A literature review on yield gaps of various root, tuber and banana crops as a background for assessing banana yield reductions due to pests and diseases at a field site in western Burundi

Guy Blomme¹*, Walter Ocimati², Alexandra Zum Felde³, David Amwonya⁴ and Deo Kantungeko⁵

¹Bioversity International, P. O. Box 5689; Addis Ababa, Ethiopia.
²Bioversity International, P. O. Box 24384, Kampala, Uganda.
³Independent consultant, Germany
⁴International Institute of Tropical Agriculture, P. O. Box 7878, Kampala, Uganda.
⁵International Institute of Tropical Agriculture, Bujumbura, Burundi.

Received 19 May, 2020; Accepted 27 July, 2020

Banana pests (corm weevil and root nematodes) and diseases (Xanthomonas wilt of banana, banana bunchy top disease and fusarium wilt) are major constraints to banana production in Central Africa. The pests cause various degrees of yield reduction, while plants affected by three of the diseases eventually die before producing an edible bunch. Studies on yield gaps for most of these constraints are currently limited. This paper reviews yield gap studies of some root, tuber and banana crops broadly and with a specific focus on biotic constraints. It also presents an initial case study conducted in Burundi to understand yield gaps due to various banana pests and diseases. Bunch weights of banana varied widely at production zones in western Burundi due to biotic constraints. Boundary line analysis revealed large yield gaps due to the various pests. The often sub-optimal, medium and small bunch sizes found in visibly healthy fields however indicate that in addition to mitigating effects of biotic constraints, significant improvements in bunch weights could be attained through the application of agronomic/field management practices that enhance soil fertility, soil moisture content and soil health. Simple and robust methods [such as the boundary line analysis] for estimating yield gaps caused by pests and diseases, and abiotic constraints on farm are crucial for informing/guiding on the need to apply agronomic and/or disease control efforts. In addition, continuous/sustained field monitoring, with the involvement of farmers, over time will be necessary for a more accurate assessment of yield gaps caused by diseases and pests.

Key words: Banana bunchy top disease, boundary line analysis, fusarium wilt, nematodes, weevils, Xanthomonas wilt.

INTRODUCTION

The Great Lakes region of east and central Africa, of which Burundi is part, is endowed with a broad...
diversity of banana cultivars spread across a wide range of altitudes and agro-ecological zones. This region also constitutes one of the secondary centers of *Musa* diversity, more specifically for the East African highland bananas. AAA-EEH (Karamura et al., 2004). Banana and plantain (*Musa* sp.) are an important staple and income-generating crop for rural communities in this region.

Banana pests and diseases however severely constrain banana production in Central Africa (Blomme et al., 2013, 2017; Ocimati et al., 2013). In Western Burundi, Xanthomonas wilt of banana, banana bumpy top disease and fusarium wilt occur together in production landscapes and farms ranging from the humid lowlands to the cooler high altitude hilly zones (Le Point et al., 2012, 2013; Ocimati et al., 2013). In addition, weevil and nematode pests are present (Blomme et al., 2012; Ocimati et al., 2013). Xanthomonas wilt has been reported to cause yield losses of up to 100% in ‘Bluggoe’/'Pisang Awak' [Musa ABB type]-dominated systems in Uganda (Blomme and Ocimati, 2018), while Fusarium wilt decimated the ‘Sukari Ndizi’ (AAB) production zones in Rwanda (Karangwa et al., 2016). Losses from banana bunchy top disease (BBTD) are significant in Cavendish production systems (e.g., in Malawi), while they are more moderate in plantain (AAB) and ‘Bluggoe’/’Pisang Awak’ dominated systems (Ngama Boloy et al., 2014; Mikwamba et al., 2020). This paper reviews yield gap studies of some root, tuber and banana crops broadly and with a specific focus on biotic constraints. It also presents an initial case study conducted in Burundi, using proven yield gap assessment methodologies, to understand yield gaps due to banana pests and diseases.

**Banana yield gap review**

**What are yield gaps?**

Though the concept of yield gaps is not new, and an increasing number of studies have calculated it for a variety of crops, particularly grain crops, a standard definition of a yield gap does not exist. Sumberg (2012) lists the following variable definitions of yield gap in his review paper:

(i) Becker et al. (2003): the difference between actual farmers’ yield and calculated average potential yield.

(ii) Cassman et al. (2003): the “exploitable” yield gap is the difference between yield potential and the actual yield achieved by farmers.

(iii) de Bie (2004): the difference of the average production situation with the anticipated best one.

(iv) Zinck et al. (2004): the gap between the actual crop yield and the expected yield.


(vi) Ortiz-Ferrara et al. (2007): the gap between farmers’ and experimental yields.

(vii) Lobell et al. (2007): the difference between average and maximum yields.

(viii) Audebert and Fofana (2009): the difference between simulated yields and observed yields.

More recently, in their introduction to the special issue on Crop Yield Gap Analysis, van Ittersum and Cassman (2013) define yield gap as the difference between the yield under optimum management and the average yield achieved by farmers, while Sheehy et al. (2015) define it as the difference between potential (and water-limited potential) yield and average farm or actual yield, under the same environment.

With this relatively wide variety of definitions, not all of which are interchangeable, it is important to clearly define what is considered a yield gap and how to calculate it, particularly as the definition also determines which methods can or should be used to calculate the yield gap.

Generally, yield gaps (Yg) can be defined as the difference between a non-limited yield or potential yield (Yp) - e.g., the yield of a crop cultivar when grown without any water or nutrients limitations and with biotic stresses effectively controlled (Evans, 1993; van Ittersum and Rabbinge, 1997) – and actual yields (Ya). For rain-fed crops, van Ittersum et al. (2013) suggest that the water-limited yield (Yw), equivalent to the maximum potential yield when water is limited, is the most relevant benchmark. They define Yw as “similar to Yp, but crop growth is also limited by water supply, and hence influenced by soil type (water holding capacity and rooting depth) and field topography (runoff)”.

