics (10.3%). Since the study was based on expert's judgement, 73.3% of respondents had at least received an MSc and 48.6% of respondents had worked on potato for at least six years (Figure 3).

A high consistency of all the Likert-based responses was evidenced by a Cronbach's alpha value of 0.94. On their opinion regarding potato yields in SSA, 65.2% of

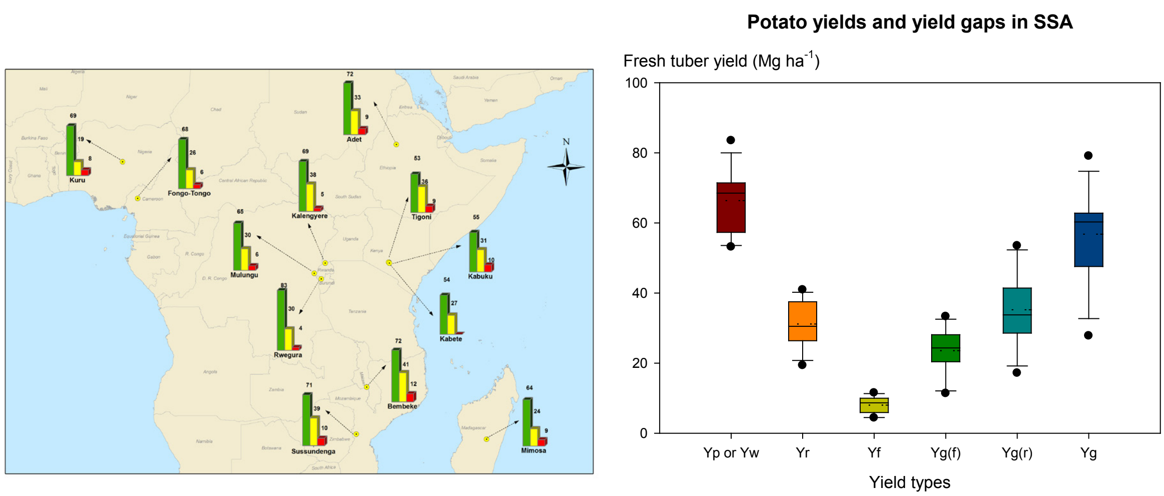

Figure 2: Levels of potato yields and yield gaps in selected sites of Sub-Saharan Africa: georeferenced graphs showing Y_p, Y_r and Y_f (**left**); box-plots (**right**). The X-Axis presents the following yields: Y_p = potential yield; Y_w = water-limited potential yield used for Winter potato (Bembeke and Sussundenga); Y_r = maximum yield attained by researchers; Y_f = actual yield obtained by farmers; $Y_{g(r)}$ = research yield gap; $Y_{g(f)}$ = farmers' yield gap; Y_g = absolute yield gap. Lines within boxes show the medians (solid line) and the boxes and whiskers represent 25th to 75th and 10th to 90th percentiles, respectively. Green, yellow and red graphs on the left represent Y_p or Y_w, Y_r and Y_f, respectively

