Food and Agriculture Organization of the United Nations

A tool to support the spatial prioritization of commodity-specific investments

An application to Uganda

FAO AGRICULTURAL DEVELOPMENT ECONOMICS WORKING PAPER 22–12

ISSN 2521-1838

November 2022

A tool to support the spatial prioritization of commodity-specific investments

An application to Uganda

Kougblenou Christophe Adjin

Francisco Pereira Fontes

Marco V. Sánchez

Food and Agriculture Organization of the United Nations

Rome, 2022

Required citation:

Adjin, K.C., Fontes, F. & Sánchez, M.V. 2022. *A tool to support the spatial prioritization of commodity-specific investments – An application to Uganda*. FAO Agricultural Development Economics Working Paper 22-12. Rome, FAO. https://doi.org/10.4060/cc2840en

ISSN 2664-5785 [Print]

ISSN 2521-1838 [Online]

ISBN 978-92-5-137165-7

Contents

Abstract v
Acknowledgements vi
1 Introduction 1
2 Conceptual description of the approach 4
3 Steps to implement the method and data requirements 8
4 Results 17
5 Conclusions and discussion 25
References 28
Annex 30

Tables

Table 1. Sectoral ranking by the impact of government investment on five socioeconomic indicators (only top-ten commodities are shown) 18
Table 2. List of commodities and approach used 18
Table 3. List of selected districts..... 22

Table A1. Data sources..... 30
Table A2. Robustness checks 38
Table A3. Robustness checks: Shares of selected districts per region 44
Table A4. Robustness checks: Selected districts and the number of specifications where they appear 44
Table A5. Robustness checks: List of selected districts by permuting steps 2 and 3..... 47

Figures

Figure 1. Iterative elimination process 7
Figure 2. Decision tree for implementing the methodology 9
Figure 3. Sequential selection of prioritized districts for banana production based on potential, unrealized potential and poverty 20
Figure 4. Sequential selection of prioritized districts for cassava production based on potential, unrealized potential and poverty 20
Figure 5. Sequential selection of prioritized districts for coffee production based on potential, unrealized potential and poverty 20
Figure 6. Sequential selection of prioritized districts for goat rearing based on potential, unrealized potential and poverty 21
Figure 7. Sequential selection of prioritized districts for maize production based on potential, unrealized potential and poverty 21
Figure 8. Sequential selection of prioritized districts for millet production based on potential, unrealized potential and poverty 21
Figure 9. Sequential selection of prioritized districts for sugar cane production based on potential, unrealized potential and poverty 22

Figure A1. Results: Potential dimension 32
Figure A2. Results: Unrealized potential dimension 34
Figure A3. Results: Selected districts..... 36
Figure A4. Robustness checks: Selected districts by permuting steps 2 and 3..... 45

Abstract

In this paper we propose a simple methodology to select a limited number of geographical areas to prioritize for commodity-specific investments in Uganda. Similar to other approaches for prioritizing investments geographically, the basic idea behind the proposed method is that districts with high agro-ecological potential, that are also far from their potential and have high levels of poverty, should be prioritized for commodity-specific investments as they are where investments are likely to have the highest impact. The methodology then proposes an iterative elimination algorithm to provide a list of suggested districts that rank high in all dimensions.

We apply this methodology to districts in Uganda and build on previous FAO evidence that ranked Ugandan agricultural sectors (and their related commodities) at the national level based on their economic and social welfare cost-effectiveness. We apply the approach to identify districts that have a high theoretical investment potential for sectors (and their related commodities) selected from the aforementioned ranking. For the illustrative purposes of this paper, the number of selected districts was set to five.

The results highlight that prioritized districts are very context and commodity specific. In certain cases (e.g. sugar cane or millet), prioritized districts tend to be highly concentrated in one geographical region, whereas they tend to be more spread out for the sectors producing other commodities (i.e. bananas, coffee, goats, cassava and maize).

The results are expected to inform a discussion with policymakers in Uganda which is expected to culminate in the selection of an even narrower set of districts for which more in-depth analyses of commodity-specific investments will be undertaken at the level of priority areas, including, among others, irrigation, mechanization, seeds and fertilizers.

Keywords: agriculture; agricultural economics; agricultural investments; Uganda; spatial prioritization.

JEL codes: Q100, Q140, Q180.

Acknowledgements

This paper was conducted under FAO's Multidisciplinary Fund Project RA20201001C00 "Strengthening governments' capacity for enabling an economic and social recovery post-COVID-19 though investments in agrifood sectors". It was further developed and finalized by the FAO's Monitoring and Analysing Food and Agricultural Policies (MAFAP) programme.

It benefited extensively from comments provided during a presentation to Uganda's National Planning Authority (NPA) and the Ministry of Agriculture, Animal Industry and Fisheries (MAAIF). Specifically, from the NPA, the authors wish to thank Pamela K. Mbabazi (Chairperson, NPA), Joseph Muvawala (Executive Director, NPA), Asumani Guloba (Director of Development Planning, NPA) and Ollen Wanda Kahurubuka (Senior Planner Agriculture, NPA), for the support provided throughout the implementation of this FAO activity, as well as for the useful feedback received. From MAAIF, the authors wish to thank Frank Kagyigyi Tumwebaze (Minister of Agriculture), Fred Mayanja (Commissioner, Planning, MAAIF) and Robinson Lufafa (Senior Statistician, MAAIF).

The authors also wish to thank colleagues from the FAO Uganda country office for the fantastic support received throughout the development of this study and useful comments, which ultimately helped improve its quality and country specificity. Specifically, the authors wish to thank António Querido (FAO Representative, FAO Uganda), Priya Gujadhur (Deputy Representative, FAO Uganda), Jean-Marie Byakweli (Policy Officer, FAO Uganda), Leila Shamsaifar (Resource Mobilization and Partnerships Specialist, FAO Uganda) and Dominique Reumkens (Associate Professional Officer, FAO Uganda).

The document also benefited from comments from an anonymous referee as part of the ESA Working Paper Series peer-review process, as well as colleagues in the FAO Agrifood Economics Division (ESA), specifically Emiliano Magrini and Martín Cicowiez.

Finally, the authors are also grateful to Craig Lawson (ESA, FAO) for copy-editing the paper and Daniela Verona (ESA, FAO) for the editorial and layout support, as well as publishing coordination.

Corresponding author: Kougblenou Christophe Adjin (Kougblenou.Adjin@fao.org)

1 Introduction

The impressive economic growth and poverty reduction witnessed in Uganda over the last 20 years has brought the country close to achieving "middle-income" status. Past analyses have highlighted that the agriculture sector played an important role in this process (World Bank, 2016). For the 2021–2025 period, under the auspices of the Third National Development Plan (NDPIII), the Government of Uganda aims to roll out an ambitious investment programme that is expected to stimulate the economy and help the country reach "middle-income" status. To do this, the government aims to achieve a total investment (i.e. including both public and private investment) of 411.7 trillion Ugandan shillings (of which UGX 276.9 trillion is of public financing),[1] with a yearly level of total investment equivalent to 30 percent of GDP (NPA, 2020).

Such investments are critical to stimulate the economy in normal times and help the country achieve middle-income status, but they have become even more necessary for the recovery in the aftermath of the COVID-19 pandemic. While Uganda's agriculture does not seem to have been affected to the same extent as other sectors in the economy as a result of the pandemic, overall economic growth was negatively affected in 2020 and 2021 (by about 0.2 percentage points in 2020/21, according to the IMF Article IV mission). Furthermore, the IMF's recent projections for Uganda (IMF, 2022) have downgraded the forecasted speed of recovery for the 2021/22 fiscal year.[2] In such times, countercyclical fiscal spending can alleviate the pandemic-induced negative impacts on the economy and support the post-COVID-19 recovery, and investments in agriculture are likely to be important drivers of economic growth due to the importance of this sector in Uganda. However, the economic recession brought about by the pandemic, and the resulting more binding fiscal constraint, will demand more discipline and effectiveness at the time of spending and investing public resources. Prioritizing expenditures in agriculture, in particular, is therefore a must in its own right, as well as creating more enabling conditions for private investments.

Under the NDPIII, the Government of Uganda seeks to invest UGX 18.7 trillion[3] (of which 9.2 will be funded with public resources) in the agrifood sector and expects to, among other objectives, attain an agricultural growth rate of 6 percent, create 800 000 jobs, and triple the value of exports over a five-year period (NPA, 2020). These ambitious goals epitomize the importance given to the sector and this should come as no surprise, given that agriculture remains a sector of critical importance in Uganda. The agriculture sector accounts for 72 percent of total employment and 24 percent of GDP and, despite steady growth in the past, is still characterized by a vast, unexploited potential (World Bank, 2020; Fiala and Apell, 2017). In addition to this, evidence abounds on the poverty-reducing potential of agriculture, both in the academic and policy literature. Studies generally highlight that agricultural growth has a higher poverty-reducing potential than growth in other sectors of the economy (Ligon and Sadoulet, 2018; Dorosh and Thurlow, 2018; Ivanic and Martin, 2018).

However, in the same way that there is a general consensus on the poverty-reducing effects of agriculture, there is also a consensus that within agriculture, the composition of investments matters and may lead to very different outcomes. Although the evidence on this is more limited,

[1] Equivalent to a total investment of USD 107.9 billion, of which USD 72.5 billion is of public financing.

[2] The most recent IMF Article IV mission document downgraded the growth projection from an initial 4.1 percent to 3.8 percent.

[3] Equivalent to USD 4.9 billion.

studies show that functional, geographic, and commodity composition of investments lead to very heterogeneous impacts (Fan, Gulati and Thorat, 2008; Fan, Yu and Saurkar, 2008; Fan and Zhang, 2008; Mogues *et al.* 2012; Pernechele *et al.* 2021). Overall, at the risk of oversimplifying, three general stylized facts emerge from the literature. First, the economic returns on investments in public goods (e.g. including, among other investments, those on research and development, extension and infrastructure) outweigh those of private goods (e.g. fertilizer subsidies) (Jayne *et al.* 2018; Fan, Gulati and Thorat, 2018). Second, the economic returns to investment on different crops differs substantially and is very country specific. This was shown by Diao *et al.* (2010) in the case of Rwanda and evidence provided in Sánchez, Cicowiez and Fontes (2022) suggests that this is also the case for Uganda.[4] Finally, while the evidence on this topic is still limited, evidence suggests that investments in poorer and more underdeveloped areas tend to have higher returns in terms of poverty reduction and agricultural GDP growth (Mogues *et al.* 2012). Even when this is not the case, poverty rates can be used as a proxy of the urgency of an investment (Maruyama *et al.* 2018).

As part of FAO's support to the Government of Uganda in the framework of post-COVID-19 recovery and the Monitoring and Analysing Food and Agricultural Policies (MAFAP) programme's policy prioritization's work, a workstream has been established to support Ugandan decision-makers in prioritizing agrifood investments. More specifically, FAO is providing evidence on the cost-effectiveness of investments, and this implies looking at different dimensions of investment composition, including commodity (or production sector), geographic and functional composition of investment. This process started with the development of an economy-wide and multisectoral analysis to identify the production sectors where investments would be most cost-effective to achieve several agricultural transformation outcomes under current macroeconomic constraints, such as overall GDP growth, agrifood GDP growth, rural poverty reduction and export growth (Sánchez, Cicowiez and Fontes, 2022). From the financing of investment point of view, it is important for policymakers to understand which investments in agriculture have the potential to generate economic and revenue growth at the lowest cost. This analysis ultimately resulted in a ranking of sectors/commodities that Ugandan policymakers can use to guide the prioritization of their sub-sectoral investments.

However, this ranking is only the departing point to further guide decisions of policymakers who, in addition to choosing a commodity, need to decide **what** and **where** to invest. These are precisely the two questions that FAO support seeks to address as part of its ongoing cooperation with the Government of Uganda. This paper specifically addresses the question of **where** to invest for those commodities that ranked high in the economy-wide and multisectoral analysis. A subsequent study will then look at **what** are the investment needs in areas that were identified as high priority for investments.

Making use of the ranking provided in Sánchez, Cicowiez and Fontes (2022), this study relies on a methodology for geographical prioritization of commodity-specific investments and applies it in the Ugandan context. Similar to other tools used in studies that focus on the geographical prioritization of investments (i.e. Maruyama *et al.* 2018; Marivoet *et al.* 2019), the basic underlying idea of the proposed methodology is that geographical areas where unrealized potential and poverty are high, are likely to be suitable areas for investment. We propose a simple measure of absolute commodity-specific unrealized potential and then use an iterative elimination approach based on three dimensions likely to be strongly correlated

[4] This observation seems to apply also to middle income countries. For evidence in this regard, see the case of Mexico's agriculture in Sánchez and Cicowiez (2022).

with returns to investment, namely potential, poverty and unrealized potential. We then propose, for each commodity, five locations with a high **expected return to investment** and, therefore, high theoretical potential for investment. These findings will then be discussed with policymakers and sector experts to assess their relevance and accuracy, before selecting a limited number of geographic units where specific commodity investment needs will be assessed at a later stage.

The rest of the paper is structured as follows. Section 2 discusses the selection of commodities, presents the methodology at a conceptual level, explains how it is implemented in practice, and highlights its limitations. Section 3 explains the methodology in more detail and describes some special cases where a deviation from the suggested methodology is necessary. Section 4 presents the results from applying the methodology for Uganda for seven selected commodities (i.e. bananas, cassava, coffee, maize, millet, goats and sugar cane), and culminates in a table that identifies five districts with a theoretical high potential for investment for each commodity. Section 5 concludes and proposes potential areas to expand the tool.

2 Conceptual description of the approach

With the recent improvements in the availability of, and accessibility to, spatial data, analysing the spatial dimension of phenomena has become more common and this has led to the development of approaches to help governments spatially prioritize investments, policies and interventions. Two recent applications to agriculture are the approaches proposed by Maruyama *et al.* (2018) and the approach by Marivoet *et al*. (2019).

The approach proposed by Maruyama *et al.* (2018) is essentially an approach that uses three dimensions (estimated potential, estimated unrealized potential and poverty) to create a typology of geographical units and classifying them different categories (e.g. areas with high poverty and low potential are considered "critical" and coloured in red in their proposed typology). This typology is then used to highlight how the differences in poverty, potential and unrealized potential call for different policies to overcome the challenges in different places.

Marivoet *et al.* (2019) propose a method that is conceptually similar in that it seeks to classify geographical units into high/medium/low priority and high/low agricultural potential. The main difference between the approaches relates to the dimensions and the way they are calculated or estimated. The proposed approach in Marivoet *et al.* (2019) aims to prioritize areas based on the dimensions of food security and use availability (proxied by production potential), access (proxied by food consumption) and utilization (proxied by anthropometric data) as the main dimensions to create the typology. The other main difference between the Marivoet *et al.* (2019) and the Maruyama *et al.* (2018) approaches is the definition of potential used. In Marivoet *et al.* (2019), the authors start by identifying the pixels under crop cultivation, assume that all pixels cultivate the same share of cassava, rice, maize, beans and plantain and then multiply these crop-specific assumed areas by the potential yield of crops obtained in local field stations of the national agricultural research institute. As a result, the authors obtain a measure of potential food energy production per capita in a given geographical area.

While these existing methods are useful, they suffer from two important drawbacks that limit their usefulness when looking at crop-specific spatial prioritization. The limited usefulness when focusing crop-specific recommendations is an important drawback, given that commodity selection is often key when governments prioritize investments in agriculture. First, both methods are constructed in order to provide results for agriculture as a whole, rather than a specific crop. In the case of the Maruyama *et al.* (2018) approach, this is because required samples to accurately estimate potential are often lacking for specific crops and this issue is compounded by the fact that samples are unlikely to be representative at very granular geographical levels for specific crops. In the case of Marivoet *et al.* (2019), since their definition of potential focuses on a combination of crops, without modifications, this method cannot be used for crop-specific analyses. Second, the data requirements of both methods are quite onerous which means that the time required to carry out the analysis may not align with the needs of policymakers, who may have pressing needs for evidence to inform their decisions. Both the Maruyama *et al.* (2018) and Marivoet *et al.* (2019) require in-depth cleaning of large datasets. In the case of Marivoet *et al.* (2019) in addition to this, data on spatially disaggregated yields from research stations are needed for the main commodities, which may not be easily obtainable or not available to researchers altogether in many low-income countries, certainly not in sub-Saharan Africa.

This paper aims to develop a spatial prioritization tool that is commodity specific, easier to generalize and less data intensive. On a conceptual level, similar to the approach proposed by Maruyama *et al.* (2018), our approach relies on three key concepts, namely that of **potential,**[5] **unrealized potential and poverty** to determine locations (or districts, in our Uganda application) with a theoretically high cost-effectiveness for investments from the point of view of primary production. These concepts are deliberately chosen for several reasons. First, as highlighted in Maruyama *et al.* (2018), most governments seek to reduce poverty and the level of poverty is often a good proxy for the urgency of an intervention. In addition to this, in areas with a large number of poor people, investments are more likely to lead to reductions in poverty than in areas with a low number of poor people. Second, a large unrealized potential (gap in the value of production as per our method) in a given location indicates that, on average, a given geographical area is still very far from their productive potential. The main idea is that, if a given location is already very close to their potential, a large amount of investment will be needed to achieve a modest result (in other words, the marginal returns to investment are likely to be low). On the other hand, choosing a location that is still far from its maximum potential is likely to avoid the pitfalls of stark diminishing returns to investment and will also focus on areas where agro-ecological potential is high to start with.

However, **potential, unrealized potential and poverty** need to be defined and measured. Potential and unrealized potential are typically not observed, so we need to rely on proxy variables in an attempt to represent them. For potential, we use the current market value of the maximum attainable yield[6] of a given crop as a measure of potential. In the case of unrealized potential, we use the yield gap valued at market prices.[7] Finally, for the dimension on poverty, we use the headcount poverty ratio of the most recent year available.

Conceptually, the reliance on potential, unrealized potential and poverty is quite similar to other geographical prioritization frameworks (e.g. Maruyama *et al.* 2018, 2019). Our approach, however, has an explicit focus on specific commodities, aims to be much simpler and less data intensive than other approaches, allows for cross-commodity comparisons (to a certain extent) and, in an ideal setting, is informed by a previous analysis that has helped to rank sectors (and their commodities) based on the potential economy-wide and poverty-reduction gains associated with investing in these sectors.

The main idea of our approach is therefore to integrate publicly available spatial data with price data[8] to obtain the layers on potential and unrealized potential, as well as separately obtain data on poverty rates. Once these data are obtained, the variables are constructed and mapped, and we subsequently use a simple iterative elimination process to select a given

[5] However, as will be explained later, our definition of potential differs from that used by Maruyama *et al.* (2018). Specifically, our definition of potential focuses on crop-specific attainable yields, whereas that of Maruyama *et al*. (2018) focuses on overall potential market revenue/profits across all crops.