Often, yields derived in yield experiments in agricultural research stations, under the best management conditions currently known and applicable for any given crop growth environment are defined as Yp, though Yengoh and Ardö (2014) refer to these as the maximum attainable yield. Based on Lobell et al. (2009) and van Ittersum et al. (2013) distinguish four methods to estimate yield gaps at the local level: (1) field experiments, (2) yield contests (where participating farmers compete against each other to achieve maximum yields for a certain crop in a given season), (3) maximum farmer yields based on surveys, and (4) crop model simulations. For a given crop, the first step in each method is to estimate potential yields (Yp and Yw) in a certain location or region. Yg is then calculated as the difference between farmer’s actual yields (Ya) and estimated potential yields (Yp or Yw).

Alternatively, the Yp can be simulated using crop models. However, there are no published guidelines about standard sources and quality of data input for weather, soil, actual yields, and cropping-system context, or requirements for calibration of crop models used in such studies (Grassini et al., 2015). Nonetheless, a robust approach to simulate accurate crop yield potential and estimate Yg requires: (i) input data that meet minimum quality standards at the appropriate spatial scale, (ii) agronomic relevance with regard to cropping...
system context, (iii) proper calibration of crop models used, and (iv) flexibility and transparency to account for different scenarios of data availability and quality (Grassini et al., 2015).

**Yield gaps in roots, tubers and banana (RTB) crops**

Most of the yield gap studies have looked at the yields of grain crops (esp. wheat, maize or rice), while the yields and yield gaps of RTB crops have generally been given less attention. Recently, potato (Solanum tuberosum L.) yields and yield gaps have been studied in some depth (Haverkort et al., 2014; Haverkort and Struik, 2015; Svubure et al., 2015), and other RTB crops (cassava (Fremont et al., 2009) and banana (Wairegi et al., 2010)) have received some attention. In general, two types of yield gap studies have been done: very specific, looking at one constraint (e.g. % losses due to different levels of pests or diseases on a specific cultivar) or model based, looking at a variety of theoretical constraints and attributing them levels of loss (e.g., based on crop models to estimate potential yields).

**Yield gaps in potato - simulated and actual yields**

Haverkort et al. (2014) determined potato yield gaps across a range of agroecological zones in six distinct potato-growing areas in Chile using a crop simulation model. They defined a yield gap as being the difference between the potential simulated yield ($Y_{pot}$) and actual yield ($Y_{act}$). Actual yields were collected by surveying growers, whose cropping systems were also characterized with regard to field size (smaller or larger than 10 ha per season grown), technology usage levels and aim of potato production (seed, early or late potato). The potential yields were calculated using the LINTUL-POTATO simulation model into which monthly weather data (maximum and minimum temperatures, solar radiation, precipitation and evaporation) acquired from local meteorological stations, soil data and planting and harvesting dates were fed.

Actual and potential yields, as well as yield gaps varied among locations and the latter were generally larger on small farms than on "not so small farms". The average yield across productions systems was 31 t ha$^{-1}$ (range: approx. 15-50 t ha$^{-1}$), while the average potential yield was 74 t ha$^{-1}$ (range: approx. 20-120 t ha$^{-1}$), leading to a yield gap of 43 t ha$^{-1}$ or of 58%.

In areas with higher potential yields, actual average yields were about 40% of potential yields, that is, a 60% yield gap, while in areas with relatively low potential yields, actual yields approached potential ones, leading to a much smaller yield gaps. Individual contributing factors to the yield gaps were not identified (that is, diseases, low soil fertility, etc.), but the authors noted that daily growth rate explained almost 70% of the variation in reported actual yields; while the length of the growing season was not clearly related to tuber yields. Increasing levels of technology (e.g., application of fertilizers and/or irrigation water) helped to close the yield gap.

In this paper, yield gaps are graphically presented as a ratio of $Y_{act}/Y_{pot}$, which ranges from approximately 0.15 to 1.20, with 17 of the 20 systems analyzed ranging between 0.20 and 0.70. Haverkort and Struik (2015) consider a $Y_{act}/Y_{pot}$ ratio of about 0.6 in potatoes to be close to an economic optimum (that is, where growers apply all inputs and carry out all necessary cultural practices in a balanced and optimal manner).

In the same vein, they define potential yields ($Y_{pot}$) as "those not necessarily economical but the highest yields when all inputs are supplied such that the crop faces no shortage" (Haverkort and Struik, 2015). So the potential yield of potato is the theoretical yield that can be calculated or modelled for a certain cultivar grown in a certain environment without any limiting or reducing factor being present. They list a number of papers in which this approach is described in greater detail, and highlight two models for potatoes: “the simple, robust LINTUL-POTATO model” (Kooman and Haverkort, 1995) – also used in Haverkort et al. (2014), and “the more complicated, but also more versatile model GECROS” (Khan et al., 2014).

Svubure et al. (2015) also used the LINTUL-POTATO model to simulate the yield of potato – both potential yield ($Y_p$) and water-limited yield ($Y_w$) – for different agro-ecological regions in Zimbabwe and calculate yield gaps. Actual tuber yields ranged from 8 to 35 t ha$^{-1}$, while the simulated potential yield ranged from 88 to 96 t ha$^{-1}$ across the seven studied agro-ecological regions. Water-limited potential yield ($Y_w$) was simulated for four agro-ecological regions with different climates. A theoretical planting date (15th day of the month) was assumed in this simulation, and water-limited potential yields were compared to potential yields, revealing that $Y_w$ followed the same pattern as $Y_p$ but is dependent on the rainfall pattern in all the areas. The two simulated yields were very similar when planting took place during the rainy months of the year (September through January) while the crop completely failed ($Y_w = 0$ t ha$^{-1}$) in the dry winter plantings of April through July and rose with the summer rains thereafter. Irrigation as practiced is therefore essential for year-round potato production in Zimbabwe. Svubure et al. (2015) report yield gaps from 65 to 92% between simulated potential potato yield and actual yields reported by the growers, so there is a lot of room to increase yields, including measure to optimize planting dates and irrigation, use of IPM and improved/appropriate planting material. It seems that for potatoes, a number of well-developed crop simulation models have been developed, which can reliably predict potential yields; though they have not been recently used to investigate the impact of defined biotic constraints on yields, but
rather to get a general idea of theoretical yield gaps and potential yields in relation to solar radiation and water availability, particular in respect to climate change.