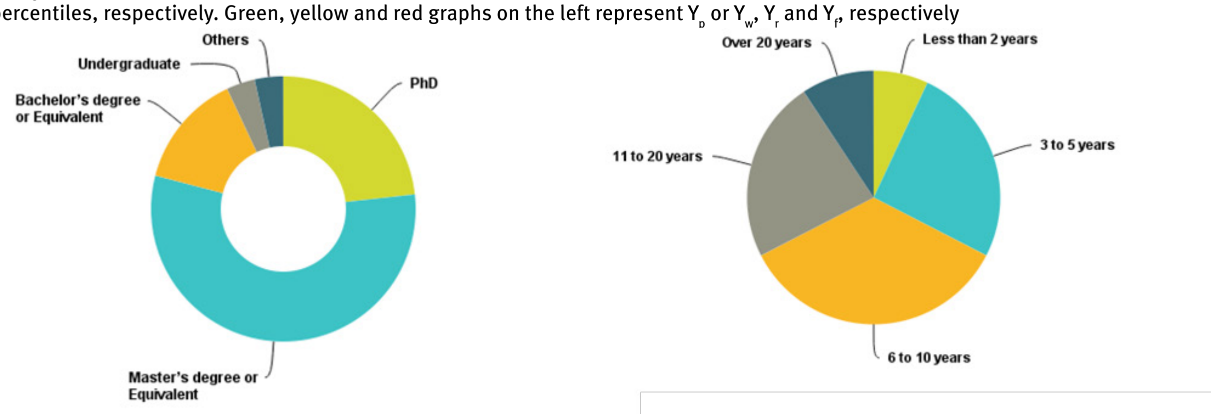

Figure 3: Biographic information on respondents: education level (left) and length of work experience with potato in SSA (right)

respondents replied that they were not satisfied. Out of the thirty drivers assessed, thirteen turned out to be the most explanatory ones for farmers' yield gaps (Table 2).

Poor quality seed, which had a very high probability, was the top-ranked yield gap driver identified by survey respondents. The expected probability of experts rating poor quality seed as an important or very important problem is at least 95%, with minor changes over agro-ecologies. Seed-borne diseases such as bacterial wilt and viruses were deliberately separated from the seed driver for mainly two reasons: they are not the only determinants of seed tuber quality, and perceptions from respondents on diseases-related drivers were based on visible observations in the field and not on laboratory test results. Seed quality as a potato yield-limiting factor has been extensively reported (Fuglie 2007; Haverkort and Struik 2015; Schulte-Geldermann et al. 2012; Thiele 1999; Thomas-Sharma et al. 2016). What was not documented so far was the quantified perception that practitioners had about its importance in SSA in a cross-analysis of drivers with data from a large number of respondents. Haverkort and Struik (2015) defined quality seed as being seed tubers with (i) good physical characteristics, (ii) the right physiological age and (iii) the best possible health (i.e. free of pests and diseases). In a study conducted in Argentina, tuber yield was found to be correlated with the physiological age of seed potato which is a combination of several factors, including types of cultivars, seed origin, haulm killing date, storage conditions and pre-planting treatments, if any (Caldiz 2000).

Bacterial wilt caused by *Ralstonia solanacearum* (Smith 1896; Yabuuchi et al. 1996) was the second most important driver, and had a high probability, which corroborates earlier findings that bacterial wilt is the most damaging biotic constraint in SSA (Fuglie 2007; Lemaga et al. 2001). However, strategies for its control exist and are especially effective when they are part of an integrated approach. These strategies include clean seed potato, pathogen-free soil, removal of wilting and/or volunteer plants, appropriate crop rotation systems with non-host plant species, negative/positive selection techniques and other agronomic practices. Other perceived drivers are listed in Table 2, but were mentioned much less frequently by survey respondents. Cloud cover, frost, hailstorms or salinity were perceived not important to affect the yield gaps.

Table 2: Probabilities of highly important and less important potato yield gap drivers in SSA

Drivers	Highlands (tropical and sub-tropical, n=66)	Mid-elevations (tropical and sub-tropical, n=46)	Lowlands (sub-tropical, n=27)
HIGH IMPORTANCE			
Poor seed quality	0.9598	0.9611	0.9639
Bacterial wilt	0.9001	0.9032	0.9097
Poor soil health	0.7545	0.7608	0.7745
Late blight	0.7455	0.7519	0.7661
Lack/inappropriate use of fertilizers	0.7352	0.7418	0.7564
Viruses	0.7083	0.7154	0.7308
Soil amendments (sub-optimum use of lime)	0.7007	0.7079	0.7236
Low yielding varieties	0.6910	0.6983	0.7143
Pests (aphids, leafminers, potato tuber moth)	0.6725	0.6800	0.6965
Farmers knowledge (Extension)	0.6292	0.6371	0.6548
Poor timeliness of operation	0.6143	0.6223	0.6403
Lack of access to market (as incentive)			0.6083
LOW IMPORTANCE			
Too much and persistent clouds	0.7649	0.7588	0.7443
Frost	0.7536	0.7472	0.7323
Hailstorm	0.7204	0.7135	0.6974
Salinity	0.6445	0.6367	0.6185

n=number of respondents. Soil health includes fertility, fauna and flora.