[6] As will be explained later, for most crops we focus on the current market value of the maximum attainable yield. The concept of maximum attainable yield can broadly be defined as the yield of a crop when grown under favourable conditions without growth limitations from water, nutrients, pests, or diseases, and is therefore determined by solar radiation, temperature, and water supply (Lobell *et al.* 2009). As will be explained later, the concept is somewhat different in the cases of livestock and certain crops with no available spatial data, where we use maximum predicted livestock density as a proxy (livestock) or market revenues (crops with no spatial data), respectively.

[7] Yield gap is defined as the difference between the maximum attainable yield and the observed yield. As will be described later, for livestock, since yield gaps are not observed, we focus instead on density gaps.

[8] To obtain the current market value of the maximum attainable yield and current value of production, we use observed market prices. These do not include how prices would have responded to changes in domestic supply.

number of prioritized geographic units. We have applied the approach using districts as the geographical unit, where a given number (*D*) of districts are eliminated in each stage.[9]

For each commodity, we follow a four-step process, as depicted in Figure 1.[10] In the first step, we use a filter that excludes districts where the share of farmers cultivating a given commodity is below a given threshold. The use of a threshold acts as a proxy for the cultural acceptability of the given commodity, as further explained below. In a second step, for the subset of districts that meet the threshold in step 1, we keep the tercile of districts[11] with the highest overall yield potential for a given commodity. In the third step, for the subset of high overall potential districts, we select the tercile of districts with the highest level of unrealized potential. Finally, in the fourth step, within the subset of districts with both high potential and high unrealized potential, we give priority to high poverty districts and select the five districts with the highest poverty levels.

It is important to make two considerations regarding the first and the last steps of the process. The first step is, strictly speaking, optional, which means that the methodology can be applied both with and without filters. Not including the filter means that the approach will provide a list of districts where theoretical potential is likely to be higher but will not take into account important issues such as cultural acceptability of the commodity. Similarly, it will not account for the fact that, if in a given location or the fact that if the commodity is not yet cultivated, the cost of promoting this commodity in this area and training farmers to cultivate it effectively, are likely to be very high. In our case, we opt for the filter because it acts like a proxy for cultural acceptability and it also is likely to prioritize those districts where closing the unrealized gap could lead to the highest increase in total volumes is high, rather than just the per hectare value of these gaps.[12] We also note that while we use the share of households cultivating a given commodity as a filter, the approach is very flexible and other filters can be applied, as long as there are layers that link these filters to the analysed districts.

There are also several reasons why we prioritize poverty rates as the last criterion in our four-step process. The first is that, since the commodity selection is based on the results of a previous economy-wide modelling exercise that also ranked sectors and their commodities based on the simulated poverty impact stemming from investments in these sectors, poverty impacts are, to a certain extent, already embedded in the commodity selection. The second reason is that we argue that a certain degree of potential is essential for commodity-specific investments to work. If poverty is used as the first criterion, and if areas with high poverty are also those with the lowest potential, the selected districts for given commodities may have little or no potential. Finally, since the poverty layer is the only static layer across all the layers used

---

[9] D is calculated using the following formula: $D = \sqrt[3]{\left(\frac{N-Exc}{d}\right)}$, where *N* is the total number of geographical areas, *Exc* denotes the number of districts that do not meet the filter threshold defined below, *d* is the desired number of districts at the end of the iterative process, and *D* is the factor by which the districts need to be reduced at each stage. For example, if, as in our case, we wish to have five districts and there are 135 districts, *D* must be equal to three. This means that at each stage (except the last stage) we need to divide the number of districts by three to proceed with the iterative elimination process. For the last stage, because we apply the filter (which means that the initial number will not always be 135), rather than divide by three, we keep five observations (i.e. exclude (N_3-5) observations, where N_3 is the number of observations left for the last step.

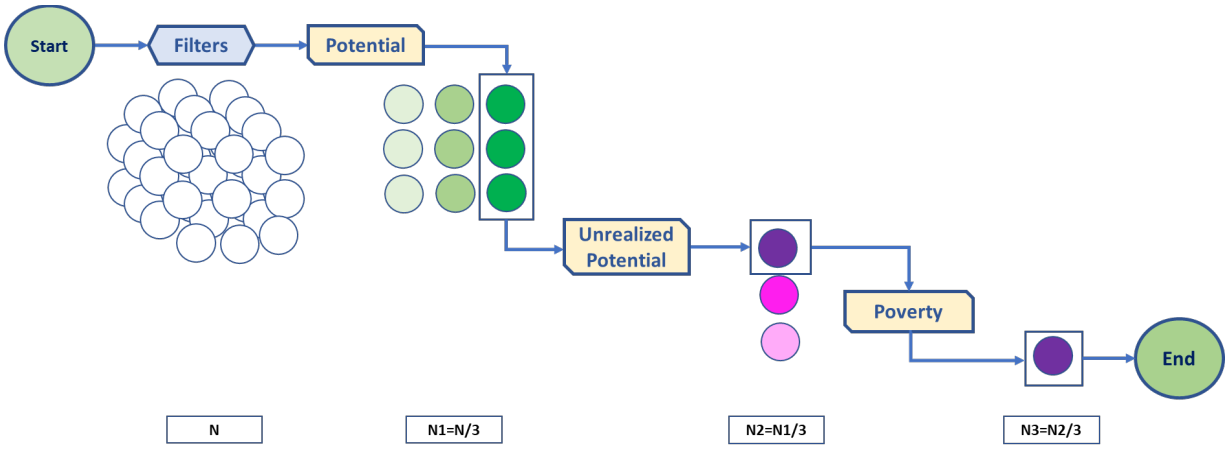
[10] In Figure 1, the sample size is equal to nine for illustration purposes only. In our actual application of the method, *N* will refer to the sample sizes in the dataset used for Uganda.

[11] The choice of a tercile of districts is informed by the formula that defines *D* (i.e. number of districts) since we want to have five districts in the end and have 135 districts to start with.

[12] Beliefs and acceptance have been shown to be important in the context of Uganda for the introduction of new technologies and crop varieties, as shown in Ndaula *et al.* (2019) and Mulugu *et al.* (2022).

in the approach, if it is used as the first criterion, all the suggested geographic units for investments will always be concentrated in the quantile containing the poorest districts (i.e. the 33 percent poorest districts in our case), hence severely limiting the choice and variation of targeted districts.

Figure 1. Iterative elimination process



Source: Authors' own elaboration.

This iterative elimination process thus produces a list of prioritized geographical units (i.e. districts, in our case), which we represent in maps. The idea is that this data-driven and easy-to-visualize approach to prioritization should not be the end of the exercise, but should instead then be used for discussion with local policymakers and experts to assess the relevance and pertinence of the list and maps containing the selected districts.

It is important to stress, however, that the maps resulting from this approach may not always be a perfect match with high performance areas and, therefore, may not always fully align with the expectations of local experts. The main reason for this is that, by definition, high performance areas are already performing well (or, at the very least, better than average) and this implies that the *per hectare* unrealized potential in these areas may be low. As such, the iterative elimination process that uses the three layers (high potential, high unrealized potential and high poverty rates) may result in one or more districts that may not be obvious. While the methodology can easily be adjusted to ensure that only obvious choices are kept (i.e. a filter can be used to limit the set of districts to a very narrow set of obvious districts), we argue that a data-driven agnostic approach is valuable in that it can potentially highlight areas with high potential that are less obvious for policymakers, but where investments could be successful, and the commodity is widely culturally accepted. In addition to this, it can stimulate an interesting debate among practitioners that can improve the allocation of resources to those districts where they can have the highest impact. Finally, this also highlights the need to "ground-truth" the results. As accurate as geospatial data are, they are not perfect and there are several additional factors, beyond the three variables we use, that can be important to determine the location of investments (e.g. conflict areas, protected areas, distance to markets, distance to processing facilities, and so forth). Nevertheless, we argue that the tool we propose here remains useful in that it provides a simple, data-driven and easy-to-visualize way to reduce the set of districts where policymakers can invest.

The next section explains the methodology in more detail and outlines some special cases where a deviation from the methodology devised is required.

3 Steps to implement the method and data requirements

In this section we first explain the approach used for the initial filter, before describing the data needs and how to implement the methodology in three different cases. A summary of the different datasets and their respective data sources used for all layers are presented in the Annex (Table A1).

3.1 Defining the threshold for the filter

This step is optional in the methodology, but to ensure that selected areas are not areas where the commodity is not culturally acceptable, we use a filter to restrict the list of districts to a set of districts where there is already a sizeable proportion of farmers cultivating a given commodity. To do this, using the 2008/09 Agricultural Census (UBOS, 2010), we first start by calculating the share of households that cultivate a given commodity in each district. Based on the distribution of these data, we use a crop-specific threshold. If the proportion of farmers cultivating that given commodity in a given district in the 2008/09 Agricultural Census was below that threshold, the district is automatically excluded from the next steps of the analysis. Otherwise, if the proportion of farmers cultivating a given commodity in that district was above the threshold in the 2008/09 Agricultural Census, the district is kept in the analysis.

While it may seem strange to use the 2008/09 Agricultural Census for the first step of the analysis, there are two reasons why this was done. The first is that, using more recent datasets that are not representative at the district level would have cast doubts on the validity of the shares of households cultivating a given commodity. Second, the main aim of the filter is to reduce the set of districts such that districts where the commodity is likely not to be culturally accepted are not part of the analysis. In this sense, we argue that, if a large share of households were engaged in the cultivation/rearing of a given crop/livestock a decade ago, then this is a good proxy for cultural acceptability given that crop/rearing patterns tend to change slow.

However, including this filter raised two challenges. The first issue related to administrative boundaries, as these have changed since the 2008/09 Agricultural Census (UBOS, 2010), which is the latest dataset with representative data at the district level available to calculate the threshold. Since 2008/09, several new districts were created and these are included in the administrative boundaries map that we use, which is based on the 2020 administrative borders. As a result, while most districts remained constant, for some, there is a mismatch between the 2008/09 districts and the ones in the 2020 map. To solve this mismatch, we use the 2009 value of the parent district for all new districts. The second issue is related to determining a commodity-specific threshold. To make the thresholds meaningful, we need to ensure that the number of included districts remains high, while also making sure that the proportion of farmers that cultivates the commodity is sufficiently high for the suggested district to be meaningful. We thus proceed as follows. For those commodities (i.e. bananas, cassava, maize and goats) that were widely cultivated or reared (more than 35 percent of farmers cultivate or rear the commodity in the median district), we use a threshold of 40 percent. For millet and coffee, we set thresholds close to the respective medians, at 15 percent (millet) and 10 percent (coffee), respectively. Finally, for sugar cane, where the production is much more geographically concentrated and where the proportion is 0 or close to 0 in most districts, we use a threshold of 5 percent, which is close to the 75th percentile of the distribution of districts.

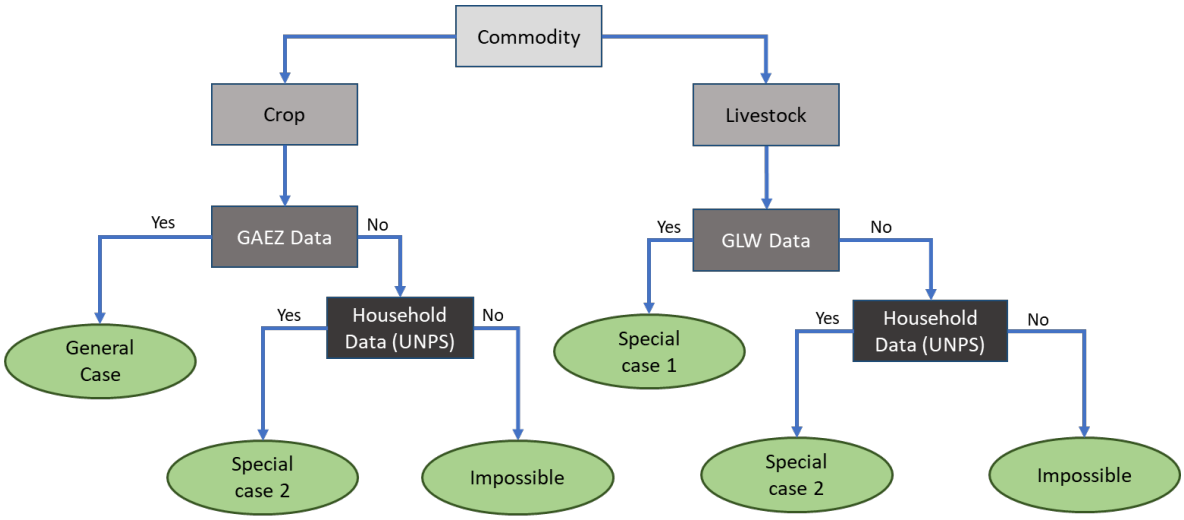
3.2 Deriving the layers: How to implement the methodology depending on the data availability

In this subsection we describe how to implement the methodology in three different cases, namely:

1) The general case – When both Global Agro-Ecological Zones (GAEZ)[13] and price data are readily available.
2) Special case 1: Livestock – We discuss how to implement the methodology for the livestock sub-sector, using the Gridded Livestock of the World (GLW)[14] and price data.
3) Special case 2: Crops when GAEZ data are not readily available.

In terms of when to apply each of these special cases of the approach and to understand when each of them can be used, we also rely on a decision tree (Figure 2). In its current form, the methodology cannot be easily implemented for fisheries or forestry due to the lack of a fixed geospatial layer with information on a measure that is equivalent to a yield gap. However, should such a layer be available, the method could easily be applied to these sub-sectors.

Figure 2. Decision tree for implementing the methodology



Source: Authors' own elaboration.

3.3 General case: description of the approach and data needs

The suggested approach is based on the combination of three variables (potential, unrealized potential and poverty). In the general case, when all the necessary data are available for a given crop, the approach consists of three steps.

In the first step, it is necessary to estimate the potential. In our approach, we define potential as the current market value of the maximum attainable yield for a given commodity valued at market prices. In the second step, it is necessary to estimate unrealized potential for a given commodity, which we define as the yield gap valued at market prices. In order to do so, for each

[13] These data are sourced from FAO's Global Agro-ecological Zones v4 Data Portal, available at https://gaez.fao.org/pages/data-viewer

[14] These data are sourced from the Gridded Livestock of the World dataset, available at https://dataverse.harvard.edu/dataverse/glw_3

district we calculate the difference between the potential yield and the observed yield and then multiply this by the most recent market price available. Algebraically, the potential and unrealized potential for commodity *c* in district *d* can be expressed as follows:

$$P_{cd} = (PYield_{cd}) * Price_{cd} \quad (1)$$

$$UP_{cd} = (PYield_{cd} - Yield_{cd}) * Price_{cd} \quad (2)$$

Where:

- P_{cd} represents the maximum value of production per hectare for commodity *c* in district *d*
- UP_{cd} represents the unrealized potential (UP) for commodity *c* in district *d*;
- $PYield_{cd}$ represents the potential yield of commodity *c* in district *d;*
- $Yield_{cd}$ represents the observed yield of commodity *c* in district *d*; and
- $Price_{cd}$ is the price of commodity *c* in district *d*.

In order to calculate unrealized potential, we calculated the yield gap $(PYield_{cd} - Yield_{cd})$ using data on potential and observed yields from GAEZ[15] for 2015, which is the latest year for which both variables were available. For potential yield, our definition focuses on the attainable yield for rainfed agriculture derived from crop-growth models that capture the agro-ecological conditions of a given district, assuming a low level of input use.[16] The actual yields were obtained using the data on the yield achievement ratio, which are also available from GAEZ.[17] In order to obtain the price data for the same year, we collected information for the year 2015 from the Uganda National Panel Survey (UNPS) 2015–2016. Specifically, for all selected crops, we calculate median district sale prices for that given crop.[18] The full set of information was available for sugar cane, bananas, maize, cassava and millet, which represent five of the seven commodities analysed in this paper. For the other two commodities, coffee and goats, data on the potential, were not available, and had therefore to be estimated (as further explained in the next subsections). As will be discussed further below, these commodities were selected as they ranked high in terms of simulated cost-effectiveness in the economy-wide model in Sánchez, Cicowiez and Fontes (2022).

[15] The maximum theoretical yield is based on crop-growth models taking into account the characteristics of the agro-ecological zones in which the district is located. GAEZ version 4 defines two types of potential yields: i.e. the agro-climatic potential yield and the suitability attainable yield. The agro-climatic potential yield is based on eco-physiological crop growth model and spatially detailed climate characteristics (radiation, temperature and precipitation) during different crop development stages. It also accounts for temperature limitations and moisture constraints, yield reducing effects due to pests, diseases and weeds as well as climate related workability constraints. On the other hand, suitability attainable yield combines agro-climatic potential yields with the results of soil and terrain evaluation. Suitability attainable yield accounts therefore for constraints induced by soil limitations and prevailing terrain-slope conditions (Fischer *et al.* 2021). We use the attainable yields under rainfed conditions with low quantities of inputs for the purposes of our exercise. In GAEZ, this is associated with acronym (c_cruts_lr).

[16] In GAEZ, there are various definitions of potential yield, based on different assumptions related to irrigated and rainfed agriculture, different measures of input intensity and management efficiency. In our case, we opted for the layer using rainfed conditions and low input use, which we considered most likely to proxy the prevailing conditions of most small-scale farmers. We tested the sensitivity of this choice in the results section.

[17] The apparent yield gaps are closely related to the calculated yield achievement factors, both summing up to 100 percent. For instance, a yield achievement factor of 75 percent would imply an apparent yield gap of 25 percent (Fischer *et al*. 2021).

[18] Whenever the sample size of sales of a crop was too low in a given district, we used region prices instead.

Once we have calculated the potential and unrealized potential, we combine this with available poverty data[19] and apply the iterative elimination process described earlier to obtain a restricted number of districts.

We note that, in the absence of price data, we can still calculate and compare a proxy of unrealized potential based solely on the yield gaps. However, this will mean that the approach will not be comparable across crops, which highlights a key advantage of adding the prices in the estimation.

3.4 Other cases: Livestock and cases where data are missing

When there are no missing data, the general case can be used. However, a layer of complexity is added for those commodities (e.g. crops not included in GAEZ and all livestock commodities) where data necessary to compute a yield gap are either not readily available or when the concept of a yield gap itself is more complex (e.g. livestock). In these cases, an additional step needs to be undertaken to estimate the potential and this subsection walks through three special cases, focusing on 1) livestock; 2) crops when no GAEZ data are available; and 3) cases when no price data are available.

3.5 Special case 1: Applying the methodology for livestock

In the case of livestock, the overall methodology is slightly different and involves an additional step with respect to the general case.

Since, to our knowledge, there are no data on spatial production and productivity of the main livestock products, we use a proxy for which spatial data exist, namely livestock density. The assumption is that livestock density is highly correlated with the livestock (and by-products) revenue potential. However, the issue is that spatial data on livestock density are only available for **observed** density, which means that we need to estimate the potential first, in order to be able to calculate the unrealized potential.