Yield gaps in cassava - boundary line analysis

Another RTB crop whose yield gaps have been looked at recently is cassava, in Uganda and Kenya. According to Fremont et al. (2009), the "ideal cassava plant" should have a fresh root yield of 75-90 t ha\(^{-1}\), while recorded survey yields ranged from 6.1 to 11.7 t ha\(^{-1}\) – translating to theoretical yield gaps of 84-93%. However, not only are the production conditions in Uganda and Kenya not necessarily ideal for cassava, the ideal cassava plants are also not necessarily available or planted, so the authors define the cassava yield gap as the gap between the actual and the attainable yield (\(y_{\text{max}}\)). The latter was defined as: "the maximum yield observed in a given agro-ecological zone with a given management intensity".

Actual yields were farmers' estimates of average cassava yield in their fields for past seasons. Attainable yields were defined as the greatest yields achieved in trial fields receiving "management level 3" (the use of improved crop establishment and genotypes) in two consecutive sets of on-farm cassava trials in Kenya and Uganda. Though this definition is straightforward, the authors did not simply subtract one yield from another, but used an adapted boundary line approach based on Webb (1972), van Asten et al. (2003) and Shatar and McBratney (2004) to identify in detail the contribution of individual abiotic, biotic and management factors to the yield gap.

This approach consisted in defining boundary lines that represented the maximum yield response (the dependent variable) to the various independent variables (e.g., rainfall), after having sorted the independent variables in ascending order and removed outliers. Boundary lines were fitted through selected boundary points (Schnug et al., 1996) following the model:

\[
y_i = \frac{y_{\text{max}}}{1 + (K \exp(-Rx))}
\]

Where \(y_i\) was the boundary line, \(y_{\text{max}}\) the observed attainable yield level at management level 3 (improved crop establishment and genotypes), \(x\) the independent variable and \(K\) and \(R\) constants. For each field and each independent variable, individual boundary lines were used to calculate the maximum cassava yield that could have been obtained if production would only have been limited by the independent variable in question (\(y_{\text{max,i}}\)). These individual boundary lines were then combined in order to create a multivariate model, assuming responses according to von Liebig's law of the minimum (von Liebig, 1863; Shatar and McBratney, 2004), and the model was used to predict yields for each field.

Finally, the yield gap caused by each independent variable in each field was determined by subtracting the \(y_{\text{max,i}}\) from the attainable cassava yield (\(y_{\text{max}}\)). Using this approach, they established that using their "complete management package" – consisting of improved crop establishment, an improved genotype and application of NPK fertilizer, average yields on farmer fields more than doubled, from ca. 9 to 21 t ha\(^{-1}\), and simulated attainable yields increased from ca. 18 to 37 t ha\(^{-1}\), in both the Kenyan and Ugandan sites. Cassava yield gaps therefore ranged from 43-76% and are caused by a multitude of production constraints. Fremont et al. (2009) conclude that abiotic constraints and related crop management practices are far more important than perceived by farmers and scientists and that efforts to improve productivity should therefore be reappraised and be geared towards combining approaches to identify and overcome the most important constraints simultaneously, rather than focusing on single constraints, and particularly on specific pests and diseases.

Yields of Musa spp.

To analyze yield gaps in banana and plantains, we must first come to a standard definition of yield. For annual and perennial crops with a distinct harvest period, defining yield is not usually a problem: yield is the mass of harvested product from a defined area after a single harvest (season) (e.g., Mt or kg ha\(^{-1}\)). For banana and plantain (genus Musa), which are harvested throughout the year, but whose production is nonetheless affected by a number of factors, including planting time, the predominant perennial nature of the crop, seasonal rainfall, cultivar type and crop cycle, defining yields is not straightforward. In their 2010 paper, Hauser and van Asten note that "a standardized approach on harvesting banana yields and a common and comparable way to express them is important". Unfortunately, their advice has not filtered into many studies, and many publications on the subject fail to clearly define Musa yields.

In addition to including a unit of time in the calculation of Musa yields, Hauser and van Asten (2010) suggest:

1. Weighing fresh bunches using a cultivar-specific peduncle cut-off point, and give two examples:
   (i) Plantains (Musa AAA): peduncle to be cut off halfway between the first and second empty bract above the first fruit hand.
   (ii) East African highland bananas (Musa AAA-EAHB): peduncle to be cut off where it enters the pseudostem at the point where the two youngest leaf petioles cross.
2. Converting fresh bunch weight data into edible dry matter (conversion calculations are cultivar/banana-type-dependent).
3. Clearly defining plot areas, borders, and plant densities (incl. reporting non-producing plants).
Yield gaps in East African highland bananas (EAHB) - boundary line analysis

Wairegi et al. (2010) also used the boundary line approach to distinguish and identify the importance of different biotic and abiotic factors affecting the yields of East African highland bananas. To do so, they monitored fresh banana yields (t ha\(^{-1}\) yr\(^{-1}\)), as well as biotic and abiotic constraints on 159 plots on on-farm demonstrations and ordinary farmer fields in three distinct banana-growing regions in Uganda. Farmer yields ranged from 10 to 20 t ha\(^{-1}\) yr\(^{-1}\) in the regions. In the model calculations, the maximum yield \(y_{\text{max}}\) (in Fremont et al., 2009) was replaced with the attainable yield \(y_{\text{att}}\), defined as “the highest yield observed in each region”, which ranged from 31 to 37 t ha\(^{-1}\) yr\(^{-1}\).