4 Conclusion

The use of modelling tools was crucial for achieving the study goals. The validated methodology can be implemented with online tools to substitute costly face-to-face workshops. The participatory approach through a CoP proved to be effective for accessing a wealth of knowledge. With farmers' productivity estimated to be about 12% of the potential yield in SSA, there is no doubt that these gaps can be reduced with sound technological (seed, fertilizers, pesticides, etc.), service delivery, policy, infrastructure and capacity building interventions.

Acknowledgements: We would like to thank the CGIAR Research Programs on Integrated Humid Tropics (HT), Roots, Tubers and Bananas (RTB) and Climate Change, Agriculture and Food Security (CCFAS) and the CGIAR Fund donors for funding this research. For a list of CGIAR Fund donors please see: http://www.cgiar.org/who-we-are/cgiar-fund/fund-donors-2/.

Authors are very grateful to all participants to workshops and respondents to the online survey. We sincerely thank Professor Emeritus Jean-François Ledent of the Catholic University of Louvain, Belgium for providing valuable comments on earlier versions of the manuscript.

Conflict of interest: Authors state no conflict of interest.

References

Anderson T.W., Estimating linear statistical relationships, Ann. Stat., 1984, 12, 1–45

Caldiz D.O., Analysis of seed and ware potato production systems and yield constraints in Argentina, PhD thesis, Wageningen University, Wageningen, The Netherlands, 2000

Chai T., Draxler R.R., Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? - Arguments against Avoiding RMSE in the Literature, Geosci. Model Dev., 2014, 7, 1247-1250

Condori B., Hijmans R.J., Ledent J.F., Quiroz R., 2014. Managing Potato Biodiversity to Cope with Frost Risk in the High Andes: A Modeling Perspective, PLOS ONE, 2014, 9, 1, e81510

Condori B., Hijmans R.J., Quiroz R., Ledent J.F., Quantifying the expression of potato genetic diversity in the high Andes through growth analysis and modeling, Field Crops Res., 2010, 119, 135-144

Cronbach L.J., Coefficient alpha and the internal structure of tests, Psychometrika, 1951, 16, 297-334

Devaux A., Kromann P., Ortiz O., Potatoes for Sustainable Global Food Security, Potato Res., 2014, 57, 185-199

FAO, FAOSTAT online databases www.fao.org/faostat/, 2017

FAO, International Year of the Potato 2008: New light on a hidden treasure, FAO, Rome, Italy, 2009

Fleisher D.H., Condori B., Quiroz R., Alva A., Asseng S., Barreda C., et al., A potato model intercomparison across varying climates and productivity levels, Glob. Change Biol., 2017, 23, 1258-1281

Fuglie K.O., Priorities for potato research in developing countries: Results of a survey, Am. J. Potato Res., 2007, 84, 353

Grassini P., van Bussel L.G.J., Van Wart J., Wolf J., Claessens L., Yang H., et al., How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis, Field Crops Res., 2015, 177, 49-63

Harahagazwe D., Ledent J.F., Rusuku G., Growth analysis and modelling of CIP potato genotypes for their characterization in two contrasting environments of Burundi, Afr. J. Agric. Res, 2012, 7, 46, 6173-6185

Haverkort A.J., Struik P.C., Yield levels of potato crops: Recent achievements and future prospects, Field Crops Res., 2015, 182, 76-85

Hochman Z., Gobbett D., Holzworth D., McClelland T., van Rees H., Marinoni O., et al., Reprint of "Quantifying yield gaps in rainfed cropping systems: A case study of wheat in Australia." Crop Yield Gap Anal. - Ration, Methods Appl, 2013, 143, 65-75