As such, the first step consists of calculating a proxy of livestock value per hectare (ha) by multiplying livestock density by livestock prices, which gives us the observed value of livestock per ha in a given district. Subsequently, we estimate a potential value of density using a stochastic frontier model. The latter helps estimate the potential livestock density per ha using land use (pasture and arable area) and the unit price of the livestock as predictors. We do this using the stochastic frontier approach (Aigner, Lovell and Schmidt,1977; Meeusen and van den Broeck, 1977; Khumbakar and Lovell, 2000). Specifically, in our case we focus on a simple single-output frontier production function proposed by Aigner *et al.* (1977):

$$y_{ld} = f(\boldsymbol{x}_{ld}, \boldsymbol{\beta}) exp(v_{ld} - u_{ld}) \quad (3)$$

where y_{ld} is the livestock density for livestock *l* in district *d*, $\boldsymbol{x}_{ld}$ is a vector of predictors at the district level related to livestock type *l*; u_{ld} is an inefficiency term; v_{ld} is a random error term; and, β is the parameter associated with variables x.

[19] District-specific head count poverty of 2021 (Uganda Bureau of Statistics, 2022).

Once the potential density is estimated, we multiply the estimated inefficiency term by the estimated potential to obtain a livestock density gap valued at market prices as follows:

$$UP_{ld} = Pdens_{ld} * u_{ld} \tag{4}$$

where UP_{ld} is the unrealized potential of livestock type *l* in district *d;* $Pdens_{ld}$ is the estimated potential livestock density of livestock *l* valued at market prices in district *d;* and, u_{id} is the estimated technical inefficiency for livestock *l* in district *d*.

Once this variable is estimated, we proceed analogously as with the general case (presented above for crops) by applying the iterative elimination procedure.

For livestock commodities, we use livestock densities from the World Livestock Gridded Dataset[20] (Gilbert *et al.* 2018). This method is applied to one of the seven analysed commodities, namely goats. As with other analysed commodities, goats were selected as they ranked high in terms of simulated cost-effectiveness in the economy-wide model in Sánchez, Cicowiez and Fontes (2022).

3.6 Special case 2: Applying the methodology when no data on potential yields are available

When data on crop potential yields are not available from GAEZ, the approach is very similar to the case of livestock. In other words, prior to calculating the unrealized potential, we need to estimate a variable that captures potential. However, this approach can only be implemented when one has a sufficiently large number of observations at the household level, which is our case as we use three waves of the Uganda National Panel Survey (UNPS), namely 2011–2012, 2013–2014 and 2015–2016. The slight difference relative to the livestock case is that we focus on a crop revenue gap, rather than a livestock density gap. In this case, the first step of the methodology (the estimation of potential) is very similar to the approach described in Maruyama *et al.* (2018).

Thus, we first derive a revenue frontier at the household level using stochastic frontier analysis. Then we extrapolate the estimated statistical relationship at the district level, using the following formulas:

$$y_{it} = f(x_{it}; \beta) exp\ (v_{it} - u_{it}) \tag{5}$$

$$u_{it} = \delta z_{it} \tag{6}$$

$$v_{it} = \exp(\gamma w_{it}) \tag{7}$$

where y_{it} denotes the commodity revenue of household i in year t; x_{it} is a matrix of independent variables (in our case, similar to Maruyama *et al.* (2018), we use commodity-specific unit prices, crop cultivated land area, long-term observed climatic conditions proxied by the Normalized Difference Vegetation Index-NDVI and/or cumulative rainfall, land-use variables, year and region fixed effects) for a household i in year t; v_{it} is a random error term associated with a set of exogenous variables w_{it}; u_{it} is an inefficiency term, which also depends on a set of exogenous variables, z_{it} such as household characteristics (family size, head age, education, assets, market access, altitude and NDVI/rainfall deviations); and β, δ

[20]Data were retrieved at the district level from https://dataverse.harvard.edu/dataverse/glw

and γ are the set of estimated parameters associated respectively with variables x, z and w. After estimating these parameters at the household level, we use small area estimation[21] to extrapolate the frontier and the inefficiency (calculated as one minus the technical efficiency) so that we have these values at the district level.

Once the potential and the inefficiency term are estimated, we can use the following formula to derive a measure of unrealized potential:

$$UP_{cd} = u_{cd} * PYield_{cd} \tag{8}$$

where $PYield_{cd}$ is the extrapolated potential yield frontier of commodity *c* in district *d*, u_{cd} is the extrapolated technical inefficiency (1 – technical efficiency) of commodity *c* in district *d*.

We then combine this with poverty data and apply the iterative elimination procedure to select a restricted set of districts.

3.7 Proposed robustness checks

Since this methodology is developed to support policy decisions, generally leading to investments with long-term impacts, we also propose several robustness checks to assess the stability of the results obtained whenever we implement the general case. Broadly speaking, we propose three main types of robustness checks, to test the reliability of the results to 1) changes in the definition of potential; 2) changes in the price level used; and 3) changes in the methodological approach.

With regards to the definition of agro-climatic potential, we use the historical climate model of the agro-climatic attainable yield of current cropland with low level of input use and under rainfed conditions (c_cruts_lr).[22] However, the agro-climatic potential depends heavily on assumptions related to the level of input-use, climate model used (including present vs. future potential), and assumptions on the management of efficiency. As such, we propose to also use alternative definitions of potential to understand how sensitive the results are to this change. Specifically, we re-run our estimation procedure using the following additional definitions of potential[23] (names in brackets below refer to the variable names based on variables names in GAEZ):

- Agro-climatic potential yield under low input, rainfed and historical CRUTS32 climate model (m_cruts_lr);

[21] Small area estimation refers to statistical techniques that involve the estimation of parameters for small sub-populations of interest (i.e. at the district level, in our case). The main idea, in our case, is to estimate a statistical relationship (stochastic frontier) for the sample as a whole and then use the parameters (the betas) and use the variation in the values explanatory across the sub-populations of interest to predict the outcome of interest.

[22] The reasons for choosing this model include: the management assumptions (rainfed and low input use), which are close to the current conditions of most farmers in Uganda; the historical model, which should give an accurate overview of the actual state of Ugandan agriculture; and moreover, the attainable yields which comprise the agro-climatic potential yields and combine them with the results of soil and terrain evaluation.

[23] The agro-climatic potential yield (defined by variables m_cruts_lr and m_clim_lr) are derived using an eco-physiological crop growth model. The results are then an agronomically potential yield under given agro-climatic, soil and terrain conditions and under specific management assumptions and agronomic input levels. These conditions also include soil moisture conditions together with other climate characteristics (radiation and temperature) during different crop development stages. On the other hand, the agro-climatic attainable yields (defined by variables c_cruts_hr, x_cruts_lr, and x_en_hr_oo) combine the agro-climatic potential yield with soil/terrain evaluation results, i.e. yield reduction factors due to the constraints induced by soil limitations and prevailing terrain-slope conditions (https://gaez.fao.org/pages/modules).

- Agro-climatic potential yield under high input, rainfed and CLIMATE (2010–2040) climate model (m_clim_lr);
- Agro-climatic attainable yield of current cropland under high input level, rainfed and historical CRUTS32 climate model (c_cruts_hr);
- Agro-climatic attainable yield of best occurring suitability class in grid cell under low input level, rainfed and historical CRUTS32 climate model (x_cruts_lr);
- Agro-climatic attainable yield of best occurring suitability class in grid cell under high input level, rainfed and ENSEMBLE (2010–2040) climate model, with CO_2 fertilization (x_en_hr_oo).

The second set of robustness checks relates to prices. Our price data come from the Uganda National Panel Survey (UNPS) surveys and, for some commodities (e.g. millet) the number of sellers in a given district may be low. Despite our best efforts to address this issue, there could be price outliers which may affect the results. In order to assess whether this is likely to be an issue, we also re-run the estimation procedure using the national median price, instead of district-level prices, and compare how different the set of results are.

Finally, we also assess the robustness of the results to using different methods. Specifically, in our case, whenever possible, we also run a crop-specific version of the approach proposed by Maruyama *et al.* (2018) to check how different the results would be – and by doing so, understand the benefits or advantages of our approach. In summary, there are three main differences between our approach and the approach proposed by Maruyama *et al.* (2018). First, our approach is run at the commodity level, rather than focusing on aggregate agriculture revenues. Second, for most commodities (whenever there are GAEZ data) the potential and unrealized potential are estimated in a deterministic, rather than stochastic manner and partially grounded on agronomic models. Third, rather than deriving a typology, the main aim of this methodology is to restrict the subset of districts of interest for policymakers, which is why we use the iterative typology construction, rather than devising a typology.

3.8 Strengths and limitations of the approach

As with every method, when evaluating the results of the methodology, it is important for the readers to be aware of what the methodology does and does not do, and where it represents an improvement on other existing approaches.

With regards to the strengths of this approach, we argue that it has several important strengths, listed below:

- **Simplicity of the approach** – First and foremost, this approach was developed to provide support to policymakers on locations with high theoretical potential for investments. As such, it was conceived such that it is simple to understand and easy to implement. In this regard, the data requirements of such a methodology are less onerous than other approaches that have been developed for geographical prioritization (Maruyama *et al.* 2018, 2019).
- **Flexibility of the approach –** As was shown earlier, our approach can be easily replicated and applied to different crop and livestock commodities, even if at times it may entail slight methodological deviations depending on data availability.
- **Commodity specific –** When going from higher-level planning and prioritization to implementations, policymakers inevitably need to make choices on selected commodities and types of investment. This method provides a tool to look at theoretical potential of

commodity-specific investments and, as done in this paper, it can be informed in a top-down fashion, with the selection of commodities being informed by a complementary analysis that ranks commodities based on the potential economy-wide effects (including both economic and public revenue growth) and poverty reducing potential of investing in these commodities at national level.

- **Cross-crop comparison –** Importantly, for the majority of crops (but not livestock commodities), the proposed methodology allows for a comparison across crops provided all the data that are required are available. This is important because in some cases, a given geographical unit may rank very high on a specific crop, but this crop may overall have limited potential compared to other crops. In such cases, in absolute terms, it may be better to invest in a different crop with higher overall potential, even if this specific geographical unit ranks lower in this crop.
- **Accounting for changing climate –** A critical aspect is that many investments (e.g. infrastructure) are long-term investments. As such, it may be important to have an idea of how conditions are expected to change and develop with climate change (i.e. optimal places today may not be optimal in the future). For certain crops, the approach developed in this document is able to incorporate how climate change is expected to impact agro-ecological potential.

This being said, this methodology also has its limitations – most of which are also shared by the alternative spatial prioritization approaches, specifically:

- **Unrealized potential is sensitive to the base year** – This approach provides a snapshot of the situation in the year for which data are available. While unrealized potential is likely to be highly correlated over time (at least for years close to one another), it can be affected by factors such as extreme weather events in that specific year. As such, it is key to discuss the results with stakeholders that are knowledgeable about the context where this method is being applied to avoid "obvious" (to local experts) wrong recommendations.
- **Crops and livestock cannot be compared** – Another issue is that, while the methodology, in principle, allows for a comparison across all crops included in GAEZ,[24] it does not allow for a comparison in two cases. First, it does not allow for a comparison between crops not covered in GAEZ with crops covered in GAEZ. Second, it does not allow for a comparison between specific crops and specific livestock commodities; they are measuring different things and units are not comparable. In the case of crops covered in GAEZ, the unrealized potential can be thought of as a "productivity gap valued at market prices". However, in the case of crops not covered in GAEZ, unrealized potential is given by a marketed revenue gap which may not be representative at the same geographical scale. In the case of livestock, it is given by a "livestock density gap valued at market prices", which is clearly not comparable to a yield gap.
- **Good data are key for good results** – While not a weakness of the approach itself, another issue worth mentioning relates to the quality of the data. Specifically, the data rely a lot on GAEZ data and price data. As such, if there are data quality issues in either (or both) of those data types, these will affect the end product.
- **Data needed does not get updated regularly –** Frequency of data updates is another issue. Unfortunately, GAEZ data do not get updated yearly and the latest update is from

[24] For certain crops, a conversion factor may be needed.

2015. Price data are often not collected frequently at the degree of granularity needed for this approach either. As such, if very rapid changes in yields have occurred since the last year available from GAEZ, the results may not be as current as would be desirable for policymakers. Similarly, if sub-regional level prices are not collected very regularly, this might force people applying such a methodology to either use older data, use either simplifying assumptions related to prices or drop prices altogether, none of which is ideal.

- **Fisheries and forestry are currently excluded –**This does not mean that these sectors cannot be included from a conceptual perspective, only that they are more challenging, and, to our knowledge, there is no readily available spatially distributed layer with a variable that can be used as a proxy for unrealized potential in both cases. In the case of forestry, the two main challenges relate to the fact that we are not aware of any available proxy that can be used for unrealized potential of timber/logging as well as sustainability concerns. For fisheries, the main challenge lies in the fact that fish stocks are not static and thus it is very difficult to estimate a spatially explicit measure of potential. For fisheries, one potential option for future research is to use sub-national statistics on outputs and inputs of fishery production (e.g. boats and catch volume value), which could serve as a proxy for potential and then estimate the unrealized potential using Data Envelopment Analysis (as there are likely to be a small number of observations).
- **Results emanating from the method need to be contextualized –** An additional limitation of the approach is that it requires contextualization to analyse the results. Given the way it is constructed, it does not incorporate important factors such as conflict in the analysis, as well as other variables that may be important when deciding where to invest. As such, it does not replace the knowledge of national experts, which remain key to help contextualize the results, which is why the results in this study were reviewed, discussed and validated by Ugandan Government experts.

4 Results

4.1 Main results

Table 1 shows the ranking that Sánchez, Cicowiez and Fontes (2022) proposed using their economy-wide, multisectoral-modelling analysis to prioritize investments across Uganda's agricultural sector. They ranked sectors according to the impact that the same public investment in productive infrastructure in each of them would have on four key variables: i.e. private consumption per capita, GDP, agrifood GDP, exports and rural poverty. We apply our spatial prioritization of commodity-specific investment approach to commodities produced by seven of sectors in the ranking[25] (i.e. sugar cane, bananas, maize, goats, millet, cassava and coffee),[26] which were selected according to any of the following rules:

1. **Rule 1: The commodity ranked first for at least one key variable**. This rule covers three out of the seven selected commodities, namely sugar cane, coffee and cassava. The exception to this rule is coffee, which ranks second for exports in the ranking that came out of the economy-wide model's analysis. However, it is included under this rule given that data are not available for tea (which ranks first in the export dimension). The inclusion of coffee is further justified by the fact that it was selected as a priority commodity under the NDPIII and is thus of high relevance to policymakers.
2. **Rule 2: The commodity was ranked in the first ten commodities across at least three dimensions.** This rule covers the remaining four commodities (i.e. bananas, goats, maize and millet). The only exception to this rule is the commodity group "vegetables", because, unfortunately, it is not possible to have a more disaggregated breakdown for vegetables in the dataset of the economy-wide model and given that our approach is commodity specific, we are unable to apply it to a category as broad as vegetables.

The commodities selected are shown in bold in Table 1.

[25] This rule was applied to keep the results of the work manageable. Selecting all the commodities that appear in the ranking would have resulted in 20 different commodities and one commodity group (vegetables).

[26] We note with interest that all but one of the selected commodities for this study were also included in previous prioritization exercises carried out by FAO under the framework of the MAFAP programme (MAFAP, 2013).

Table 1. Sectoral ranking by the impact of government investment on five socioeconomic indicators (only top-ten commodities are shown)

#	Private consumption	Gross domestic product (GDP)	Agrifood GDP	Exports	Rural poverty
1	**Sugar cane**	**Sugar cane**	**Sugar cane**	Tea	**Cassava**
2	Cattle	**Goats**	Sorghum	**Coffee**	Potatoes
3	**Bananas**	Cattle	Rice	Cocoa	**Sugar cane**
4	**Goats**	**Bananas**	**Coffee**	Vanilla	**Bananas**
5	Vegetables	**Maize**	Cotton	**Sugar cane**	Vegetables
6	**Maize**	Tea	**Millet**	Cotton	Beans
7	Potatoes	Simsim	Tea	Flowers	**Maize**
8	**Cassava**	Vegetables	Cocoa	Sorghum	**Millet**
9	Poultry	**Millet**	Soybeans	**Goats**	**Goats**
10	Beans	Groundnuts	Flowers	**Maize**	Poultry

Source: Sánchez, M.V., Cicowiez, M. & Pereira Fontes, F. 2022. *Productive public investment in agriculture for economic recovery with rural well-being: an analysis of prospective scenarios for Uganda*. FAO Agricultural Development Economics Technical Study. Rome, FAO. https://doi.org/10.4060/cb8730en

However, for the reasons explained in the previous section, it was not always possible to apply the general case approach to all commodities. More specifically, we could only apply the general case approach for five commodities. We therefore followed the decision-tree in Figure 2. The approach applied for each commodity is summarized in Table 2.

Table 2. List of commodities and approach used

	Subsector	Price data available	GAEZ data available	Approach used
Bananas	Crops	Yes*	Yes	General case
Cassava	Crops	Yes	Yes	General case
Coffee	Crops	Yes	No	Special case 2
Goats	Livestock	Yes	No	Special case 1
Maize	Crops	Yes	Yes	General case
Millet	Crops	Yes*	Yes	General case
Sugar cane	Crops	Yes*	Yes	General case

Notes: In the case of millet and sugar cane, national, rather than district prices were used due to a low number of observations and/or the presence of a large number of outliers. In the case of bananas, price data refer to 2013, rather than 2015 due to issues related to outliers in some regions of the country.

Source: Authors' own elaboration.

Figures 3–10 summarize the results for each commodity and the iterative elimination procedure,[27] while Table 3 shows the final list of selected districts. Each figure contains four maps. The first map denotes the overall potential of the different districts in the production of a commodity (current market value of the maximum attainable yields in a district), with darker shades of green denoting higher levels of potential. The second map, which uses shades of

[27] The full maps for each of the dimensions (potential, unrealized potential, and poverty) are available in the Annex (Figures A1–A3).

purple, denotes unrealized potential. In this map, using the subset of high potential districts (dark green in first map), we show the levels of unrealized potential, with districts shaded with darker purple indicating higher levels of unrealized potential. The third map includes the subset of districts kept in the first two stages and uses shades of red to map poverty. Districts in darker red indicate higher levels of poverty. Finally, the fourth map shows the five selected districts.[28]

Starting with the results for the **banana** sector (Figure 3), we notice that the potential is concentrated mostly in the central and western regions and this pattern is similar for the levels of unrealized potential (map with shades of purple). However, since poverty levels are more concentrated in the western parts of the country than in central areas, our final selection of districts includes three districts in the western region, and only two in the central region (see Table 3).

Turning to **cassava** (Figure 4), our data indicate that both the overall level of potential and unrealized potential for the production of this commodity are predominantly in the western and central regions. As a result, even when poverty rates are accounted for, we find that four out of the five selected districts are in these two regions (see Table 3).

For **coffee** (Figure 5), where the lack of GAEZ data forced us to use an alternative method (see Table 2), our estimates also indicate a higher potential in central and western regions, although there are a few districts in the eastern region with high levels of estimated unrealized potential. Once we combine this information with the poverty rates, three of the selected districts are in the western region, with the central regions containing two districts. It is important to highlight, however, that the method used in the case of coffee, entirely driven by a lack of data on agro-ecological potential, is likely to less accurately portray the agro-ecological attainable potential.

