Digital agriculture to fulfil the shortage of horticultural data and achieve food security in Sub-Saharan Africa

J. Sarron[1,2], D. Beillouin[1,2], J. Huat[1,2,3], J.M. Koffi[4,5], J. Diatta[6], É. Malézieux[1,2], E. Faye[1,2]
[1] CIRAD, UPR HortSys, F-34398 Montpellier, France; [2]HortSys, Univ Montpellier, CIRAD, Montpellier, France; [3]CIRAD, UPR HortSys, F-97455 Saint-Pierre, Réunion, France; [4]Centre National de Recherche Agronomique (CNRA), Direction Régionale de Korhogo/Station de Recherche de Lataha, BP 856 Korhogo, Côte d'Ivoire ; [5]Université Jean Lorougnon Guédé (UJLoG), Daloa, P.O. Box 150, Côte d'Ivoire ; [6]ISRA, Centre pour le Développement de l'Horticulture, Dakar, Sénégal

Abstract

In Sub-Saharan Africa (SSA), agricultural data are often scarce, irregular, unreliable, and inaccessible. Agricultural data, when existing, relate almost exclusively to yields and cultivated areas at the national level and SSA suffers a significant lack of information at sub-national levels describing the agricultural systems (area, production quality and quantity, management practices, agrobiodiversity). This study focused on horticultural crops (fruits and vegetables) for which the situation of data is worse than for annual staple crops because: i) fewer studies are carried out on horticultural crops; and ii) the complexity of horticultural systems hampering the development of accurate assessment methods. Despite the importance of fruits and vegetables for food security, it is acknowledged that the lack of horticultural data in SSA affects all stakeholders of the value chains - from farmers to the national and international institutions. We focused on digital agriculture, which relies on technologies, as one of the pathways for data improvement in SSA by supporting data acquisition, standardisation, availability, and sharing between stakeholders. For instance, data acquired by participative monitoring using smartphones might increase the availability of accurate information on yields at the field scale. Other examples have shown that remotely sensed data from drones or satellites allow to accurately map and monitor horticultural systems, including agroforests. By overcoming the lack of data, digital agriculture can help to better characterise and evaluate horticultural systems to improve food security. The robustness of these tools and their adoption by farmers are discussed to upscale digital agriculture to address food security issues in SSA.

Keywords: digital technologies, remote sensing, data collection, smartphone applications

INTRODUCTION

The major challenge of agriculture in Sub-Saharan Africa (SSA) is to meet the growing demand for food in the context of climate change and increasing anthropogenic pressures. At the same time, a large significant part of the population faces undernutrition (23% of the population), micronutrient deficiencies, or overweight (FAO, 2020). Smallholders (< 2 ha of land) represent 80% of farmers in SSA and remain mostly poor and food insecure due to low investment capacity, low labour productivity, poor market access, and high crop yield gaps (FAO, 2020).

As essential sources of nutrients (fibres, vitamins, and minerals), horticultural crops contribute to a healthier diet, improve population's food security, and prevent chronic diseases (Gómez et al., 2013). From the farmer's perspective, horticultural crops allow to diversify the sources of income, improve the resilience to extreme weather conditions and price volatility, and thus are often seen as an opportunity to favour farms' economic growth (Paut et al., 2019). In SSA, horticulture occurs under various cropping systems, some of which can be highly diversified and complex (e.g. agroforestry). Indeed, smallholders mix several cultivated and natural species increasing the variability across the plot and along the year (different growing cycles) on a unit of area (Sarron et al., 2018). However, despite the importance of fruits and vegetables for food security, there is a lack of institutional policies supporting and developing the horticultural sector in SSA (Weinberger and Lumpkin, 2007).

Agricultural statistics are crucial for all stakeholders to have explicit knowledge and extensive information on the sector. For national and international institutions (agencies, governments, and NGOs), agricultural data are required to monitor development policies and evaluate outcomes (Desiere et al., 2016). For researchers, data are needed to measure the performances of agrosystems and quantify the limiting factors that impact these performances (Carletto et al., 2015). Farmers can also benefit from accessible and accurate agricultural information, especially for decision-making (World Bank, 2019).