In addition, each boundary line function was used to predict the maximum yield possible \(y_{\text{att}}\) for each biophysical factor \(i=1, 2, \ldots, n\) in each plot. For each biophysical factor and each plot, the gap between \(y_{\text{att}}\) and \(y_{\text{ob}}\) was calculated. The yield gap was then expressed as percentage of \(y_{\text{att}}\) to allow for comparison among regions. For each plot, the minimum predicted yield \(y_{\text{min}} = \min (y_{x1}, y_{x2}, y_{xn})\) was also identified. Two types of yield gaps were presented: the explainable yield gap was defined as the difference between the maximum attainable yield \(y_{\text{att}}\) and the predicted minimum yield \(y_{\text{min}}\) while the unexplained yield gap represented the difference between the predicted minimum yield \(y_{\text{min}}\) and observed yield \(y_{\text{ob}}\).

Average yield gaps (expressed as percentage of \(y_{\text{att}}\)) corresponding to different factor by region combinations (factors: nematodes, weevils, weeds, mulch, fertilizer, N-total, K/(Ca + Mg), banana population, pH, SOM, N-total, clay and rainfall) were identified by plotting separate boundary lines for each factor and region. Depending on the region, there were only significant differences in average yield gaps between control and demonstration plots for the factors fertilizers (yield gap from 4.8 to 53.1%), weeds (4.6 to 36.1%) and mulch (10.4 to 28.0%). As the distribution of yield gaps varied among factors and regions, median and not average yields were used to graphically represent yield gaps associated with various factors for a typical plot in each region. The explained average yield gap ranged from 10.9 to 18.4 t ha\(^{-1}\) year\(^{-1}\) and the unexplained yield gap averaged 1.9 to 5.0 t ha\(^{-1}\) yr\(^{-1}\). Total yield gaps, that is, difference between attainable \(y_{\text{att}}\) and observed \(y_{\text{ob}}\) yields, ranged from 5.9 to 16.2 t ha\(^{-1}\) yr\(^{-1}\), when the attainable yield was assumed to be the highest yield observed in each region (31 to 37 t ha\(^{-1}\) yr\(^{-1}\)).

The authors conclude that the boundary line approach is more appropriate for data confined to a single agroecological zone and less appropriate for data covering a wide geographical region. They suggested that management decisions should not be only based on visual assessment of single constraints but based on a comprehensive understanding of causes of yield reduction.

Wairegi et al. (2010) identify soil fertility as a major constraint in all regions, while pests (weevils and nematodes) are more important in some regions than in others. They state that in general, the low banana yields observed in Uganda are due to abiotic factors (soil fertility and moisture) to a greater extent than biotic factors (pests and diseases), though these can vary greatly from plot to plot. They highlight that although yield gaps attributed to root necrosis had a median of 7.9% in one region, the outliers suggested that in some farms, yield loss attributed to nematodes could be as high as 80%. This suggests the need to target units [of analysis] smaller than regions when diagnosing constraints and subsequently developing recommendations” (Wairegi et al., 2010). The effect of diseases was not analyzed in this paper, as they did not seem to be a major problem in the monitored plots. Nonetheless, Wairegi et al. (2010) do recognize the potential yield limiting effects of diseases and suggest that banana systems be continuously monitored to prevent potentially devastating disease outbreaks.

Yield gaps associated with specific pests and diseases of banana and plantains

Nematodes

Substantial yield increases (20-75%) in production areas where nematicides were applied revealed the extent of production losses due to nematodes (Broadley, 1979; McSorley and Parrado, 1986; Sarah, 1989; Gowen, 1994). However, yield gaps were not defined in these studies, though various levels of infestation, as revealed both by assessing % of root necrosis, the relative proportion of dead to functional roots and counting the number of plant parasitic nematodes per 100 g root tissue have been associated with general negative effects on yields (Gold et al., 1994; Speijer and De Waele, 1997).

A particular research focus has been given to the effects of nematodes on commercial banana plantations in lowland tropical areas, where the most damaging nematode is the burrowing nematode, Radopholus similis. In these plantations, nematode populations are monitored on a monthly basis, and a nematode threshold of 10,000 R. similis/100 g functional root was established.
by Tartré and Pinochet (1981). This threshold was used to determine the need for nematicide applications on plantations in Latin America (Chávez and Araya, 2001).

Nonetheless, despite years of research on plant parasitic nematodes that attack Musa spp., a direct relationship between the number of nematodes in roots or the % of necrotic roots and certain levels of yield losses has not been established for any particular banana cultivar nor for any particular nematode species. In addition to the multitude of different cultivars and species that would come into play, not to mention pathotypes of certain nematodes with variable levels of aggressiveness towards banana plants, levels of nematode resistance and tolerance to nematodes also differ from cultivar to cultivar. Environmental factors such as flooding, drought, and particularly temperature in relation to altitude, as well as plant density and agronomic practices, such as mulching, can also affect the health of banana plants in general, and subsequently their tolerance of and resistance to nematodes (Gaidashova et al., 2009).

Since even establishing a clear relationship between the presence and density of nematodes in roots and levels of root necrosis has at times been difficult, with positive plant growth conditions at times masking the negative effects of nematodes (Gaidashova et al., 2009), establishing a clear relationship between levels of yield loss and nematodes has never been a research priority and no such calculations have been published. While it is not disputed that nematodes cause both a decline in plant health and productivity – with weaker plants and root systems, longer production cycles and smaller fruit bunches all being cited as effects of nematode infestations, no clear equations have been published or proposed to quantify this relationship.

**Insect pests**

Among the various insects that directly attack Musa spp., the Banana Weevil, *Cosmopolites sordidus*, is most destructive, particularly in banana producing areas of Africa. This insect has been intensively studied over the years, and many attempts have been made to relate levels of field infestation and corm damage to yield losses, particularly by the group of researchers working with Clifford S. Gold (Gold et al., 1994; Rukazambua et al., 1998; Kiggundu et al., 2003; Gold et al., 2004; Gold et al., 2005). Among the yield losses the group associates with *C. sordidus* damage are mat disappearance, plant loss and reductions in bunch weight.