With regards to **goats**, estimates suggest that potential is scattered across the country, with pockets of high potential in the western region, eastern regions and some area is the northern region (Figure 6). The highest levels of unrealized potential, however, are concentrated in the western, eastern and the central regions. As a result, all of the identified districts are in these three regions (see Table 3).

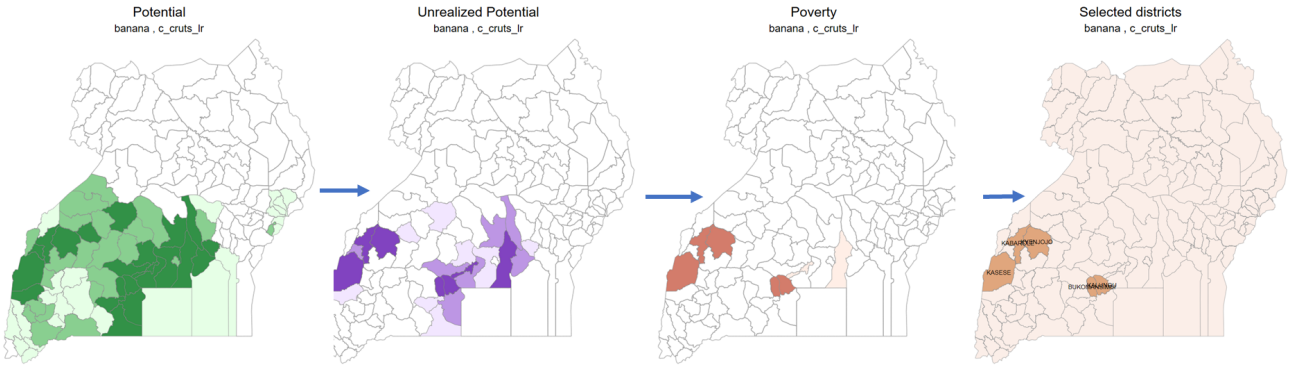
On **maize** (Figure 7), we note that while potential exists in almost every region, there is a concentration of high-potential districts in the central region. However, many of the districts in the central region have low unrealized potential. The unrealized potential is mostly concentrated in the eastern and northern regions, where higher levels of poverty are also prevalent. This explains why our five selected districts (see Table 3) are located in the northern and eastern regions.

Millet shows a very different pattern of potential compared to other crops (Figure 8). Specifically, the potential of millet is predominantly concentrated in the eastern region of the country and the same applies to the unrealized potential. As a result, all five selected districts are located in the eastern region, close to each other (see Table 3).

Finally, the results for **sugar cane** (see Figure 9) highlight high levels of potential mostly in districts in the Eastern region and the highest unrealized potential is concentrated in the eastern region in the districts surrounding the Jinja district, where all five selected districts are located (see Table 3).

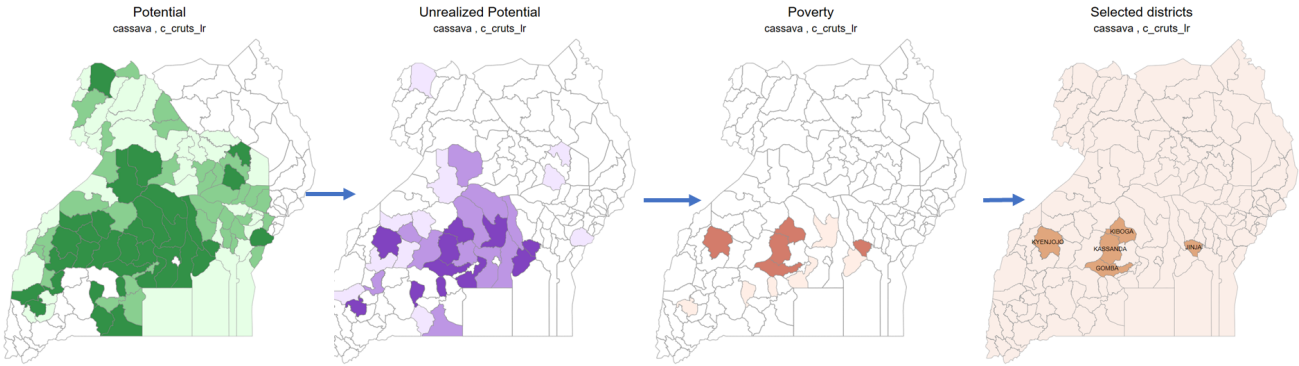
[28] We note that in some cases there may be three or four, rather than five, districts. In these cases, this is because there was a tie in the poverty rate across multiple districts.

Figure 3. Sequential selection of prioritized districts for banana production based on potential, unrealized potential and poverty



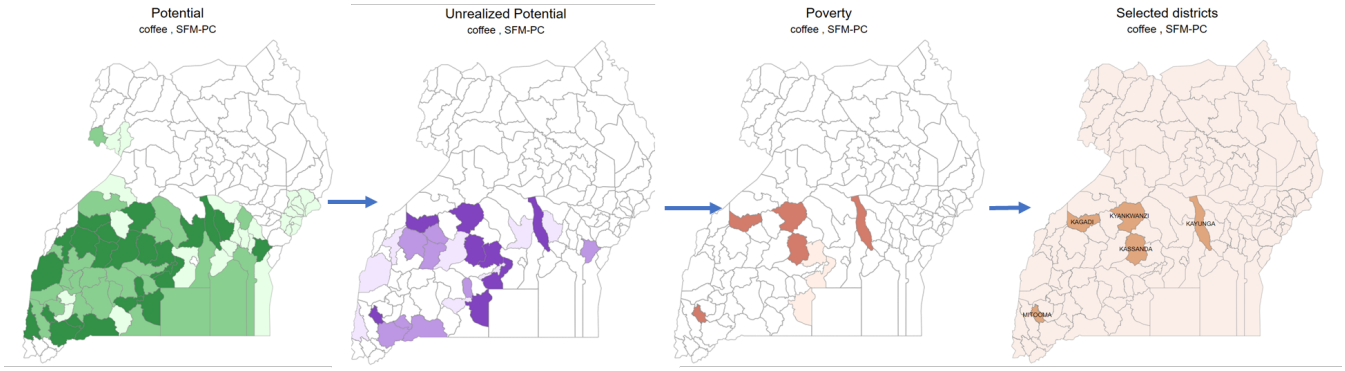
Source: OCHA (United Nations Office for the Coordination of Humanitarian Affairs). 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Figure 4. Sequential selection of prioritized districts for cassava production based on potential, unrealized potential and poverty



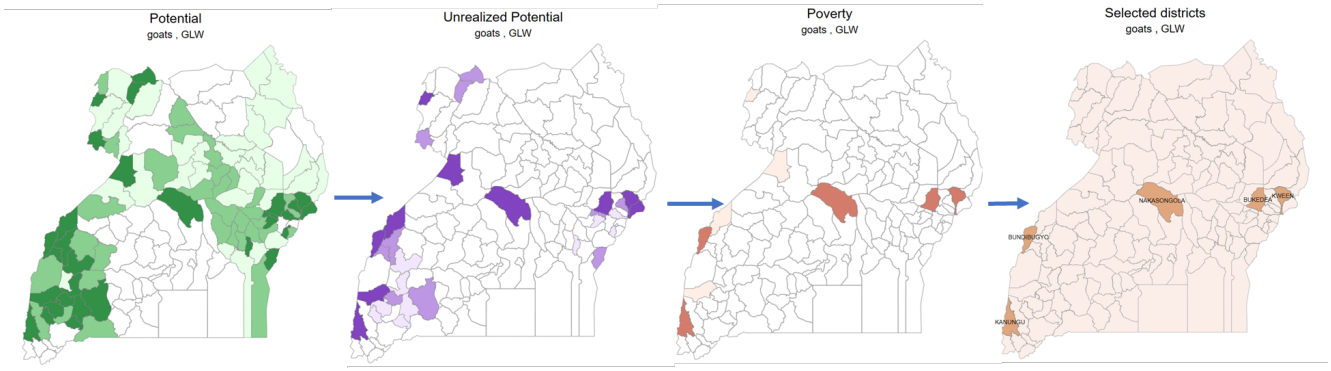
Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Figure 5. Sequential selection of prioritized districts for coffee production based on potential, unrealized potential and poverty



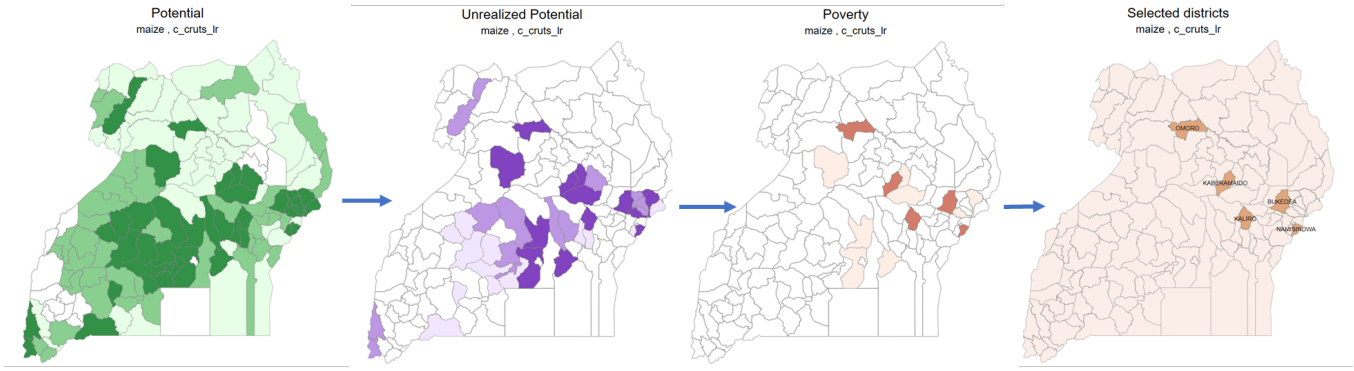
Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Figure 6. Sequential selection of prioritized districts for goat rearing based on potential, unrealized potential and poverty



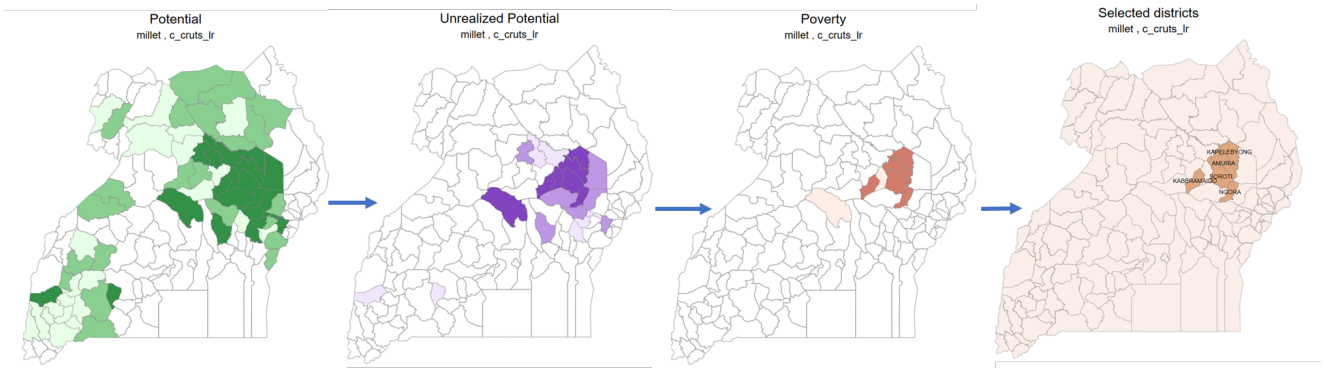
Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Figure 7. Sequential selection of prioritized districts for maize production based on potential, unrealized potential and poverty



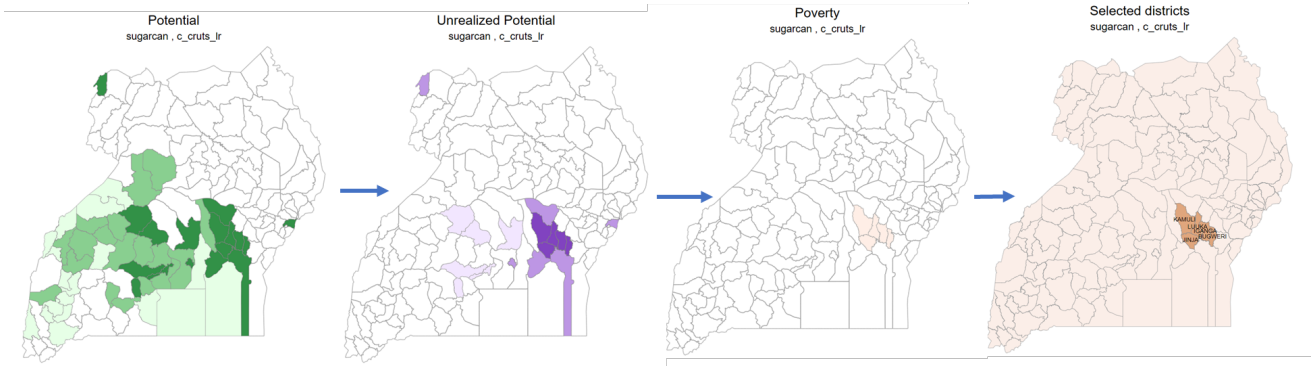
Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Figure 8. Sequential selection of prioritized districts for millet production based on potential, unrealized potential and poverty



Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Figure 9. Sequential selection of prioritized districts for sugar cane production based on potential, unrealized potential and poverty



Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Table 3. List of selected districts

Bananas [PC, GDP, PO]	Cassava [PC, *PO*]	Coffee [*Ex*, AGDP]	Goats [PC, GDP, EX, PO]	Maize [PC, GDP, EX, PO]	Millet [GDP, AGDP, PO]	Sugar cane [*PC, GDP, AGDP*, EX, PO]
Kasese (Western – Rwebitaba)	Gomba (Central – Mukono)	Kyankwanzi (Central – Mukono)	Kween (Eastern – Buginyanya)	Kaliro (Eastern – Buginyanya)	Ngora (Eastern – Nubin)	Bugweri (Eastern – Buginyanya)
Kabarole (Western – Rwebitaba)	**Kassanda** (Central – Mukono)	**Kassanda** (Central – Mukono)	**Bukedea** (Eastern – Nubin)	**Kaberamaido** (Eastern – Nubin)	**Kaberamaido** (Eastern – Nubin)	Luuka (Eastern – Buginyanya)
Kyenjojo (Western – Rwebitaba)	Kiboga (Central – Mukono)	Mitooma (Western – Mbarara)	Kanungu (Western – Kachwekano)	Omoro (Northern – Ngetta)	Soroti (Eastern – Nubin)	Iganga (Eastern – Buginyanya)
Kalungu (Central – Mukono)	**Kyenjojo** (Western – Rwebitaba)	Kayunga (Central – Mukono)	Nakasongola (Central – Mukono)	Namisindwa (Eastern – Bugninyanya)	Amuria (Eastern – Nubin)	Kamuli (Eastern – Buginyanya)
Bukomansimbi (Central – Mukono)	**Jinja** (Eastern – Buginyanya)	Kagadi (Western – Bulindi)	Bundibugyo (Western – Rwebitaba)	**Bukedea** (Eastern – Buginyanya)	Kapelebyong (Eastern – Nubin)	**Jinja** (Eastern – Buginyanya)

Notes: The acronyms in parentheses below the name of the commodity refer to the dimensions of the impact analysis using the economy-wide model in Sánchez, Cicowiez and Fontes (2022), namely private consumption (PC), gross domestic product (GDP), agrifood GDP (AGDP), exports (EX) and rural poverty (PO). If the acronym is placed below the commodity name, it means that this commodity was in the top-ten commodities where simulated investments were most cost-effective in terms of impacts in that dimension. When the acronym is in italics, it means that the commodity in question ranked first in a specific dimension. For instance, picking the case of cassava [PC, PO] means that cassava ranked in the top-ten in terms of private consumption and ranked first in terms of rural poverty reduction. Districts in bold (e.g. Kyenjojo) refer to districts that appear as a selected district for two different commodities (i.e. sugar cane and maize in the case of Serere).

Source: Authors' own elaboration.

4.2 Robustness checks

We also test the sensitivity of the results to different definitions of potential, changes in assumptions regarding prices, and changes in the way the potential and unexploited potential are generated (i.e. going from applying formulas [1]-[2] to estimating a crop-specific version of the Maruyama *et al.* [2018] approach using formulas [5]-[8]). These results are summarized in Tables A2, A3, A4, and A5, and Figure A4 in the Annex.

Overall, millet aside, we note that the results are quite robust to changes in the definition of potential used (first five columns in Tables A4a, A4b and A4c). This is the case both for the potential (Table A4a) and unrealized potential dimensions (Table A4b), which often leads to several common districts in the restricted list of five districts.[29] Even in cases where the districts differ, they are often in a similar geographical area to the five selected districts.

With regards to using a unique price, rather than considering variations in prices across districts, we note that, with the exception of maize, maps related to potential remain fairly similar (Table A4a). With regards to unrealized potential, the maps remain quite similar, except in the cases of maize and millet (Table A4b). However, since even a small change in the unrealized potential ranking can lead to a different final selection, there are some commodities (i.e. maize and millet) for which the selected five districts are either different in a similar geographical area (millet) or completely different (maize). In all other cases, we find that at least one or more (often two or more) common districts make it to the final list of five selected districts.

With regards to the sensitivity to using different approaches to estimating potential, we notice that the distribution of potential is quite similar (despite the difference in approaches and outcome of interest) for all commodities for which there is a fairly large number of observations (i.e. maize and cassava). For other commodities, however, especially those with few observations (e.g. millet), the results differ more. This is not necessarily surprising as the Maruyama *et al.* (2018) approach relies on an econometric estimation procedure, and thus requires a large number of observations to make the results meaningful. However, despite this, for the three commodities where the approaches are more comparable (i.e. maize, cassava and bananas), the results from the two approaches either include at least one common district in the final list of five districts or they at least tend to indicate districts that concentrated in roughly similar regions, although in the case of bananas, the identified districts in the western region are quite different.

We also calculated the shares of selected districts by region for each of the robustness check specification and compared those shares with the main specification (see Table A3). On average, the average shares of districts remain quite stable across most specifications. We also look at the number of specifications where a given district is selected and find that for commodities such as bananas, cassava, and sugar cane, all the selected districts are found at least once in the specifications of the robustness checks. For bananas, for example, the district of Bukomansimbi appears in six of seven robustness checks specifications, and for maize, the district of Bukedea appears in four out of six robustness checks (see Table A4).