Nevertheless, agricultural data are often scarce, unreliable, or inaccessible in SSA (Carletto et al., 2015; Desiere et al., 2016; World Bank, 2010). When they exist, data relate almost exclusively to surveyed yields and cultivated areas at the national level (e.g., FAOSTAT). At sub-national levels (district, landscape, and field levels), there is a significant lack of information describing the agricultural systems – area, production quality and quantity, crop management, agrobiodiversity - and their surrounding environment - soil and climate (Carletto et al., 2015). Most data are also irregular because they depend on short-term and punctual investments or projects, which in turn hampers the collection of long-time series data. Horticultural data face a worsened situation as there is a lack of accurate methods to characterise horticultural systems, especially the complex ones.

Over the last few decades, digital agriculture has become increasingly present among farmers and other stakeholders (World Bank, 2019). Digital agriculture can be defined by the use of technologies - such as sensors, machine learning algorithms, or Information and Communications Technologies (ICTs) - within the agriculture value chain. Digital agriculture not only focuses on acquiring data but also on sharing and reusing them in open access. Open data represent an opportunity for the horticultural sector and food security in SSA by facilitating access to information for the greatest number of stakeholders.

In this paper, we argue that digital agriculture is one pathway for better data reliability and accessibility contributing to achieving food security in SSA. We first provide an overview of the current status of horticultural data in SSA and point several factors of data degradation and their effect on the development of sustainable horticultural systems. In the second part, we detail how digital agriculture can contribute to fulfilling the shortage of horticultural data. We also provide some concrete examples of the application of digital agriculture within horticultural systems in SSA.

CURRENT STATUS OF HORTICULTURAL DATA IN SUB-SAHARAN AFRICA

Open access databases

Table 1 provides a non-exhaustive list of open access databases that are freely available and related to horticultural systems (e.g., farm structure, management practices, and yields) and crop species or cultivars in SSA. The ten databases mentioned in Table 1 are the ones that provide observational, survey, participatory, or simulated data. Although useful for researchers, databases on experiments are not included because they often relate to specific conditions. In addition, databases from long-term observation sites are not presented in this study due to their low spatial coverage.

Eight out of the ten databases contain variables related to horticultural cropping systems such as harvested areas, management practices, production, and yields. Five of them are exclusively based on surveys at national, sub-national or field levels. Nevertheless, the three databases that are finer in scale than the country (CountryStat, AgroMAPS, LSMS) are limited to a few countries (Table 1). Three databases (Data Africa, GAEZ, EarthStat) provide geospatial data on the cropping systems. However, these databases are based on FAOSTAT national data disaggregated using downscaling models (You et al., 2009).

Two other databases do not relate directly to the cropping system but instead crop species and cultivars in SSA (Table 1). These databases enable to gather information and observational knowledge such as cultural practices, plant descriptions, or plant genetics in the form of technical sheets of links to other specific information sources. The ASS database is a repository of several databases worldwide that permits the mutualization of information on plants useful in agroforestry (Kindt et al., 2019).

Limitations on SSA horticultural data

Although there are a certain number of available databases related to the horticultural sector in SSA (Table 1), these data faced many limitations that concern: i) the scarcity and irregularity of the data; ii) their reliability and quality; and iii) their accessibility. On the one hand, observation and survey data are available only at the national level and can be subject to numerous measurements or reporting biases. On the other hand, geospatial data are mostly simulated and concern only environment and land cover variables.