Kooman P.L., Haverkort A.J., 1995. Modelling development and growth of the potato crop influenced by temperature and daylength: LINTUL-POTATO, In: Haverkort A.J., MacKerron D.K.L. (Eds.), Potato Ecology And Modelling of Crops under Conditions Limiting Growth: Proceedings of the Second International Potato Modeling Conference (17-19 May 1994, Wageningen, The Netherlands), Dordrecht, 1994, 41-59

Lemaga B., Siriri D., Ebanyat, P., Effect of soil amendments on bacterial wilt incidence and yield of potatoes in southwestern Uganda, Afr. Crop Sci. J., 2001, 9

Licker R., Johnston M., Foley J.A., Barford C., Kucharik C.J., Monfreda C., et al., Mind the gap: how do climate and agricultural management explain the 'yield gap' of croplands around the world?, Glob. Ecol. Biogeogr., 2010, 19, 769-782

Lobell D.B., The use of satellite data for crop yield gap analysis. Crop Yield Gap Anal. - Ration, Methods Appl., 2013, 143, 56-64

Lobell D.B., Cassman K.G., Field C.B., Crop Yield Gaps: Their Importance, Magnitudes, and Causes, Annu. Rev. Environ. Resour., 2009, 34, 179-204

McCullagh P., Regression models for ordinal data, Ournal R. Stat. Soc. Ser. B Methodol., 1980, 42, 109-142

Monfreda C., Ramankutty, N., Foley, J.A., Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000, Glob. Biogeochem. Cycles, 2008, 22, 1-19

Neumann K., Verburg P.H., Stehfest E., Muller C., The yield gap of global grain production: A spatial analysis, Agric. Syst., 2010, 103, 316-326

Parsa S., Morse S., Bonifacio A., Chancellor T.C.B., Condori B., Crespo-Pérez V., et al., Obstacles to integrated pest management adoption in developing countries, Proc. Natl. Acad. Sci., 2014, 111, 3889

Sadras V.O., Cassman K.G., Grassini P., Yield gap analysis of field crops: Methods and case studies, FAO, Rome, Italy, 2015

Schulte-Geldermann E., Gildemacher P.R., Struik P.C., Improving Seed Health and Seed Performance by Positive Selection in Three Kenyan Potato Varieties, Am. J. Potato Res., 2012, 89, 429-437

Svubure O., Struik P.C., Haverkort A.J., Steyn J.M., Yield gap analysis and resource footprints of Irish potato production systems in Zimbabwe, Field Crops Res., 2015, 178, 77-90
Thiele G., Informal potato seed systems in the Andes: Why are they important and what should we do with them?, World Dev., 1999, 27, 83-99
Thomas-Sharma S., Abdurahman A., Ali S., Andrade-Piedra J.L., Bao S., Charkowski A.O., et al., Seed degeneration in potato: the need for an integrated seed health strategy to mitigate the problem in developing countries, Plant Pathol., 2016, 65, 3-16
Tittonell P., Giller K.E., When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. Crop Yield Gap Anal. - Ration, Methods Appl., 2013, 143, 76-90
van Ittersum M.K., Cassman K.G., Grassini P., Wolf J., Tittonell P., Hochman Z., Yield gap analysis with local to global relevance - A review. Crop Yield Gap Anal. - Ration, Methods Appl., 143, 4-17
van Ittersum M.K., Rabbinge R., Concepts in production ecology for analysis and quantification of agricultural input-output combinations, Field Crops Res., 1997, 52, 197-208
van Wart J., van Bussel L.G.J., Wolf J., Licker R., Grassini P., Nelson A., et al., Use of agro-climatic zones to upscale simulated crop yield potential. Crop Yield Gap Anal. - Ration, Methods Appl., 2013, 143, 44-55
Willmott C.J., Matsuura K., Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, Clim. Res., 2005, 30, 79-82
Winsor C.P., The Gompertz Curve as a Growth Curve, Proc. Natl. Acad. Sci. U. S. A., 1932, 18, 1-8
Yin X., Goudriaan J., Lantinga E.A., Vos J., Spiertz H.J., A flexible sigmoid function of determinate growth, Ann. Bot., 2003, 91, 361-371