Additionally, we also check the sensitivity of the results for the main specification to the permutation of steps two and three. The maps (see Figure A4) and list of selected districts

[29] For goats and coffee, these columns in the robustness checks are missing. The reason for this is that the alternative definitions of potential used are based on GAEZ and for these two commodities GAEZ data on potential were not available.

from this permutation (Table A4) show that, besides bananas, the main results are also quite robust to this change in the proposed methodology, and several similarities can be observed between the final maps. For cassava, coffee, and goats, switching the order of steps two and three results in three common districts (out of five) and in the cases of maize and sugar cane, changing the order results in four common districts. For millet two out of five districts are common. Finally, for bananas, although the selected districts are completely different, they are all located in the same regions as for the main results, namely the central and eastern regions.

5 Conclusions and discussion

When deciding on investments in agriculture, policymakers face difficult decisions, ranging from deciding on the targeted commodity selection, the investment type, to the geographic location of the investment. This is made even more challenging as these decisions are often made in data-scarce environments and the tools that exist often do not provide answers at the granularity that policymakers require to make these decisions.

In this paper we presented an approach designed to provide evidence to support policymakers in their decision on where to prioritize commodity-specific investments within the national geographic boundaries of their country. We applied this approach in the context of Uganda. To do this, we start by filtering out districts where the share of farmers cultivating the commodity is below a certain threshold, thereby reducing the likelihood of issues related to cultural acceptability of the commodity in selected districts. We then combine spatial data, alongside price data to measure commodity-specific unrealized potential at district level. Combining this measure of unrealized potential with existing poverty at district level, we then apply an iterative elimination approach in order to select, in a data driven way, a restricted number of districts that should be prioritized (i.e. five districts).[30]

Our methodology is conceptually similar to that proposed by Maruyama *et al.* (2018), but rather than focusing on maps for agricultural potential as a whole, we create commodity-specific maps. We argue that the approach proposed in this paper has several benefits compared to spatial approaches that focus on unrealized potential. First, whenever GAEZ data are available, the data needs to carry out the proposed approach are drastically reduced compared to the alternative approaches (Maruyama *et al.* 2018, 2019). Second, in most cases (i.e. "general case", as we call it in this paper), the method does not rely on estimations. This means that the approach overcomes issues related to small samples and sample representativeness which often limit the usefulness of applying econometric-based methods (e.g. Maruyama *et al.* 2018) for spatial prioritization. Third, the proposed approach allows to factor in some dimensions of predicted climate change effects and allows comparison across several (but not all) commodities. As a result, this allows policymakers to take into account climate change induced predicted changes to agro-climatic potential yields, which is especially important when considering multi-year investments such as in infrastructure. Finally, another advantage of the way in which we have applied this approach is that we benefited from a recently conducted economy-wide analysis for Uganda, which provides a ranking of the agricultural sectors where the same public investment in productive infrastructure would provide the highest economic and poverty reduction returns. Using this approach to focus on commodities that are produced by sectors selected from the said ranking (i.e. millet, maize, sugar cane, bananas, coffee, cassava, and goats) ensures that selected commodities are those where investments have the potential to have the largest macroeconomic impacts at the lowest cost.

In terms of policy recommendations, the results highlight that the spatial distribution of selected districts is very commodity specific, which reflects the heterogeneity in growing conditions and economic development in Uganda. We argue that this result is of particular relevance in the Ugandan context for two reasons. First, it highlights the importance of using commodity-specific approaches, as those approaches focusing on agriculture as a whole using revenues or profits (e.g. Maruyama *et al.* [2018]) would not bring to light such heterogeneity and areas

[30] We chose five for illustrative purposes, but the number of districts can be adjusted flexibly.

with harsher growing conditions (rainfall and temperature), for example, may not be selected, whereas some semi-arid regions are shown to be important for crops like millet. Second, in Uganda there seems to be a trend of moving policymaking closer to the people through decentralization and this is perhaps best epitomized by the decision to embark on a Parish Development Model, which was conceived as a model for effective provision and service delivery at the Parish level (Government of Uganda, 2021). While a model at such a decentralized level is laudable, it is also likely to be associated with challenges in terms of allocating resources, which is where spatial approaches capable of supporting policymakers can play a critical role. The approach proposed in this paper, while obviously not covering every important variable in the resource allocation process is one such tool, does allow for concrete results to support an evidence-based selection of prioritized areas.

Specifically, the proposed approach suggests that certain districts in **the eastern and/or northern regions of Uganda would be suitable candidates for further assessment of potential investments in the sectors producing millet, maize and sugar cane**.

In the cases of **bananas, coffee and cassava, however, suitable candidate districts are predominantly located in the western and central regions**, whereas for goats there does not seem to be a clear regional pattern, but rather pockets of districts across different regions.

Given potential shortcomings related to price data or the very definition of potential, we also undertake several robustness checks. Overall, we find a very similar picture in terms of the overall levels of unrealized potential when different definitions of potential are used. With respect to prices, our results are more sensitive, but the overall distribution of unrealized potential remains similar. Nevertheless, given that even small differences can affect the list of five districts, while for some commodities we find similar lists, there are commodities (e.g. maize) where the final list of districts is more sensitive to these different assumptions. Overall, however, the rough spatial location of selected districts remains relatively similar across the different robustness checks,[31] which provides some assurance on the spatial prioritization.

Beyond the results, it is also important to discuss how to go from the results of the approach and translating these into policies/investments. In our view, there are three aspects worth considering with regard to this point.

First, there is a need to ground-truth the approach, as while spatial data are continuously getting better, they are not perfect and the approach remains first and foremost a desk-based, data-driven methodology. As a result, the approach does not replace the role and the knowledge of local experts, who are very knowledgeable about the country context, but rather provides a tool to support their decision-making process.

A second aspect is that, while the methodology provides a useful first step towards identifying suitable areas for investments, it remains silent regarding the types of investment that are needed in a sector that produces a specific commodity at a given district. The information on the types of investment is critical for policymakers wishing to invest in specific districts and the next step of FAO's work seeks to fill this gap and carry out focus group discussions in five districts in Uganda to understand the types of investment that are most necessary in those districts.

[31] In the case of the Maruyama *et al.* (2018) approach, however, for commodities with few observations in the UNPS, we notice that discrepancies are larger. We argue that in this case the issue is likely to be related to sample size in the UNPS, which means that the Maruyama *et al.* (2018) approach, being an econometric method, is less reliable, as it relies on a large sample size to obtain an accurate estimate of potential. However, for commodities where the sample size is high in the UNPS (i.e. maize, cassava and bananas), the two approaches (our general case approach and Maruyama *et al.* [2018]) are more comparable.

Third, the approach presented in this paper is a simplification of reality and does not take into account all the equally important factors that are taken into account in deciding where to invest. This is mainly the case because each government gives different weight to different factors and because geographical layers may not be available for some factors (e.g. location on commodity-specific processing units).

We argue, however, that there are solutions to the three aspects highlighted above. To address the issue of ground-truthing and types of investment, once districts are selected, it is possible to organize visits to the districts to meet with key informants and value chain stakeholders to understand whether the final result of the approach makes sense and identify the main constraints and investment needs in these areas. Combining an economy-wide approach for commodity selection with this spatial approach and then carrying out open-ended interviews with stakeholders in these districts, therefore, has the potential to provide very concrete support to policymakers on policy-specific investments by identifying the commodity-district-investment combination that is most likely to have a high impact on performance.

To address the issue of embedding additional variables in the analysis, we argue that the approach proposed here is very flexible and that, if spatial data or a specific rule exist, in most cases, these can be factored in the approach as filters before the analysis or as a filter in the analysis. For instance, in our illustration, we notice that there are few districts selected from the northern region, which is the most arid region in Uganda and tends to be associated with lower potential yields.[32] From the perspective of unrealized potential, it is understandable as potential yields are likely to be higher in areas with higher rainfall. However, equity is often an important determinant of resource allocation for political or social reasons, then this can be embedded in the approach as an additional filter. Rather than keeping the top performers across the country, the approach could easily be tweaked to allow for a minimum number of districts per region by embedding this rule in the iterative elimination approach.[33] Similarly, important aspects such as whether geographic areas experience conflict or are protected areas (for environmental reasons) can all be embedded in the approach as a filter, as long as a layer exists with these data.

Ultimately, we believe that the approach we propose in this paper has the potential to be useful to inform commodity-specific spatial prioritization, while remaining sufficiently flexible to adapt to local contexts. It can also be implemented quickly, thereby making it particularly useful in policy-contexts where, more often than not, evidence needs to be generated quickly in order to feed into on-going or forthcoming policy processes.

By combining this approach with economy-wide models for commodity selection and the knowledge of local experts and the ultimate beneficiaries in the districts, policymakers can obtain concrete investment recommendations that identify investments in highly relevant districts that are aligned with the needs of farmers in selected districts and that target the sectors which have the highest potential impact on macroeconomic outcomes per USD invested.

[32] There are also other reasons why fewer districts from the northern region were selected in the final list of five, not least the commodity selection, with most of the chosen commodities (with the exception of maize, goats and cassava) being predominantly cultivated in other regions. Had sorghum been chosen, the northern region would have almost certainly been the region with the most selected districts.

[33] In practical terms, one would have to either add a rule that a minimum number of districts from each region are kept at each stage or, alternatively, run the iterative elimination procedure for each region separately and then select the desired number of selected districts from each region.

References

Aigner, D., Lovell, C.K., & Schmidt, P. 1977. Formulation and estimation of stochastic frontier production function models. *Journal of econometrics*, 6(1): 21–37.

Lobell, D.B., Cassman, K.G. & Field, C.B. 2009. Crop yield gaps: their Importance, magnitudes, and causes. *Annual Review of Environment and Resources*, 34(1): 179–204.

Diao, X., Fan, S., Kanyarukiga, S. & Yu, B. 2010. Agricultural Growth and Investment Options for Poverty Reduction in Rwanda. IFPRI Research Monograph. Washington, DC, IFPRI (International Food Policy Research Institute).

Dorosh, P. & Thurlow, J. 2018. Beyond agriculture versus non-agriculture: Decomposing sectoral growth–poverty linkages in five African countries. *World Development*, 109: 440–451.

Fan, S., Gulati, A. & Thorat, S. 2008. Investment, subsidies, and pro-poor growth in rural India. *Agricultural Economics,* 39(2): 163–170.

Fan, S., Yu, B. & Saurkar, A. 2008. Public Spending in Developing Countries: Trends, Determination, and Impact. *In* S. Fan, ed. *Public expenditures, growth, and poverty. Lessons from developing countries*, pp. 20–55. Baltimore, USA, Johns Hopkins University Press.

Fan, S. & Zhang, X. 2008. Public Expenditure, Growth and Poverty Reduction in Rural Uganda. *African Development Review*, 20(3): 466–496.

MAFAP (Monitoring and Analysing Food and Agricultural Policies). 2013. *Review of food and agricultural policies in Uganda*. MAFAP Country Report Series. Rome, FAO. https://www.fao.org/3/at594e/at594e.pdf

Fiala, N. & Apell, D. 2017. *Transforming Uganda's agricultural sector for sustained economic growth*. London, International Growth Centre.

Fischer, G., Nachtergaele, F.O., van Velthuizen, H.T., Chiozza, F., Franceschini, G., Henry, M., Muchoney, D. & Tramberend, S. 2021. *Global Agro-Ecological Zones v4 – Model documentation*. Rome, FAO. https://doi.org/10.4060/cb4744en

Government of Uganda. 2021. *The Parish Development Model: Community mobilization and mindset change operational manual*. Kampala. https://kamuli.go.ug/sites/files/NParish%20Model%20MANUAL-OCTOBER%202021-final%20draftNEW.pdf

IMF (International Monetary Fund). 2022. Uganda: 2021 Article IV Consultation and First Review under the Extended Credit Facility Arrangement and Requests for Modifications of Performance Criteria-Press Release; Staff Report; and Statement by the Executive Director for Uganda. Cited 4 May 2022. www.imf.org/en/Publications/CR/Issues/2022/03/15/Uganda-2021-Article-IV-Consultation-and-First-Review-under-the-Extended-Credit-Facility-515168

Ivanic, M. & Martin, W. 2018. Sectoral productivity growth and poverty reduction: National and global impacts. *World Development*, 109: 429–439.

Jayne, T.S., Mason, N.M., Burke, W.J. & Ariga, J. 2018. Taking stock of Africa's second-generation agricultural input subsidy programs. *Food Policy*, 75: 1–14.

Kumbhakar, S.C. & Lovell, C.A.K. 2000. *Stochastic Frontier Analysis*. Cambridge, UK, Cambridge University Press.

Ligon, E. & Sadoulet, E. 2018. Estimating the relative benefits of agricultural growth on the distribution of expenditures. *World Development*, 109: 417–428.

Gilbert, M., Nicolas, G., Cinardi, G., Van Boeckel, T.P., Vanwambeke, S., Wint, W.G.R. & Robinson, T.P. 2018. Gridded Livestock of the World (GLW). *In: Harvard Datavers.* Cambridge, USA. Cited 21 November 2021. https://dataverse.harvard.edu/dataverse/glw

Marivoet, W., Ulimwengu, J. & Sedano, F. 2019. Spatial typology for targeted food and nutrition security interventions. *World Development*, 120: 62–75.

Maruyama, E., Torero, M., Scollard, P., Elías, M., Mulangu, F. & Seck, A. 2018. *Frontier analysis and agricultural typologies*. ZEF-Discussion Papers on Development Policy No. 251. Bonn, Germany, Center for Development Research (ZEF).

Meeusen, W. & van Den Broeck, J. 1977**.** Efficiency estimation from Cobb-Douglas production functions with composed error**.** *International economic review,* 435–444.

Mogues, T., Yu, B., Fan, S. & McBride, L. 2012. *The impacts of public investment in and for agriculture. Synthesis of the existing evidence*. FAO-ESA Working paper No. 12–07. Rome, FAO.

Mulugo, L., Kibwika, P., Kyazze, F.B., Bonaventure, A.O. & Kikulwe, E. 2022. The contestations of diversity, culture and commercialization: why tissue culture technology alone cannot solve the banana Xanthomonas wilt problem in central Uganda. *Agriculture and Human Values*, 1–18.

NPA (National Planning Authority). 2020. *Third National Development Plan (NDPIII) 2020/21–2024/25*. Kampala.

Ndaula, S., Matsiko, F. & Sseguya, H. 2019. Farmers' multidimensional beliefs in orange-fleshed sweetpotato acceptance among rural households in Uganda. *Advances in Agricultural Science*, 7(2): 100–112.

OCHA (United Nations Office for the Coordination of Humanitarian Affairs). 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga

Pernechele, V., Fontes, F., Baborska, R., Nkuingoua, J., Pan, X. & Tuyishime, C. 2021. *Public expenditure on food and agriculture in sub-Saharan Africa: trends, challenges and priorities*. Rome, FAO. https://doi.org/10.4060/cb4492en

Sánchez, M.V., Cicowiez, M. & Pereira Fontes, F. 2022. *Productive public investment in agriculture for economic recovery with rural well-being: an analysis of prospective scenarios for Uganda.* FAO Agricultural Development Economics Technical Study. Rome, FAO. https://doi.org/10.4060/cb8730en

Sánchez, M.V., Cicowiez, M. & Ortega, A. 2022. Prioritizing public investment in agriculture for post-COVID-19 recovery: A sectoral ranking for Mexico. *Food Policy*, 109: 102251. https://doi.org/10.1016/j.foodpol.2022.102251

UBOS (Uganda Bureau of Statistics). 2010. *Uganda Census of Agriculture 2008/2009.* Kampala.

World Bank. 2016. *The Uganda Poverty Assessment Report 2016 Farms, cities and good fortune: assessing poverty reduction in Uganda from 2006 to 2013*. Washington, DC.

World Bank. 2020. World development indicators 2020. In: *World Bank*. Washington, DC. Cited 20 June 2022. https://databank.worldbank.org/source/world-development-indicators

Annex

Table A1. Data sources

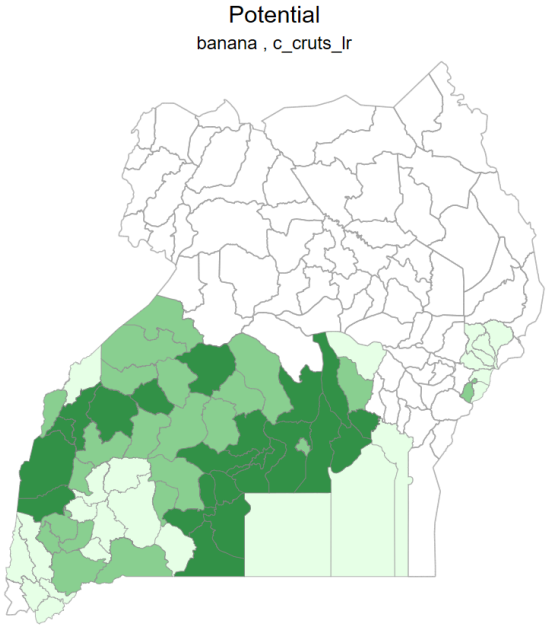
Layers	General case	Special case 1	Special case 2
Districts Administrative Boundaries	Retrieved from OCHA Humanitarian Data Exchange (OCHA [2020] with source shapefiles from the Uganda Bureau of Statistics), and adjusted to the United Nations official borders.	Retrieved from OCHA Humanitarian Data Exchange (OCHA [2020] with source shapefiles from the Uganda Bureau of Statistics), and adjusted to the United Nations official borders.	Retrieved from OCHA Humanitarian Data Exchange (OCHA [2020] with source shapefiles from the Uganda Bureau of Statistics), and adjusted to the United Nations official borders.
Filters	The share of households cultivating a given crop was computed at the district level using data from the 2008/2009 Agricultural Census (Uganda Bureau of Statistics, 2010).	The share of households rearing a given type of livestock was computed at the district level using data from the 2008/2009 Agricultural Census (Uganda Bureau of Statistics, 2010).	The share of households cultivating a given crop was computed at the district level using data from the 2008/2009 Agricultural Census (Uganda Bureau of Statistics, 2010).
Prices	District level median prices were calculated using household level commodity sale prices of the Uganda National Panel Survey (UNPS) of 2015–2016.	District level median prices were calculated using household level commodity sale prices of the Uganda National Panel Survey (UNPS) of 2015–2016.	Household level prices for the estimation of potential were retrieved from the Uganda National Panel Survey (UNPS) of 2011–2012, 2013–2014, and 2015–2016. District level median prices were calculated using household level commodity sale prices of the Uganda National Panel Survey (UNPS) of 2015–2016.
Potential	The potential yield was retrieved at the district level from the Global Agro-Ecological Zones (GAEZ) dataset. The GAEZ specification used is the historical climate model of the agro-climatic attainable yield of current cropland with low level of input use and under rainfed conditions.	Estimated using prices information and district level livestock densities. The livestock densities were retrieved from the Gridded Livestock of the World (GLW) dataset.	Extrapolated at a district level using a household revenue stochastic frontier model (like Maruyama *et al*. 2018) which was estimated with the Uganda National Panel Survey (UNPS) datasets of 2011–2012, 2013–2014, and 2015–2016. Independent variables included in the stochastic frontier model are commodity-specific unit prices, crop cultivated land area, normalized difference vegetation index, land-use variables, and year and region fixed effects.