1. Data scarcity and irregularity

In SSA, farm surveys remain widely used to collect agricultural statistics (Carletto et al., 2015). In the field, data collection poses numerous constraints in terms of time and cost (for labour and compensation for the farmer). The yield and crop management monitoring tasks are even harder to achieve in complex horticultural systems that mix several species spread in the plot and throughout the year. Moreover, there is a lack of resources allocated in data collection by most countries (World Bank, 2010). Inadequate financial and human resources lead to public statistical agencies offices that rely on external donor funding, which are less interested in horticultural crops (Desiere et al., 2016; Weinberger and Lumpkin, 2007). Thus, manual data collection remains limited in terms of temporal frequency and spatial coverage. In addition, the variables considered, the measurement methods and the frequency of data collection vary according to the projects or the monitoring systems (Carletto et al., 2015). There is insufficient coordination between statistical agencies at the regional level, which results in the lack of harmonised and integrated data sources (World Bank, 2010). Consequently, Africa is the continent with the lowest response rates to FAOSTAT questionnaires, with only 26% of questionnaires filled compared to 51% for the world average (FAO, 2019).

2. Data reliability and quality

Observation and survey data can be subject to several measurement errors and biases. Several reasons can explain the low data reliability in SSA. Firstly, data quality can be affected by recall errors. When farmers are surveyed, they are asked to recall the details of past events, measures, or management practices (Carletto et al., 2015). Most smallholders in SSA do not record their practices and harvests or use eye-estimation (Desiere et al., 2016). Consequently, they provide information on their practices with human-induced biases that are complex to assess. In addition, yield measurement varies greatly following the type of crops (e.g., fruits, vegetable, or leaves) and the tool and means used (e.g., scale, bag, or basket). Secondly, the monitoring methodology can also conduct to numerous errors. It is especially the case for sample-based measurements such as agrobiodiversity, disease, or yield estimation (Carletto et al., 2015). Sampling strategies most often do not consider the spatial variability at the scale of interest (field, farm, or region) and consequently might misestimate the measurement. Thirdly, some studies highlighted that the importance of agricultural data to implement and support policies tends to push some administrations to misreport data (Desiere et al., 2016). According to Desiere et al. (2016), some officials may voluntarily over-estimate agricultural production to demonstrate that their reforms are working.

3. Data accessibility and reuse

In SSA, much of the data collected during public research or development projects are not accessible, hindering knowledge generation as a common good. The investment in time and money for data acquisition is often a barrier to sharing them, as data producers do not want to share them easily. In addition, open data repositories of sub-Saharan public research institutes are still largely absent. Nevertheless, some international agencies have set up their own open data repositories (e.g., gardian.bigdata.cgiar.org).

Impact of the lack of data for horticultural research and development

Lack of data has multiple impacts on the horticultural sector in SSA. At the field or farm scales, study and characterisation of the cropping systems are hampered by the limited access to reliable information on crop yields and management practices. For instance, the difficulty in obtaining data on pesticide uses in fruit and vegetable production hampers the evaluation of pesticides risks in the region (de Bon et al., 2014). The lack of quantitative data also slows down the evaluation of factors that determine yields (Karst et al., 2020). Comparative analysis of the different types of agriculture (e.g., extensive vs intensive) and the promotion of sustainable practices adapted to the conditions of SSA is then limited, especially for smallholders (Sarron et al., 2018). Finally, the lack of freely available and accurate data with a large spatial coverage and regular revisiting time-step on agrosystems restrains the making of appropriate policies for the horticultural sector in SSA (Desiere et al., 2016).

DIGITAL AGRICULTURE TO FULFILL THE SHORTAGE OF HORTICULTURAL DATA

One major factor contributing to the degradation of horticultural data in SSA is the lack of adapted assessment methods, including standard methods for measuring agricultural outputs. A solution would be to develop reliable and low-cost measurement tools based on digital agriculture, which can be a powerful lever for i) reliable and accurate data acquisition; ii) trustworthy data reporting; and iii) safe shared and open access data.

Characterisation of horticultural systems using remote sensing

Digital agriculture makes use of remote sensing in a broad sense: the acquisition and management of data from a distance, whether on the ground, in the air, or space. Remote sensing has been long recognised by researchers for its potential to acquire large amounts of multi-scale and multi-temporal data. Remote sensing has been used for many applications in horticulture, such as land cover mapping, crop production estimation, or plant health monitoring (Usha and Singh, 2013).