Rukazambua et al. (1998) calculated yield losses in East African Highland Bananas (*Musa* AAA-EAHB cv. ‘Atwala’) to *C. sordidus* in field trials in Uganda. They concluded that the banana weevil is the leading cause of banana decline and even the disappearance of banana from parts of central Uganda. After having to slightly modify their experimental design (weevils invaded designated “weevil exclusion zones”), they determined that damage to the central cylinder (CC) (that is, inner section of a root, vascular cylinder including xylem and phloem bundles) had the most important effect on yields, and that high levels of damage in the plant crop will have negative effects on ratoon plants (smaller plants and bunches – that is, no recovery observed). Yield losses increased from 5% in the plant crop to 44% in the third ratoon, with yield losses attributed to both high levels of plant loss (up to 29% in the 3rd ratoon) and reductions in bunch weight. In another study by the same group, Gold et al. (2004) found that yield loss averaged 42% during the last four years of a 7-year yield loss trial.

In their review of assessment methods for evaluating damage *C. sordidus* on EAHB, Gold et al. (2005) concluded that damage estimates on the corm periphery are not useful parameters for assessing pest status of *C. sordidus*. They go on to state that internal damage revealed in cross sections of the central cylinder is the best of the available damage predictors for assessing *C. sordidus* pest status, though even this method was not very reliable for assessing yield losses, as the relationships between the different damage parameters, plant size and bunch weight were weak.

Nonetheless, Rukazambua et al. (1998) estimated yield losses by comparing plants with high levels of *C. sordidus* damage with the controls (that is, plants with little or no weevil damage), using data collected from four generations of bananas (plant crop and 3 ratoon crops). Damage to the CC was assessed for each plant after harvest and plants of each crop cycle were grouped into damage categories. Yield losses were calculated by taking into account not only actual bunch weights, but also lost plants.

Though not explicitly named expected yield (*Y*\textsubscript{exp}) in Rukazambua et al. (1998), an expected yield was calculated by establishing the average bunch weight of plants in the lowest weevil damage category (0-5% weevil damage to CC) and multiplying this by the number of plants initially planted on the plot, to end up with an average expected yield of X kg ha\(^{-1}\) (e.g., 500 plants ha\(^{-1}\) x 10 kg bunch\(^{-1}\) = expected yield of 5000 kg ha\(^{-1}\)).

Average bunch weights for each damage category were then also calculated. Yield losses were calculated by subtracting the actual yield (*Y*\textsubscript{act}; product of the number of plants in each damage category multiplied by the average bunch weight in that category) from the expected yield (*Y*\textsubscript{exp}). For example: 500 plants are planted on a 1 ha plot. At the end of the harvest of the plant crop:

(i) 100 plants: low damage and 10 kg bunch\(^{-1}\)
(ii) 100 plants: moderate damage and 8 kg bunch\(^{-1}\)
(iii) 100 plants: heavy damage and 6 kg bunch\(^{-1}\)
(iv) 100 plants: very heavy damage and 4 kg bunch\(^{-1}\)
(v) 100 plants: dead / failed to produce a bunch

The expected yield is 500 plants ha\(^{-1}\) x 10 kg bunch\(^{-1}\) = 5000 kg ha\(^{-1}\), while the actual yield is:
(100×10) + (100 × 8) + (100 × 6) + (100×4) + (100 × 0)  
= 1000 + 800 + 600 + 400 + 0  
= 2800 kg ha⁻¹

Yield gap is therefore:  \[ Y_{exp} - Y_{act} = 5000 - 2800 = 2200 \text{ kg ha}^{-1}\] or 44%.

Though this method allows researchers to estimate yield losses due to banana weevil attack, the exhaustive monitoring necessary to calculate yield losses in this manner is not particularly adoptable to assessing yield losses or gaps under survey conditions, over large areas, and on fields with a patchy planting history.

**Identifying yield gaps in Musa spp. using models to simulate yields**

Haverkort and Struik (2015) defined the potential yield of potato as the theoretical yield that can be calculated or modelled for a certain cultivar grown in a certain environment without any limiting or reducing factor being present. In order to use simulated yields to calculate yield gaps for bananas, similar models would have to be developed and validated for different banana cultivars, however, as noted by Tixier at al. (2004):

"Banana crops represent a collection of individual plants that vegetatively propagate at their own rhythm, with stabilised but unsynchronised production of inflorescences over time. Such agrosystems cannot be simulated with existing crop models due to the unsynchronized behavior of individual plants."

Due to this problem, Tixier at al. (2004) first developed a model (SIMBA-POP) that can be used to predict harvest periods over cropping cycles and the dates of harvest peaks \( (R^2 = 0.99) \) and to simulate or compare new population management decision criteria, based on the cohort population concept, but not actual yields. They calibrated and validated the SIMBA-POP model with field data from the French West Indies (Guadeloupe and Martinique) for Musa AAA group, cv. Cavendish Grande Naine. The same team of researchers later developed various sub-modules for the SIMBA model to simulate additional variables, such as yield in Mg ha⁻¹ y⁻¹ (SIMBA-GROW; \( R^2 = 0.55 \); Tixier et al., 2008), weekly nematode populations (SIMBA-NEM; Tixier et al. 2006), weekly water potential (SIMBA-WAT; Tixier, unpublished data), pesticide risk rank \( (R_{pest} ; R^2 = 0.92; \) Tixier et al., 2007) and nitrogen dynamics (SIMBA-N; Dorel et al., 2008). SIMBA and all its modules run at a weekly time-step at the field scale. Unfortunately, no actual yields are presented in the paper, nor are they compared to actual yields of commercial plantations. While no yield gaps were calculated, such a model could help in identifying particular contributors to yield gaps in specific environments.