Layers	General case	Special case 1	Special case 2
Unrealized potential	Calculated using the yield achievement ratios and the potential yields information. Both data were retrieved from the Global Agro-Ecological Zones (GAEZ) dataset.	Calculated using the estimated potential livestock density and the estimated inefficiency terms.	Calculated at the district level using the inefficiency terms of the estimated household revenue frontier. Variables included in the inefficiency model are the household characteristics (family size, head age, education, assets, market access), altitude and normalized difference vegetation index or rainfall deviations.
Poverty	District-specific head count poverty of 2021 (Uganda Bureau of Statistics, 2022)	District-specific head count poverty of 2021 (Uganda Bureau of Statistics, 2022)	District-specific head count poverty of 2021 (Uganda Bureau of Statistics, 2022)
Results	Computed by the authors and mapped using shapefiles retrieved from OCHA Humanitarian Data Exchange (OCHA, 2020) and adjusted to the United Nations official borders.	Computed by the authors and mapped using shapefiles retrieved from OCHA Humanitarian Data Exchange (OCHA, 2020) and adjusted to the United Nations official borders.	Computed by the authors and mapped using shapefiles retrieved from OCHA Humanitarian Data Exchange (OCHA, 2020) and adjusted to the United Nations official borders.

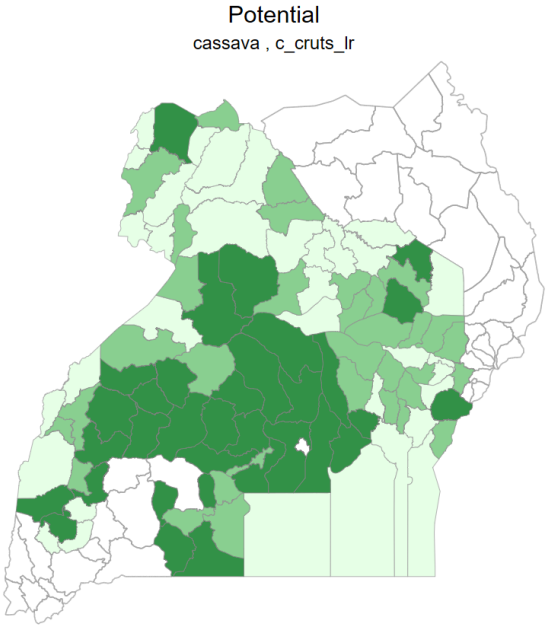
Source: Authors' own elaboration.

Figure A1. Results: Potential dimension

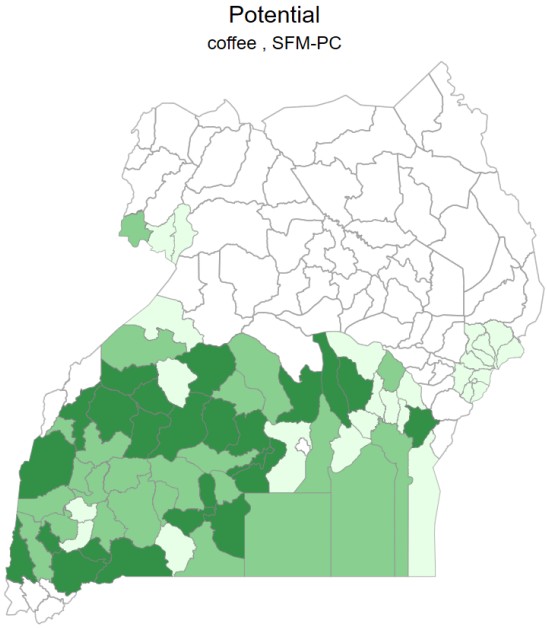
a. Bananas



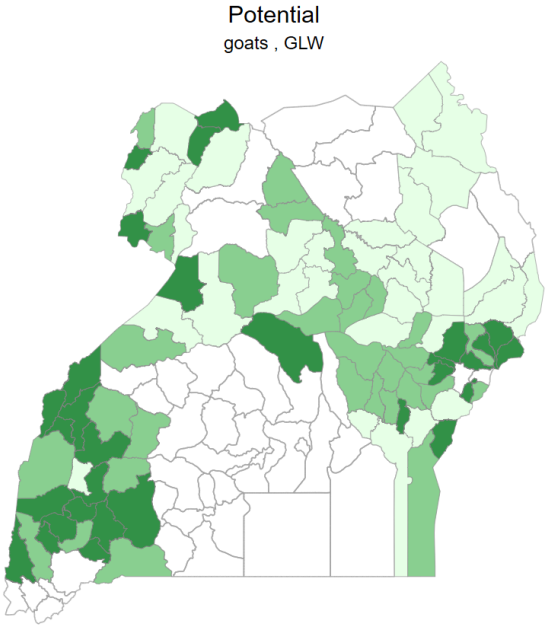
b. Cassava



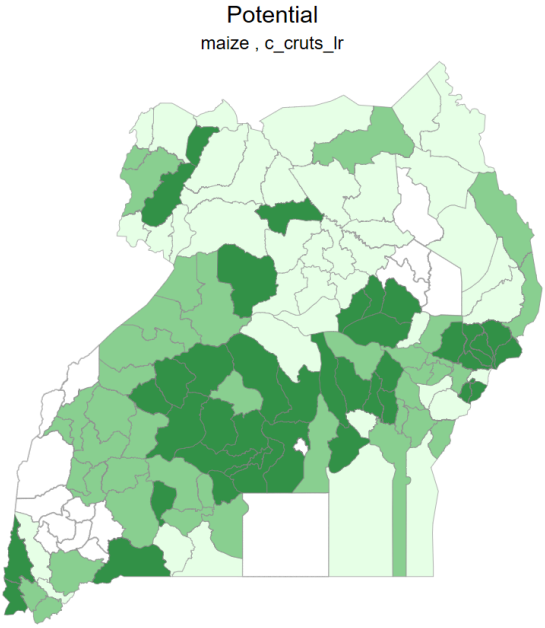
c. Coffee



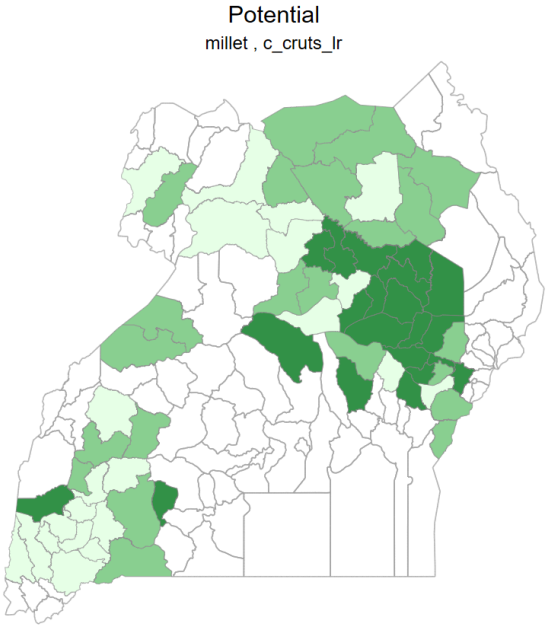
d. Goats



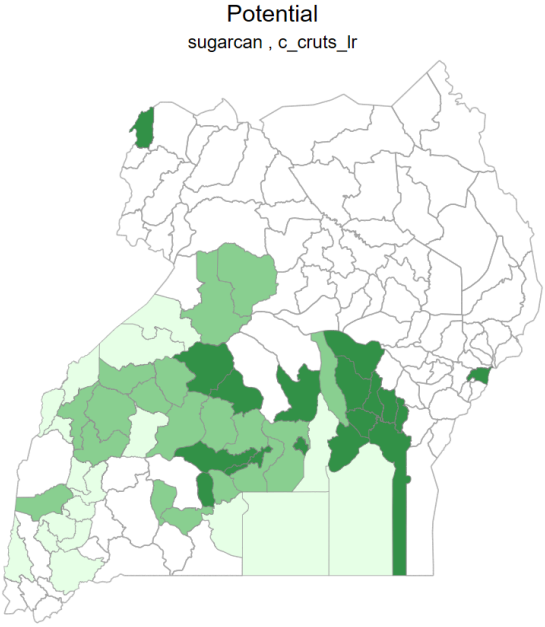
e. Maize



f. Millet



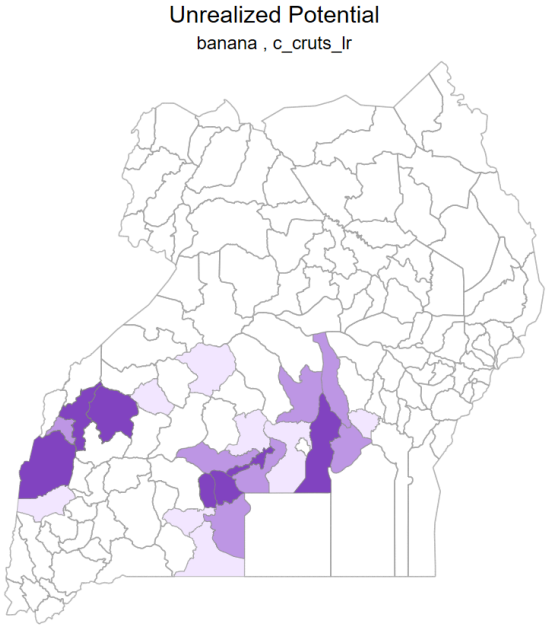
g. Sugar cane



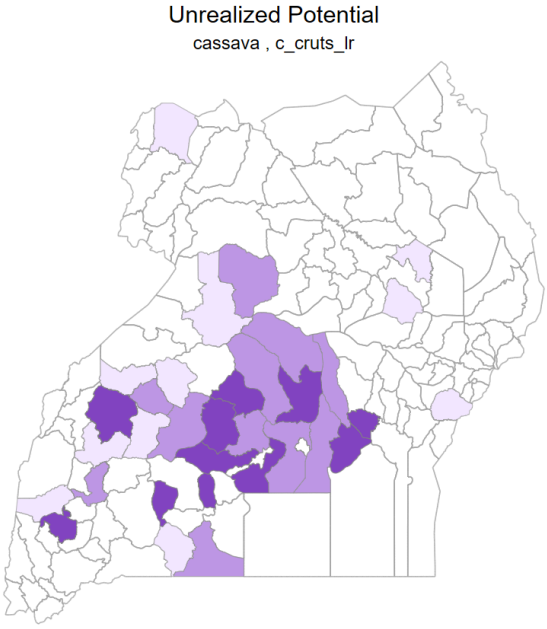
Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Figure A2. Results: Unrealized potential dimension

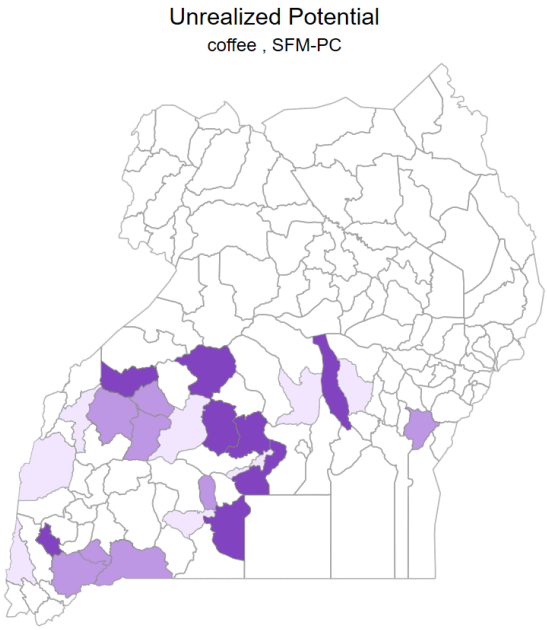
a. Bananas



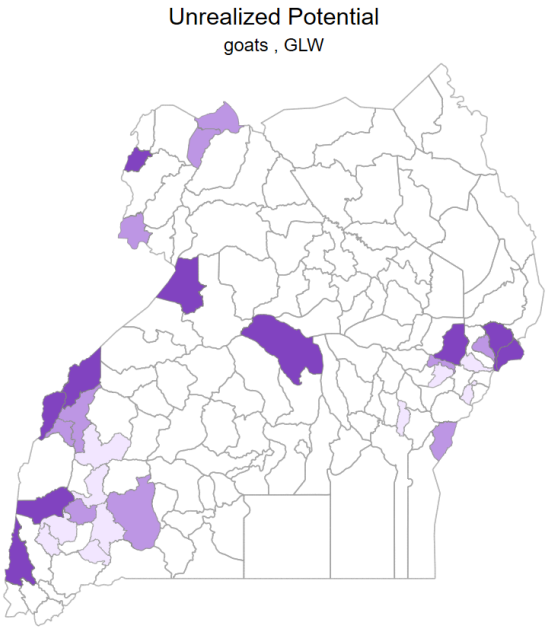
b. Cassava



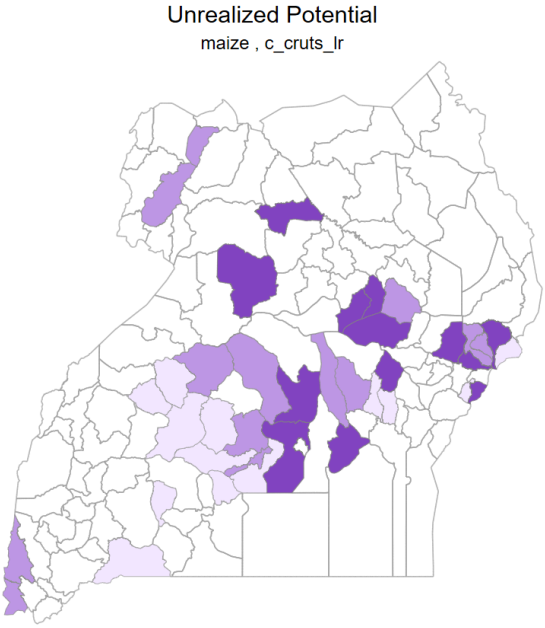
c. Coffee



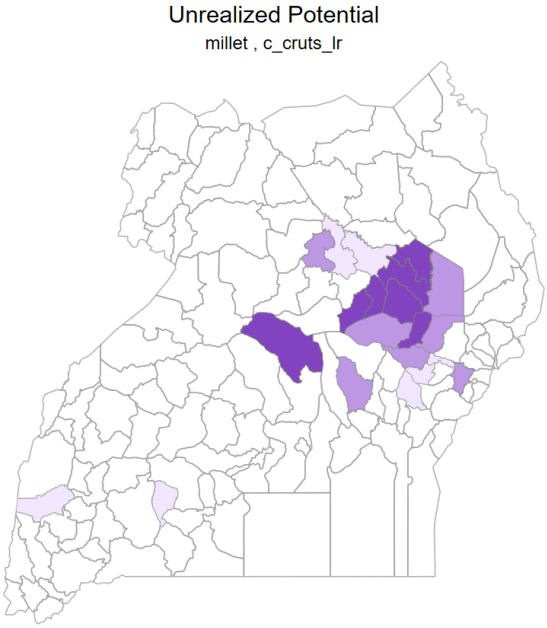
d. Goats



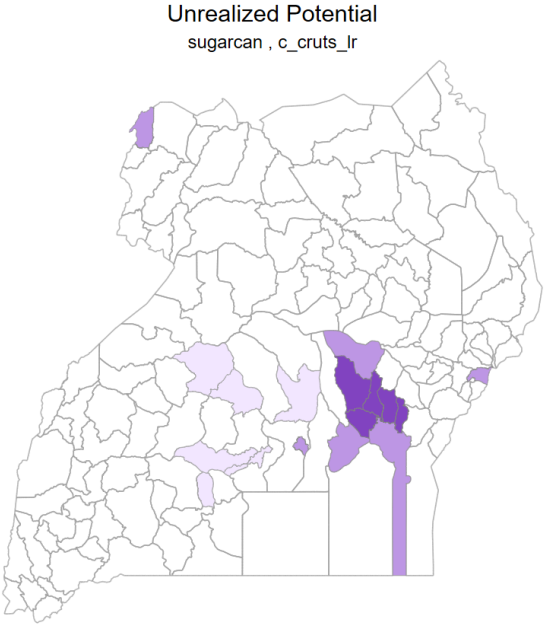
e. Maize



f. Millet



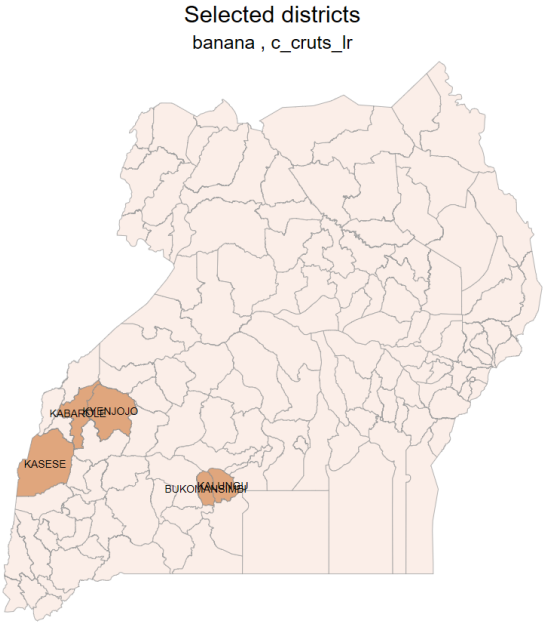
g. Sugar cane



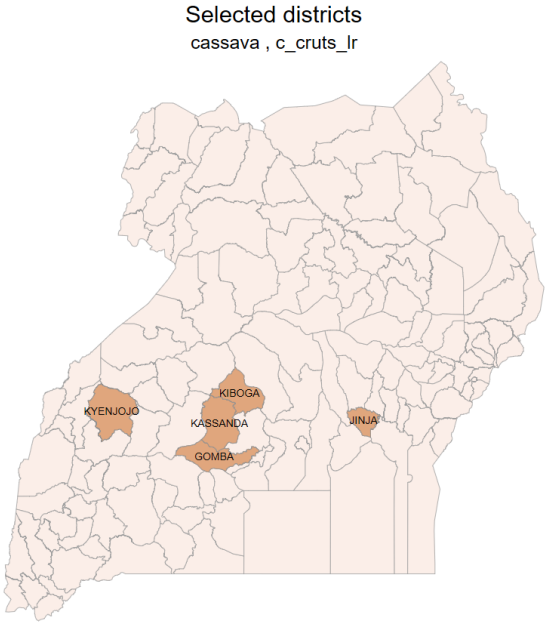
Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Figure A3. Results: Selected districts