Satellite imagery is probably the most commonly used remote sensing technology in agriculture and can provide geospatial information at various scales (Karst et al., 2020). Although satellite imagery has been mainly used for annual crops in SSA, few studies used this technology in the context of horticultural systems (Karst et al., 2020). The farm structure that is mostly small and complex and can concentrate numerous crop species partly explained this situation. In addition, the acquisition of satellite images relies on cloud-free conditions that are difficult to satisfy during the rainy season, thus limiting year-round monitoring of horticultural systems. Finally, the high cost of high-resolution images can limit their use for land cover mapping to characterise large areas with high temporal frequency.

Geospatial information can also be assessed by unmanned aerial vehicles (UAVs). The miniaturisation of sensors, low operation and equipment costs, and high flexibility in time made these tools more adapted than satellites for diversified and complex systems (Pádua et al., 2017). Thus, UAV makes it possible to acquire data at a very high resolution. For instance, UAV was used in Sarron et al. (2018) to characterise different types of mango-based orchards, including agroforests, in Senegal to extract cultivated area, tree structure and yield. However, UAVs are limited by the inherent low spatial coverage that does not allow measurement at a scale larger than the landscape. Nevertheless, UAVs can improve the understanding at the field scale by assessing the spatial and temporal variabilities in biomass and agrobiodiversity (Pádua et al., 2017).

Despite the many advantages of remote sensing, some geospatial data are obtained from simulation models for downscaling or upscaling observation data (You et al., 2009). Downscaling methods use models to desegregate coarse data (e.g., national level) using contextual information such as climate or land cover (You et al. 2009). Upscaling uses methods to interpolated field observations to non-observed areas. Thus, modelling approaches are subject to numerous biases related to the spatial coverage and representativeness of field sampling. Upscaling approach is particularly used to estimate the yield of fruit trees. Indeed, fruit tree yields remain challenging to model using spectral information and require annual sampling for calibration (Sarron et al., 2018).

ICT tools for data acquisition and management

In SSA, Internet coverage and ICTs (e.g., mobile phone) use have increased significantly in the last decade (World Bank, 2019), enabling the deployment of digital tools that facilitate data acquisition. Indeed, some mobile applications are available to facilitate manual data reporting, standardisation, and uploading to backup servers (e.g., survey applications like KoBoToolbox, www.kobotoolbox.org). The increasing use of mobile phones by farmers lead to the deployment of mobile applications that directly allow them to report and exchange information. Some tools can combine image analysis and data acquisition to extract and provide relevant information to the farmers (Mendes et al., 2020). For instance, PixFruit, a solution developed by CIRAD and the company SOWIT, uses deep learning analysis of images acquired by a smartphone and agronomic models to estimate the fruit production of a sample of trees and estimate the yield at the field scale (Faye et al., 2019). Finally, ICT tools for field data acquisition improve the reliability and the amount of data collected that can have a research interest if made available (Faye et al., 2019).

Digital agriculture to improve the horticultural sector

By collecting accurate data accessible to farmers, digital agriculture can help farmers in crop management by providing relevant information on their fields, such as disease or crop status (Mendes et al., 2020). Digital services improve access to information, knowledge, and skills and help decision making and market access. However, farmers' adoption of digital agriculture tools can be limited by numerous factors, including education to the tool and its usage, pertinence of the service provided and interest for the farmer, access to the Internet and ICTs (World Bank, 2019). The design thinking approach invites farmers to give feedback on the tool during its conception, thus improving the adoption, diffusion, and use of the solution by the most significant number of people (Bellon Maurel and Huyghe, 2017). Nevertheless, it is crucial to ensure that the tools are accessible to the greatest number of farmers, especially the most vulnerable who cannot have access to expensive technologies.

At the scale of institutions (research institutes, governments, NGOs, professional organizations), digital agriculture increases the quantity, the quality, and the temporal and spatial coverage of the data (Faye et al., 2019). In order to make such a flow of data relevant for the horticultural sector, data must be open and shared (e.g., using a data repository). Open data can improve knowledge on horticultural systems and help researchers address specific issues to improve the sustainability and efficiency of the farms. For national institutions, digital agriculture might be a lever to inform decisions (e.g., estimate yields at the district level) that contribute to developing the horticultural sector. Finally, it is essential that international institutions share their data as much as possible, especially for countries that do not have the capacity to produce and store them.