Chaves et al. (2009) published an article entitled “Modeling plantain (Musa AAB Simmonds) potential yield”, but unfortunately did not present any yield data in their paper, except to say that the average plantain yields in traditional mixed systems are about 10 t ha⁻¹. They concentrated on modelling the yields of the plant and first ratoon crops, but state that their model could be applied to previously existing plantations or to second or third production cycles; the complexity they imply is out of the scope of their work. However, they do note that the design of their model permits the incorporation of water, nutrient and pest limitations, though none of these are included in the presented model.

Due to the problematic nature of Musa crops and their often poorly defined yields and disparate harvest times, modelling these crops is particularly challenging, and cultivar specific crop models for bananas have not yet been developed. This may be in part because a lot of information is available for some cultivars (Cavendish group), while much less is available on others. However, there should be enough information already available in the banana research community to populate a model which should also include data on climate and soil (abiotic conditions, if not necessarily constraints). These data could be fed into one of the already existing models to help make the models more versatile and adaptable to different cultivars and agroecologies.

**How best to calculate yield gaps**

In their review, van Ittersum et al. (2013) emphasize the need for accurate agronomic and current yield data together with calibrated and validated crop models and up-scaling methods to larger geographical units. The protocol, including the effects on Yg of uncertainties in weather, soil, cropping system management and crop growth simulation models, remains to be tested and refined, a process which is currently undertaken in the Global Yield Gap Atlas project (www.yieldgap.org) (van Ittersum et al., 2013).

In 2012, the first phase of the Global Yield Gap and Water Productivity Atlas (Global Yield Gap and Water Productivity Atlas, 2020; www.yieldgap.org) was developed and continues to expand. This platform “provides robust estimates of untapped crop production potential on existing farmland based on current climate and available soil and water resources”, and proposes a “standard protocol for assessing yield potential (Yp), water-limited yield potential (Yw), yield gaps (Yg) and water productivity (WP)”.

For fully irrigated crops, the yield potential is defined as the yield of a crop cultivar when grown with water and nutrients non-limiting and biotic stress effectively controlled. Therefore, crop growth is determined by solar radiation, temperature, atmospheric CO₂ concentration, and genetic characteristics, but, in theory, not dependent on soil characteristics. For rain-fed crops, water-limited yield potential (Yw) is similar to Yp, but crop growth is also limited by water supply, and hence influenced by soil...
type and field topography. It is the most relevant benchmark for rain-fed crops.

The protocol, which relies on the collaboration of agronomists with knowledge of production systems, soils, and climate governing crop performance in their countries, is applied for all crops and countries based on best available data, robust crop simulation models, and a bottom-up approach to upscale results from location to region and country.

Though the first phase of the project (2012-2015) focused on cereal crops, the crop list has extended to soybean, sugarcane and potato, and the GYGA aspires for global coverage of yield gaps for all major food crops and countries that produce them. To date, data on potato are only available for Jordan and Tunisia, which reveal a 32 and 34% yield gap in irrigated potato, respectively. Both potential and actual irrigated potato yields are higher in Jordan (47.5 and 32.2 t ha\(^{-1}\)) than in Tunisia (27.8 and 18.3 t ha\(^{-1}\)), respectively. All data on yield gaps, including information of data collection site (station), climate zone and crop and type of production system (rainfed or irrigated) can be downloaded from the site (www.yieldgap.org). Data on potential yields, actual yields and resulting yield gaps are presented for each station and averaged for each country.

This type of international cooperation is very encouraging, though the actual data that have been made available, particularly for RTB crops are still very scarce. Crop models based on solid data appear to be the most promising means of calculating attainable yields for RTB crops. This type of information can be collected from breeding programs, where they exist, and such data are made available. However, the necessary agronomic data are not available for most cultivars – especially landraces – which are not necessarily included in formal breeding programs. For these cultivars, data from specific field trials or surveys will be required to calculate potential yields.

Once potential yields can be estimated, actual yields can be collected, along with supporting data to assess individual contributing factors to yield gaps. Boundary line analyses appear very interesting in this regards, as the author who used this method (Fremont et al., 2009; Wairere et al., 2010) on RTB crops seemed to have been able to generate a lot of information out of their results. However, the method in not very intuitive to follow, and requires a lot of computing and statistical skills to correctly interpret the results.

Nonetheless, before starting fieldwork to assess yield gaps and identify various contributing factors, it is necessary to clearly define the yield – simply weighing bunches is not sufficient, irrespective of the number of bunches collected, as a yield is calculated for an area, and, in the case of perennial crops like Musa spp., also for a specific amount of time. Ideally, a comparison should be made between plants of the same cultivar, of a similar age and development stage (plant crop, ratoon), under control and farmer field conditions – and all this over a certain harvest period and for a clearly defined area.

Bunch weight is not equal to yield. Particularly over time, when individual banana mats die, yields on a plot can drastically decrease, while individual bunch weights may increase (e.g., due to less competition between surviving banana plants for limited resources such as water, etc.). Nonetheless, a basis for comparison has to be made, though this depends on the aims of the work.

To gain a general idea of what portion of yields is lost to particular diseases and pests within a certain area or at specific altitude bands, first potential yields have to be estimated. This could be done either by using a model to estimate potential yields (where available) or identifying “ideal plots”, where the plants are under no biotic stress, and abiotic stresses (water and nutrient availability, etc.) are taken into account. After surveying and collecting data on actual yields within those same areas, yield losses can be calculated – but strict monitoring is necessary to allow for this, as assessing yields for bananas cannot be done with a single farm visit.

As the relationships between nematode and weevil damage and yield losses are unfortunately weak, estimating yield gaps attributable to these pests based on visual assessments (% damage to CC or roots) is not as straightforward as hoped for. Nonetheless, these methods can give an indication of what factors are contributing to yield losses, and actual yield losses observed can then be “traced back” to these pests. This study could in fact actually help to better define the portion of yield loss that can be attributed to the different diseases and pests, and identify tolerant and or resistant cultivars, and agronomic practices that negate the detrimental effects of these biotic stresses – all depending on the type and quality of data collected.