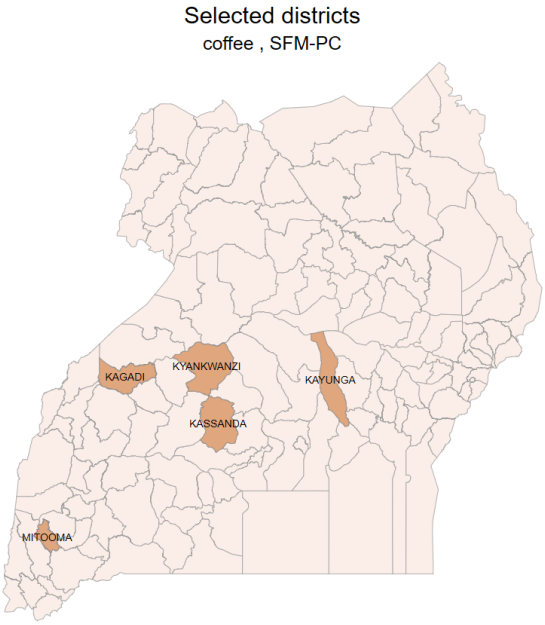
a. Bananas



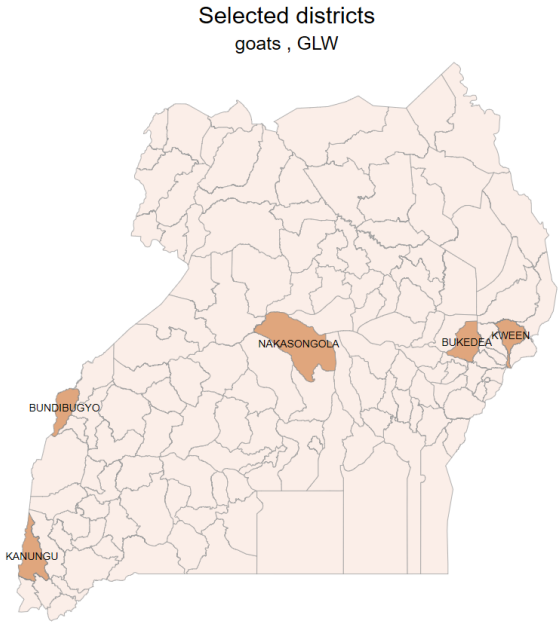
b. Cassava



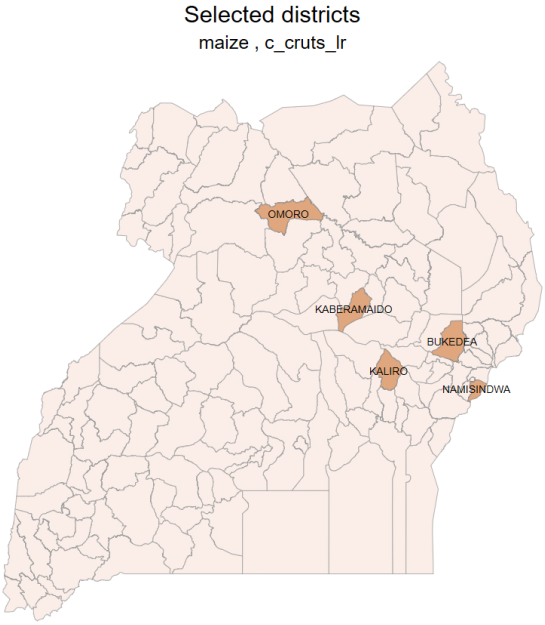
c. Coffee



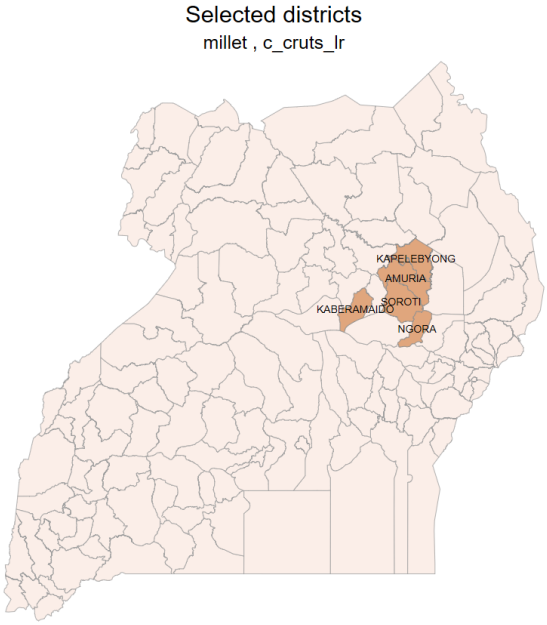
d. Goats



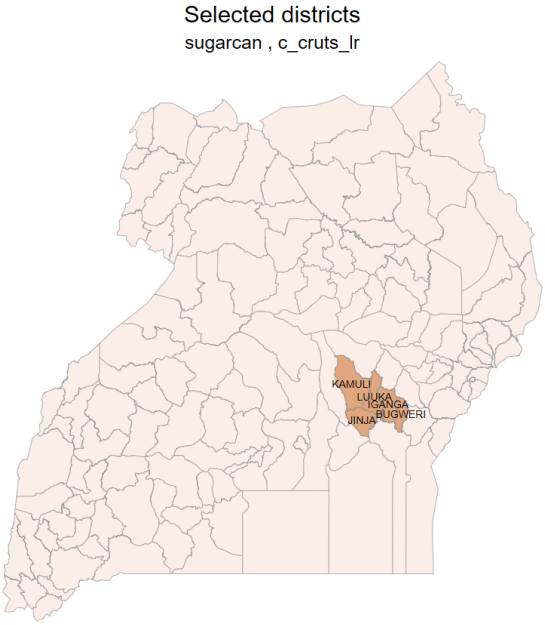
e. Maize



f. Millet



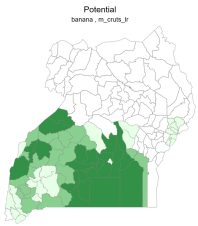
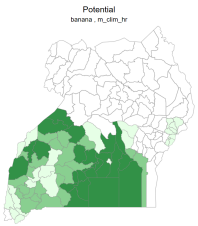
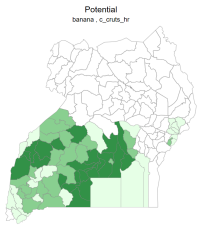
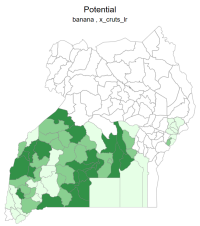
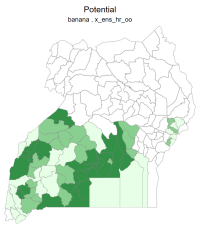
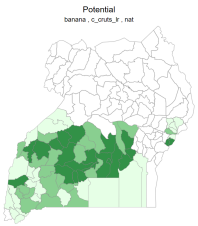
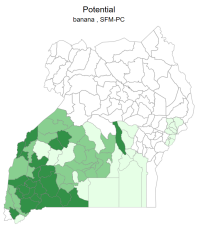
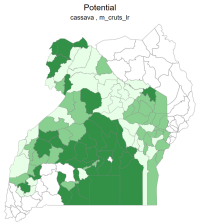
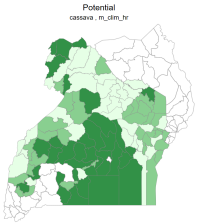
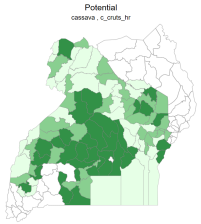
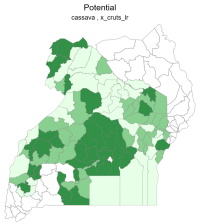
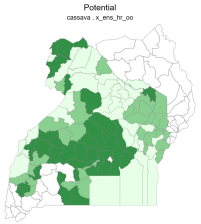
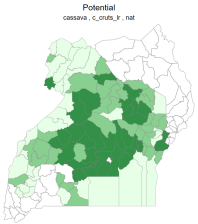
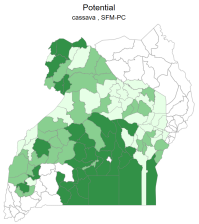
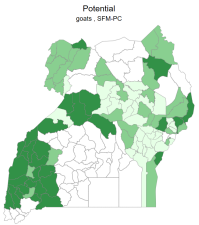
g. Sugar cane

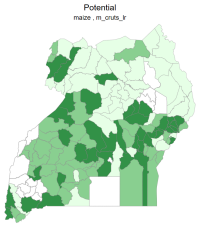
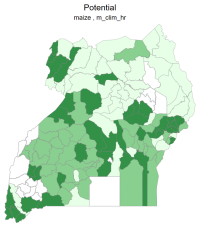
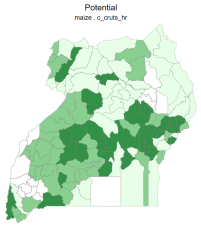
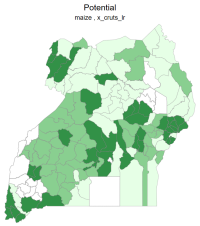
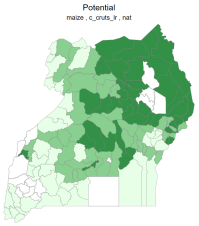
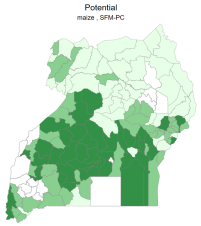
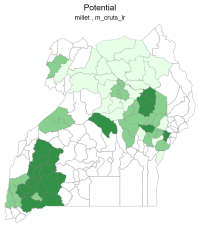
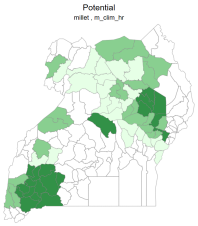
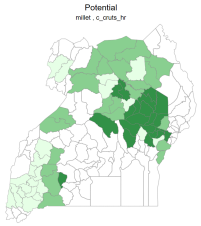
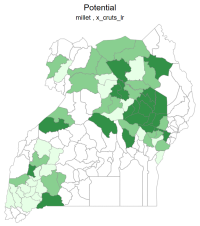
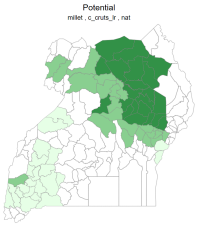
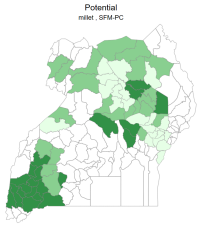
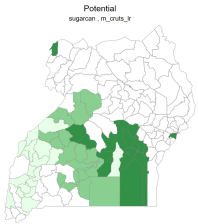
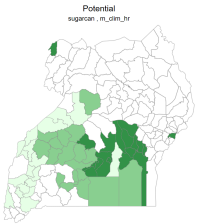
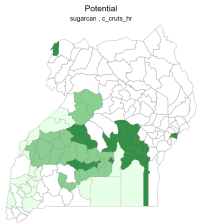
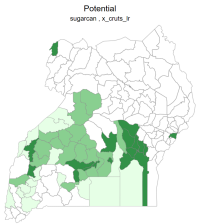
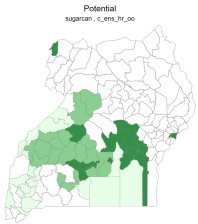
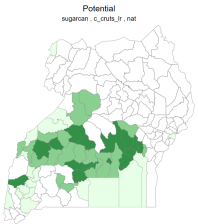


Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Table A2. Robustness checks

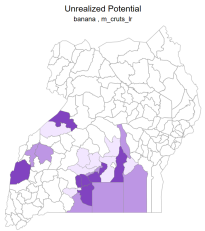
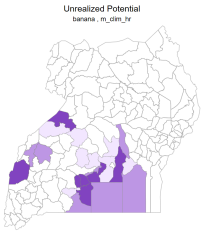
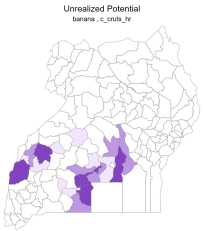
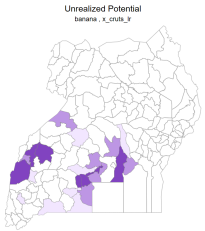
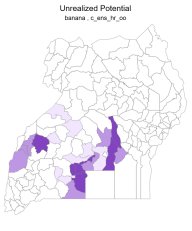
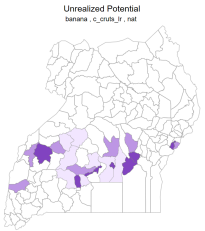
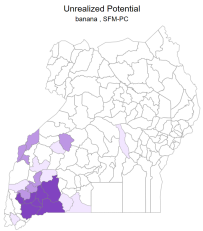
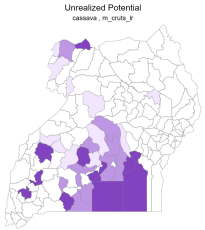
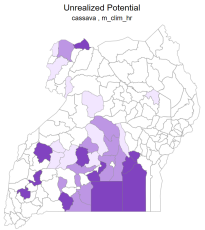
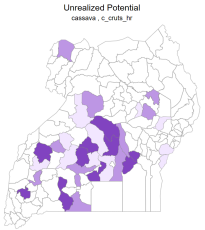
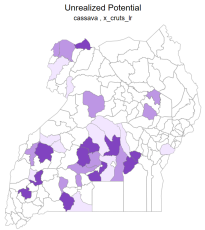
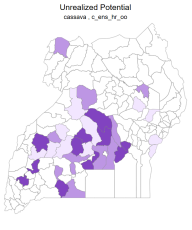
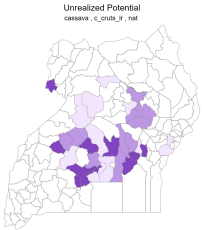
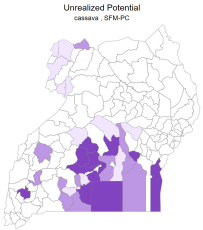
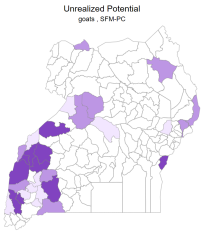
a. Robustness checks: Potential

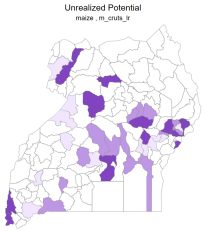
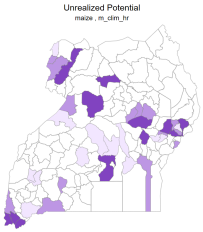
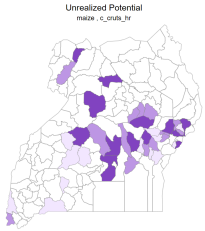
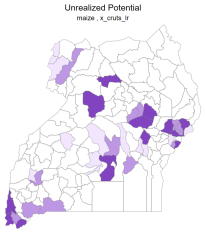
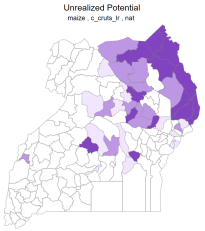
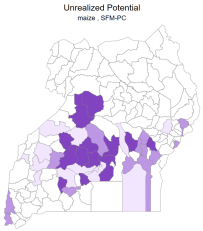
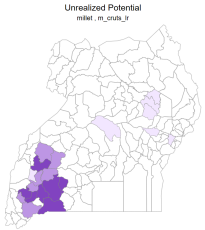
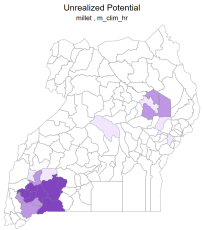
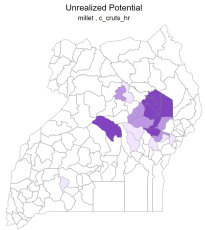
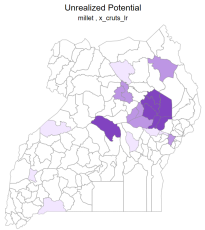
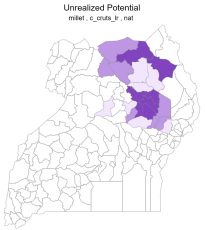
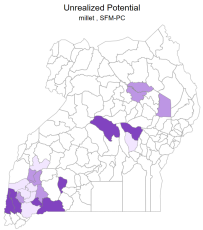
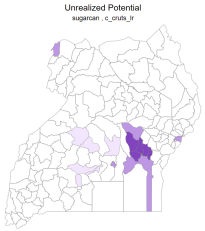
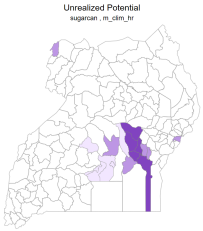
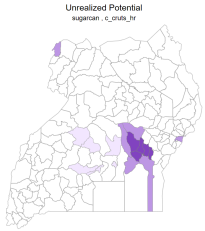
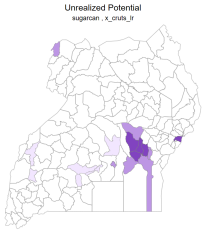
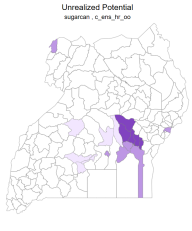
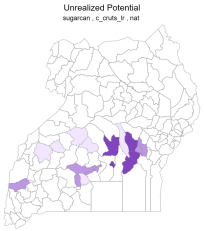
	Changes in definitions of agro-climatic potential yield					Using national prices	Maruyama *et al.* (2018)
	m_cruts_lr	m_clim_hr	c_cruts_hr	x_cruts_lr	x_en_hr_oo		
Bananas							
Cassava							
Goats							

	Changes in definitions of agro-climatic potential yield					Using national prices	Maruyama *et al.* (2018)
	m_cruts_lr	m_clim_hr	c_cruts_hr	x_cruts_lr	x_en_hr_oo		
Maize							
Millet							
Sugar cane							

Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

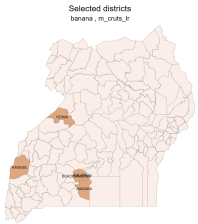
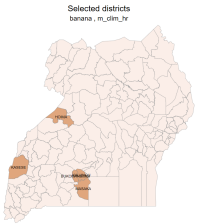
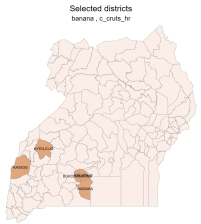
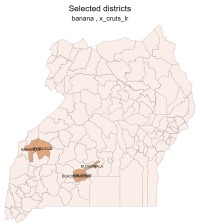
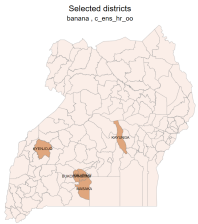
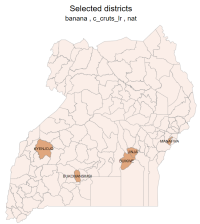
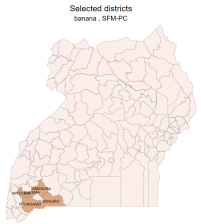
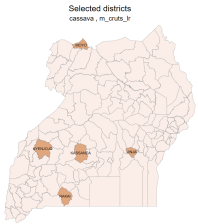
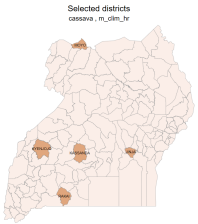
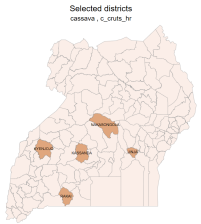
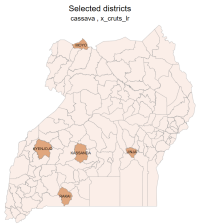
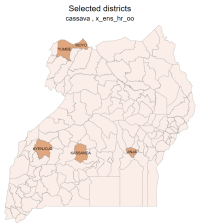
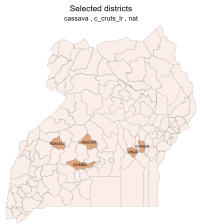
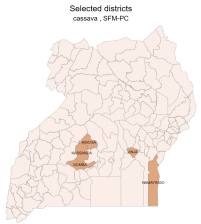
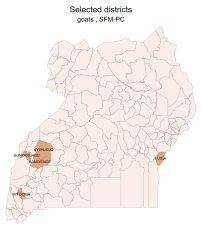
b. Robustness checks: Unrealised potential