Despite its attractiveness, the application of digital agriculture to horticulture must meet several challenges, particularly related to the robustness and costs of such tools. As seen previously, data reliability should be at the centre of attention when developing assessment tools based on digital agriculture. Finally, digital agriculture also raises numerous interrogations on the security and privacy of the data that are prerequisites for their dissemination (World Bank, 2010).

CONCLUSION

The scarcity and reliability of agricultural data are critical issues in SSA. It is especially true for horticultural crops for which there is less knowledge and measurement methods are less developed. In this study, we have highlighted the existence of some databases freely accessible for the horticultural sectors. However, they often lack spatial resolution and most databases contain only information at the national level. Several factors impact the quality and quantity of the data: the low financial resources of statistical agencies; the numerous biases and errors attributed to assessment methodologies; the lack of share and reuse of data among stakeholders. Digital agriculture can help improve the data situation in SSA by supporting accurate and standardised data acquisition and by encouraging data share and reuse through open access management. Thus, digital agriculture is one of the pathways to achieving food security by improving farm efficiency and sustainability, increasing farmers' income and informing development policies. Nonetheless, digital agriculture would address the challenge of food security if and only if it succeeds in adapting to the specificity of sub-Saharan horticultural systems and maximises its adoption by users.

ACKNOWLEDGEMENTS

This work was supported by the French National Research Agency under the Investments for the Future Program, with the project #DigitAg referred as ANR-16-CONV-0004.

Literature cited

Bellon Maurel, V., and Huyghe, C. (2017). Putting agricultural equipment and digital technologies at the cutting edge of agroecology. OCL *24*, D307.

de Bon, H., Huat, J., Parrot, L., Sinzogan, A., Martin, T., Malézieux, E., and Vayssières, J.-F. (2014). Pesticide risks from fruit and vegetable pest management by small farmers in sub-Saharan Africa. A review. Agron. Sustain. Dev. *34*, 723–736.

Carletto, C., Jolliffe, D., and Banerjee, R. (2015). From Tragedy to Renaissance: Improving Agricultural Data for Better Policies. J. Dev. Stud. *51*, 133–148.

Desiere, S., Staelens, L., and D'Haese, M. (2016). When the Data Source Writes the Conclusion: Evaluating Agricultural Policies. J. Dev. Stud. *52*, 1372–1387.

FAO (2019). Response rate to FAO questionnaires and data collection plans for 2020 (Libreville, Gabon: African Commission on Agricultural Statistics, FAO).

FAO (2020). Africa - Regional overview of food security and nutrition (Accra, Ghana: FAO).

Faye, É., Sarron, J., Diatta, J., and Borianne, P. (2019). PixFruit: un outil d'acquisition, de gestion, et de partage de données pour une normalisation de la filière Mangue en Afrique de l'Ouest aux services de ses acteurs. (Dakar, Sénégal), pp. 10–11.

Gómez, M.I., Barrett, C.B., Raney, T., Pinstrup-Andersen, P., Meerman, J., Croppenstedt, A., Carisma, B., and Thompson, B. (2013). Post-green revolution food systems and the triple burden of malnutrition. Food Policy *42*, 129–138.

Karst, I.G., Mank, I., Traoré, I., Sorgho, R., Stückemann, K.-J., Simboro, S., Sié, A., Franke, J., and Sauerborn, R. (2020). Estimating Yields of Household Fields in Rural Subsistence Farming Systems to Study Food Security in Burkina Faso. Remote Sens. *12*, 1717.

Kindt, R., John, I., Ordonez, J., Dawson, I., Lillesø, J.-P.B., Muchugi, A., Graudal, L., and Jamnadass, R. (2019). Agroforestry Species Switchboard: a synthesis of information sources to support tree research and development activities. Version 2.0. World Agroforestry Centre, Nairobi, Kenya.