Yield gaps due to banana diseases: a case study in western Burundi using yield gap assessment methods discussed in the literature review sections

This part of the paper describes a concrete field study that assessed yield levels and gaps due to selected biotic (pest and disease) constraints in a banana production landscape in Western Burundi.

MATERIALS AND METHODS

Small-scale farmers in western Burundi grow mixtures of various banana cultivars, with the dominant cultivars being ‘Igisahira’ and ‘Indarama’ (both East African Highland Banana (EAH); cooking type; Musa AAA genome), ’Km5’ a AAA genome dessert banana, ‘Pisang Awak’ (ABB, beer) and to a lesser extent ’FHIA 17’ (a tetraploid AAAA hybrid, multipurpose). Bunch weight data were collected on apparently healthy plants (that is, plants that showed no aboveground disease symptoms) in over 270 farms (with each farm having at least 50 banana mats) across 3 altitude bands (800-1,200 m asl, 1,201-1,600 m, 1,601-2,200 m). Zoning by altitude was
done to reduce within zone differences between farms. At least 10 disease-free ‘Igisahira’ ‘Indarama’, ‘Km5’ and ‘FHIA 17’ plants were assessed across various fields per altitude band. In addition, for at least 50 apparently healthy plants per cultivar, nematode root necrosis % (using the method described by Speijer and De Waele, 1997) and weevil larvae corm damage (Gold et al., 1994) were assessed. Xanthomonas wilt of banana (XW), banana bunchy top disease (BBTD) and Fusarium wilt plant incidences (that is, number of infected plants over total number of plants in an assessed field) were assessed at 35 farms per disease and per altitude band. These three banana diseases cause a total loss of the affected plants and yield gaps due to them were postulated to be proportional to plant disease incidence levels. Yield gaps due to banana pests were calculated using the boundary line analysis. The boundary line analysis procedures and concepts described by Fremont et al. (2009) and Wairegi et al. (2010) and as described in the sections above were used with a slight modification as described in the steps as follows:

(i) The most influential yield outliers were identified and dropped with the help of boxplots.

(ii) Pearson correlation analysis was then used to identify the relationship between the bunch weight and the biophysical constraints (that is, corm weevil damage and nematode root necrosis) for the different banana cultivars. Constraints either had a positive or negative influence on the bunch weight, and bunch weights were sorted, respectively, in an ascending or descending order.

(iii) This was followed by determination of the yield due to each biophysical factor ($Y_b$). $Y_b$ is the maximum yield predicted by the boundary line given the biophysical constraint.

(iv) Boundary lines were then fitted on the graphs, assuming a nonlinear relationship between bunch weight due to a factor and the corresponding factor values. This was followed by adding the boundary line obtained from the predicted yield due to each corresponding biophysical factor in the model.

The yield gap proportions were calculated as the difference between the attainable yield ($Y_a$) and $Y_b$ (Wairegi et al., 2010). $Y_a$ was the highest bunch weight measured on farmers’ fields. Where yield gap data was skewed (that is, high kurtosis > 1), median was presented instead of the mean. The statistical analyses and data visualization were carried out using R-statistical software (R Core Team, 2018) and the ggplot2 package (Wickham, 2016).

### RESULTS AND DISCUSSION

Plant incidence values for Xanthomonas wilt of banana (XW), banana bunchy top disease (BBTD) and Fusarium wilt were low and below 0.2 % at all altitude bands (Table 1). It has to be noted that plant incidence levels presented in Table 1 reflect levels at one point in time and provide a quick impression of pathogen impact. Cumulative plant incidence values over a year or years or crop cycle would be far higher. For example, large numbers of plants affected by XW had been removed in the years, months and weeks prior to the survey. Hence, percentage XW plant incidence as measured during the surveys does not reflect the overall severe damage done over time by this pathogen. A susceptible plant affected by XW, BBTD or Fusarium wilt does not yield an edible bunch. As such any plant affected by any of the three diseases can be discarded from yield calculations, and thus the incidence of these diseases can be equated to percentage yield loss.

Gold et al. (1999) reported nematode and weevil numbers and damage to build up gradually, with often cross-generation effects, making it difficult to link yield level with corm damage levels for a specific plant in a perennial banana mat. Wairegi et al. (2010) however, using the boundary line method, has been able to compute yield gaps due to damage caused by nematodes and weevils.

In the current study, nematode root damage effects were most profound on the yield of ‘FHIA17’ and moderate for the east African highland cultivars ‘Indarama’ and ‘Igisahira’ (Figure 1). ‘FHIA17’ is susceptible to the burrowing nematode R. similis (Viaene et al., 1997; Moens et al., 2005), while all highland banana types have been reported to be susceptible to nematodes (Speijer et al., 1999; Ssango et al., 2004). No nematode effects were in contrast observed on the yield of ‘Km5’ which is a robust dessert banana (Musa AAA).

Weevil damage was generally low across all the four cultivars sampled. Yields of the East African highland bananas (EAHB) declined more severely with an increasing level of weevil corm damage. The EAHB are highly susceptible to banana weevils, with more severe losses reported at the low humid altitude zones of East and Central Africa (Gold et al., 1999; Kiggundu et al., 2018).

### Table 1. Plant disease incidence (%) for Fusarium wilt, Xanthomonas wilt and banana bunchy top disease by altitude band.