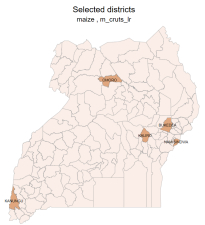
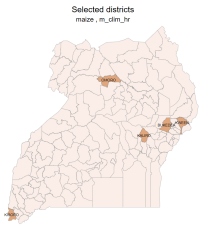
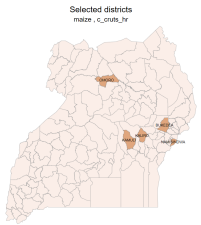
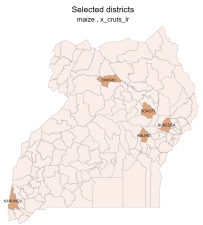
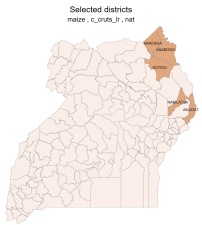
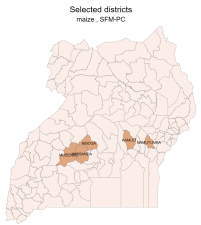
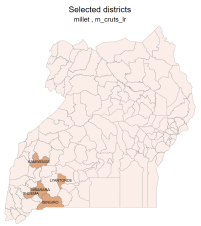
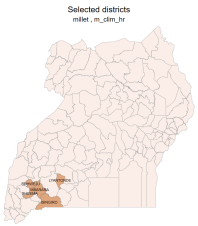
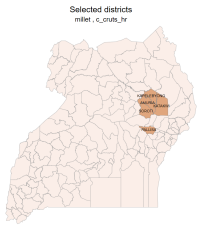
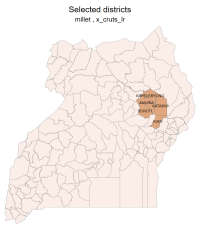
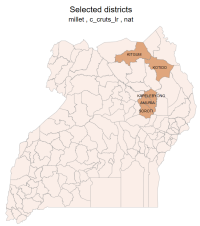
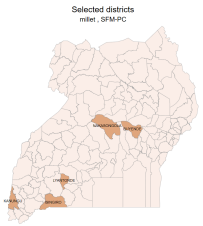
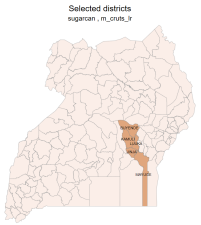
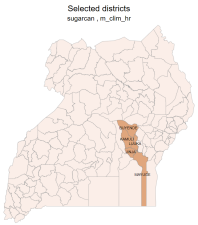
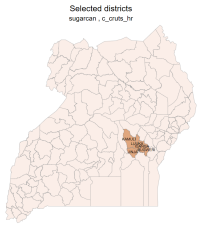
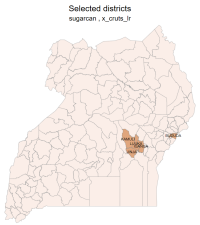
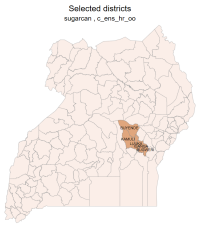
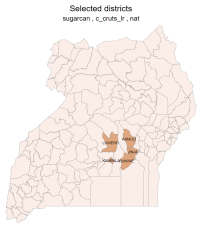
	Changes in definitions of agro-climatic potential yield					Using national prices	Maruyama *et al.* (2018)
	m_cruts_lr	m_clim_hr	c_cruts_hr	x_cruts_lr	x_en_hr_oo		
Bananas							
Cassava							
Goats							

	Changes in definitions of agro-climatic potential yield					Using national prices	Maruyama *et al.* (2018)
	m_cruts_lr	m_clim_hr	c_cruts_hr	x_cruts_lr	x_en_hr_oo		
Maize							
Millet							
Sugar cane							

Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

c. Robustness checks: Selected districts

	Changes in definitions of agro-climatic potential yield					Using national prices	Maruyama *et al.* (2018)
	m_cruts_lr	m_clim_hr	c_cruts_hr	x_cruts_lr	x_en_hr_oo		
Bananas							
Cassava							
Goats							

	Changes in definitions of agro-climatic potential yield					Using national prices	Maruyama *et al.* (2018)
	m_cruts_lr	m_clim_hr	c_cruts_hr	x_cruts_lr	x_en_hr_oo		
Maize							
Millet							
Sugar cane							

Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Table A3. Robustness checks: Shares of selected districts per region

	c_cruts_lr (main specification)	m_cruts_lr	m_clim_hr	c_cruts_hr	x_cruts_lr	x_en_hr_oo	Using national prices	Maruyama *et al.* (2018)	Always selected districts
Central	20.0%	24.0%	24.0%	20.0%	16.0%	33.3%	23.3%	31.4%	0.0%
Northern	4.0%	8.0%	8.0%	3.3%	8.0%	13.3%	30.0%	0.0%	0.0%
Eastern	60.0%	36.0%	36.0%	66.7%	56.0%	40.0%	40.0%	20.0%	66.7%
Western	16.0%	32.0%	32.0%	10.0%	20.0%	13.3%	6.7%	48.6%	33.3%

Source: Authors' own elaboration.

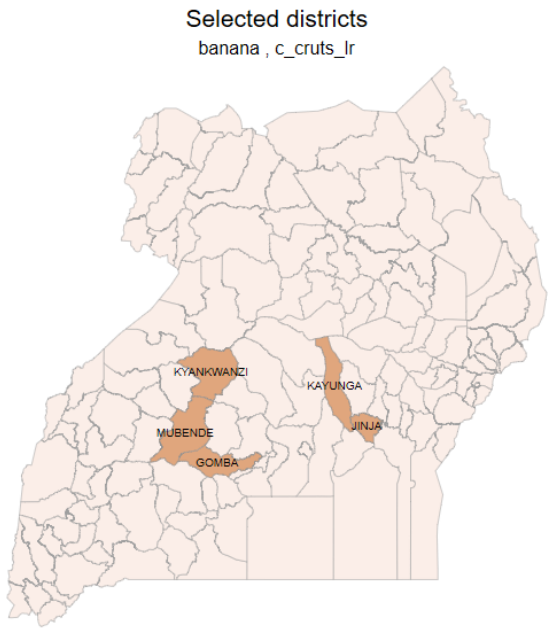
Table A4. Robustness checks: Selected districts and the number of specifications where they appear

Bananas [PC, GDP, PO]	Cassava [PC, *PO*]	Goats [PC, GDP, EX, PO]	Maize [PC, GDP, EX, PO]	Millet [GDP, AGDP, PO]	Sugar cane [*PC, GDP, AGDP*, EX, PO]
Kasese (Western) **4 out of 7**	Gomba (Central) **2 out of 7**	Kween (Eastern) **0 out of 1**	Kaliro (Eastern) **4 out of 6**	Ngora (Eastern) **0 out of 6**	Bugweri (Eastern) **2 out of 6**
Kabarole (Western) **1 out of 7**	Kassanda (Central) **6 out of 7**	Bukedea (Eastern) **0 out of 1**	Kaberamaido (Eastern) **0 out of 6**	Kaberamaido (Eastern) **0 out of 6**	Luuka (Eastern) **5 out of 6**
Kyenjojo (Western) **4 out of 7**	Kiboga (Central) **2 out of 7**	Kanungu (Western) **0 out of 1**	Omoro (Northern) **4 out of 6**	Soroti (Eastern) **3 out of 6**	Iganga (Eastern) **3 out of 6**
Kalungu (Central) **5 out of 7**	Kyenjojo (Western) **5 out of 7**	Nakasongola (Central) **0 out of 1**	Namisindwa (Eastern) **2 out of 6**	Amuria (Eastern) **3 out of 6**	Kamuli (Eastern) **6 out of 6**
Bukomansimbi (Central) **6 out of 7**	Jinja (Eastern) **7 out of 7**	Bundibugyo (Western) **0 out of 1**	Bukedea (Eastern) **4 out of 6**	Kapelebyong (Eastern) **3 out of 6**	Jinja (Eastern) **5 out of 6**

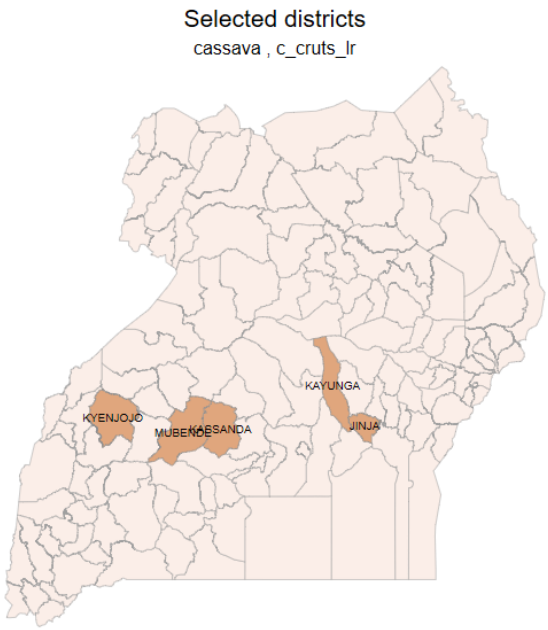
Source: Authors' own elaboration.

Figure A4. Robustness checks: Selected districts by permuting steps 2 and 3

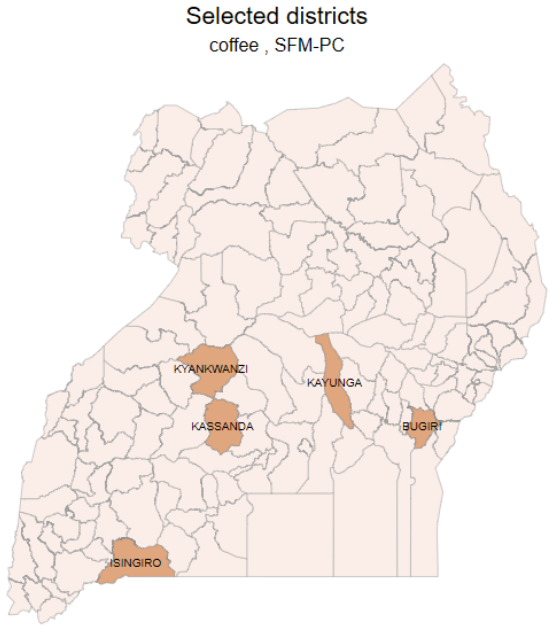
a. Bananas



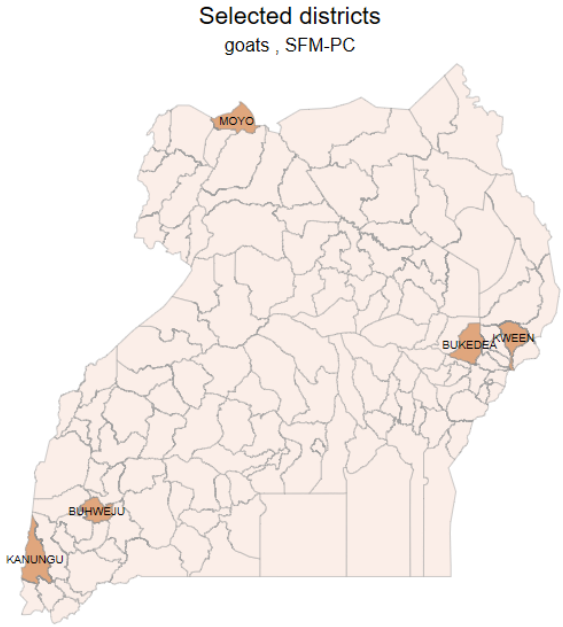
b. Cassava



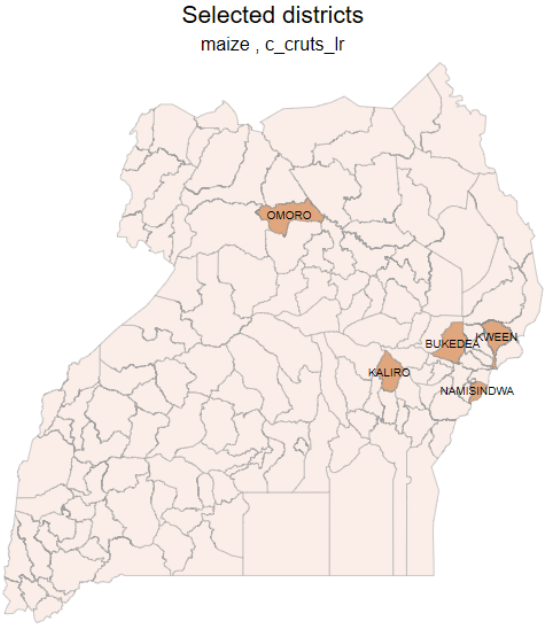
c. Coffee



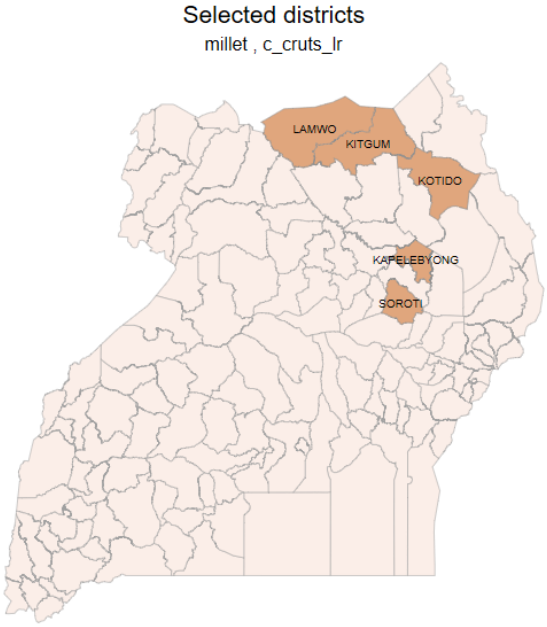
d. Goats



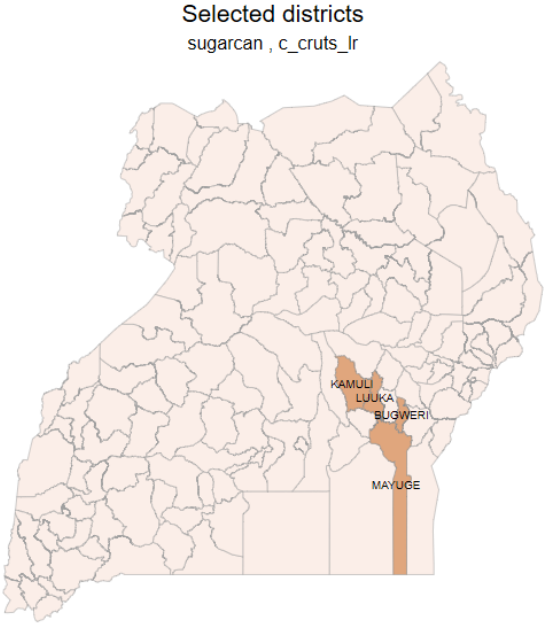
e. Maize



f. Millet



g. Sugar cane



Source: OCHA. 2020. Uganda - Subnational Administrative Boundaries. In: *OCHA | The Humanitarian Data Exchange.* Cited 12 December 2021. https://data.humdata.org/dataset/cod-ab-uga modified by the author.

Table A5. Robustness checks: List of selected districts by permuting steps 2 and 3

Bananas [PC, GDP, PO]	Cassava [PC, *PO*]	Coffee [*Ex*, AGDP]	Goats [PC, GDP, EX, PO]	Maize [PC, GDP, EX, PO]	Millet [GDP, AGDP, PO]	Sugar cane [*PC, GDP, AGDP*, EX, PO]
Mubende (Central - Mukono)	Mubende (Central - Mukono)	Bugiri (Eastern - Buginyanya)	Kween (Eastern - Buginyanya)	Kaliro (Eastern - Buginyanya)	Kotido (Northern - Nubin)	Mayuge (Eastern - Buginyanya)
Jinja (Eastern - Buginyanya)	Kayunga (Central - Mukono)	Isingiro (Western - Mbarara)	Buhweju (Western - Mbarara)	Omoro (Northern - Ngetta)	Lamwo (Northern - Ngetta)	Bugweri (Eastern - Buginyanya)
Kyankwanzi (Central - Mukono)	Kyenjojo (Western - Rwebitaba)	Kyankwanzi (Central - Mukono)	Bukedea (Eastern - Nubin)	Kween (Eastern - Buginyanya)	Kapelebyong (Eastern - Nubin)	Luuka (Eastern - Buginyanya)
Gomba (Central - Mukono)	Jinja (Eastern - Buginyanya)	Kayunga (Central - Mukono)	Moyo (Northern - Abi)	Bukedea (Eastern - Nubin)	Kitgum (Northern - Ngetta)	Kamuli (Eastern - Buginyanya)
Kayunga (Central - Mukono)	Kassanda (Central - Mukono)	Kassanda (Central - Mukono)	Kanungu (Western - Kachwekano)	Namisindwa (Eastern - Buginyanya)	Soroti (Eastern - Nubin)	

Source: Authors' own elaboration.

FAO AGRICULTURAL DEVELOPMENT ECONOMICS WORKING PAPERS

This series is produced by the Food and Agriculture Organization of the United Nations (FAO) since 2001 to share findings from research produced by FAO and elicit feedback for the authors.

It covers different thematic areas, such as food security and nutrition global trends and governance; food security and resilience; sustainable markets, agribusinesses and rural transformations; and climate-smart agriculture.

The complete series is available at:
www.fao.org/agrifood-economics/publications/working-papers

The Agrifood Economics Division (ESA) is the focal point for FAO's research and policy analysis on agricultural and economic development. The Division produces evidence-based policy analysis and strengthens the capacity of Member Nations to improve decision-making on food security and nutrition, resilience, climate-smart agriculture, sustainable markets, agribusinesses and rural transformations.

CONTACTS

Agrifood Economics Division – Economic and Social Development
ESA-Director@fao.org
www.fao.org/agrifood-economics

Food and Agriculture Organization of the United Nations
Rome, Italy

ISBN 978-92-5-137165-7 ISSN 2664-5785
9 789251 371657
CC2840EN/1/11.22