Mendes, J., Pinho, T.M., Neves dos Santos, F., Sousa, J.J., Peres, E., Boaventura-Cunha, J., Cunha, M., and Morais, R. (2020). Smartphone Applications Targeting Precision Agriculture Practices—A Systematic Review. Agronomy *10*, 855.

Pádua, L., Vanko, J., Hruška, J., Adão, T., Sousa, J.J., Peres, E., and Morais, R. (2017). UAS, sensors, and data processing in agroforestry: a review towards practical applications. Int. J. Remote Sens. *38*, 2349–2391.

Paut, R., Sabatier, R., and Tchamitchian, M. (2019). Reducing risk through crop diversification: An application of portfolio theory to diversified horticultural systems. Agric. Syst. *168*, 123–130.

Sarron, J., Malézieux, É., Sané, C., and Faye, É. (2018). Mango Yield Mapping at the Orchard Scale Based on Tree Structure and Land Cover Assessed by UAV. Remote Sens. *10*, 1900.

Usha, K., and Singh, B. (2013). Potential applications of remote sensing in horticulture—A review. Sci. Hortic. *153*, 71–83.

Weinberger, K., and Lumpkin, T.A. (2007). Diversification into Horticulture and Poverty Reduction: A Research Agenda. World Dev. *35*, 1464–1480.

World Bank (2010). Global strategy to improve agricultural and rural statistics (Washington, DC: The World Bank).

World Bank (2019). Future of Food: Harnessing Digital Technologies to Improve Food System Outcomes (World Bank).

You, L., Wood, S., and Wood-Sichra, U. (2009). Generating plausible crop distribution maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. Agric. Syst. *99*, 126–140.

Table 1. Example of some databases accessible and useful for horticultural sector in SSA

Database name	Assessment method	Variables	Spatial resolution	Temporal resolution	Dataset used	Comments
Horticultural cropping systems						
FAOSTAT[a]	Annual survey	Production, harvested area inputs, prices, trade, etc.	National	1961 - present		
PSD[b]	Annual survey	Production, trade	National	1960 - present		Only temperate fruit
CountryStat[c] and AgroMAPS[d]	Annual survey	Production, harvested area inputs, prices, trade, etc.	Sub-national	1961 - present	FAOSTAT[a]	26 countries in SSA. Lots of missing data
LSMS[e]	2-4 years survey	Production, harvested area, soil fertility, labour, etc.	Field to national	Irregular		11 countries in SSA
Data Africa[f]	Data compilation + model simulation	Production, harvested area, water supply, etc.	Sub-national	Averaged	FAOSTAT[a], CRU[k], etc.	13 countries in SSA
GAEZ[g]	Data compilation + model simulation	Production, yield gaps, climatic indices, etc.	≈ 10 km grid	Averaged	FAOSTAT[a], CRU[k], etc.	
EarthStat[h]	Data compilation + model simulation	Production, harvested area, nutrient balance, etc.	≈ 10 km grid	Averaged	FAOSTAT[a]	175 crops included
Crop species and cultivars						
Hortivar[i]	Observational	Cultivar description, production, seed sources, etc.	Points	Irregular		>70 000 technical sheet covering fruits, vegetables, roots & tubers, ornamentals, mushrooms, herbs & condiments.
ASS[j]	Observational	Species description, management practices, genetic, etc.			ICRAF[l], others	Repository of information sources on plant species

[a] www.fao.org/faostat; [b] Production Supply and Distribution: apps.fas.usda.gov/psdonline/app/index.html#/app/home; [c] www.fao.org/in-action/countrystat; [d] kids.fao.org/agromaps; [e] Living Standards Measurement Study: www.worldbank.org/en/programs/lsms; [f] dataafrica.io; [g] Global Agro-Ecological Zones: www.fao.org/nr/gaez/en; [h] www.earthstat.org; [i] www.fao.org/hortivar; [j] Agroforestry Species Switchboard (ASS): www.worldagroforestry.org/products/switchboard; [k] Climatic Research Unit, University of East Anglia: www.cru.uea.ac.uk; [l] World Agroforestry (ICRAF): www.worldagroforestry.org