<table>
<thead>
<tr>
<th>Altitude band (masl*)</th>
<th>Fusarium wilt</th>
<th>Xanthomonas wilt</th>
<th>Banana bunchy top disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>800-1,200</td>
<td>0.012</td>
<td>0.022</td>
<td>0.027</td>
</tr>
<tr>
<td>1,201-1,600</td>
<td>0.018</td>
<td>0.021</td>
<td>0.027</td>
</tr>
<tr>
<td>1,601-2,200</td>
<td>0.003</td>
<td>0.053</td>
<td>0.026</td>
</tr>
<tr>
<td>Fpr</td>
<td>0.032</td>
<td>0.054</td>
<td>0.989</td>
</tr>
<tr>
<td>Lsd</td>
<td>0.008</td>
<td>0.024</td>
<td>0.020</td>
</tr>
<tr>
<td>Cv%</td>
<td>103</td>
<td>110</td>
<td>104</td>
</tr>
</tbody>
</table>

*meter above sea level. Plant disease incidence was calculated as the number of diseased plants over the total number of plants in an assessed field.
Figure 1. Relationships between banana yield and both root necrosis percentage and corm weevil damage percentage. The dotted lines represent the boundary lines, obtained through a boundary line analysis, whereas the points represent the observed bunch weights of assessed plants. Graphs are presented for four banana cultivars (FHIA17, Km5, Indarama and Igisahira) that dominate at the studied field sites.

Weevils had minor to moderate effects on the performance of, respectively, ‘Km5’ and ‘FHIA17’. These two cultivars have been reported to be tolerant to banana weevils (Nowakunda et al., 2000; Kiggundu et al., 2003).

The percentage mean yield gaps due to nematode root necrosis were high in ‘Indarama’ (16.3%), ‘FHIA 17’ (12.7%) and ‘Igisahira’ (11.6%) while low in ‘Km5’ (1.8%) (Figure 2). Yield gaps due to the banana corm weevil damage were high for ‘FHIA17’ (10.4%), ‘Igisahira’ (6.7%) and low in ‘Indarama’ (3.3%) and ‘Km5’ (1.2%). Wairegi et al. (2010) also reported significant yield gaps due to nematodes and weevil larva, but only focused on east African highland banana cultivars grown in farmer’s fields. Pest constraints (nematodes and weevils) were reported to be particularly important in Central Uganda (1,100-1,300 masl), but not in South- and South-western
Figure 2. Percentage yield gaps due to nematode root necrosis and corn weevil damage, expressed as the percentage of the attainable bunch weights. The solid bold lines across boxes are medians. The red points represent the mean percentage bunch weight loss by each factor. The boxes represent the interquartile range (25–75th percentile), circles outside the central box represent outliers by between 1.5 and 3 times the interquartile range, while bars represent the smallest and largest observations which are not outliers. Graphs are presented for four banana cultivars (FHIA17, Km5, Indarama and Igisahira) that dominate at the studied field sites.

Uganda which has a higher altitude (1,300-1,800 masl). In Central Uganda, a yield gap median of 10.3 and 5.5% was, respectively, found for nematode and weevil damage effect (Wairegi et al., 2010).

The “FHIA17” plants had the highest bunch weights compared to the highland AAA-EAH type bananas and ‘Km5’ (Figures 1 and 3). Bunch weights collected from healthy looking plants varied greatly for the 4 assessed cultivars, and at all altitudes (Figure 3). As these plants were all disease-free, it can be assumed that the yield variation is mainly due to variations in abiotic factors (e.g., soil fertility, level of input use, altitude or temperature effects) and possibly due to inter-plant competition within mats.

A large variation in bunch weights and thus yield gaps for bunches harvested from disease-free plants for all four cultivars was observed across the three altitude bands combined (Figure 4). The boundary lines show severe declines in yields at the upper altitudes, with more profound effects observed in ‘FHIA17’ and ‘Indarama’ (Figure 3). This is supported by field observations that show increasingly poor banana plant growth and yield at altitudes above 1900-2,000 masl (Rubaihayo and Gold, 1993; Karamura et al., 1998; Sivirihauma et al., 2016). Irrespective of the altitude bands, higher yield gaps were observed for the ‘Igisahira’, ‘Km5’, ‘FHIA17’ and least for ‘Indarama’ plants with gaps of 26.9, 22.8, 22.8 and 8.9%, respectively (Figure 4). Interplant competition often results in suboptimal bunch sizes/weights. The presence in the study sites of small and large mats with varying levels of interplant competition might have contributed to the observed wide range in bunch weights for the various cultivars. For ‘Km5’ specifically, plants tended to produce a lot of suckers/lateral shoots thus increasing competition for resources, ultimately resulting in smaller bunches.

**CONCLUSION AND RECOMMENDATIONS**

Banana bunch harvests and weights vary significantly at production zones in western Burundi due to a multitude of biotic constraints. The often sub-optimal, medium and small bunch sizes found on visibly healthy mats indicate that significant improvements in bunch weights could also be attained through the application of agronomic/field management practices that enhance soil fertility, soil moisture content and/or soil health. Efforts to improve productivity should be reappraised and be geared towards combining approaches to identify and overcome the most important constraints simultaneously, covering both agronomy/enhanced field/soil management and pests and diseases control. Various integrated pest
management/disease control options are available ranging from tolerant or resistant germplasm to cultural control options. In addition, effective Integrated Soil Fertility Management/agronomic options include intercropping with N-fixing legumes, application of mulch, integration of field boundary bands of fast-growing shrubs/grasses and the integration of small ruminants. Simple and robust methods for estimating yield gaps due to pests and diseases, and abiotic constraints can inform/guide on the need to apply agronomic and/or
disease control efforts. Rigorous knowledge communication of gained insights to extension agencies and farmer communities should also be high on the research for development agenda. In the current study, the assessment of the yield gaps due to diseases was challenging as farmers continuously uproot or cut down infected plants. Field monitoring, with the involvement of farmers, over time will be necessary for a more accurate assessment of yield gaps due to diseases.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

ACKNOWLEDGMENTS

The authors are grateful for the support of our project partner ISABU-Burundi for data collection. This research was funded by the CGIAR Research Program on Roots, Tubers and Bananas (RTB) and supported by CGIAR Fund Donors. http://www.cgiar.org/about-us/our-funders/.

REFERENCES


Khan MS, Yin X, van der Putten PEL, Struijk PC (2014). An ecophysiological model analysis of yield differences within a set of
contrasting cultivars and an F1 segregating population of potato (Solanum tuberosum L.) grown under diverse environments. Ecological modelling 290:146-154.


