Assessment of the potential future sustainability of smallholder farming in the 'old cotton basin' of Mali

MSc thesis, Plant Production Systems

Dirk Jakob Hambuechen

August, 2019

WAGENINGEN
UNIVERSITY & RESEARCH

Assessment of the potential future sustainability of smallholder farming in the ‘old cotton basin’ of Mali

---

MSc thesis, Plant Production Systems

Name student	Dirk Jakob Hambuechen
Student ID	940915302080
Study	MSc Plant Sciences – Specialisation Natural resource Management
Chair group	Plant Production Systems
Code number	PPS80436
Date	30 August 2019
Supervisors	Dr. Katrien Descheemaeker, Dr. Gatien Falconnier
Examiner	Prof. Dr. Ken Giller

Disclaimer: This thesis report is part of an education program and hence might still contain (minor) inaccuracies and errors.

Correct citation: Hambuechen, D. J., 2019, Assessment of the potential future sustainability of smallholder farming in the ‘old cotton basin’ of Mali, MSc Thesis, Wageningen University, 103p.

Contact office.pp@wur.nl for access the data, models and scripts used for the analysis.

Table of content

Acknowledgement

I would first like to thank Katrien Descheemaeker for her support during all stages of the thesis, from the identification of the topic until the very end. I am especially, thankful for her elaborate feedback that taught me a lot about scientific writing and critical thinking. I also would like to thank Gatien Falconnier, whose research was the basis for this thesis. Thank you for your feedback and continuous support during the model building process. I enjoyed it very much to work with both of you! It was an instructive time.

I would further like to thank Eva Huet and Arouna Dissa for the helpful insights into the Malian agricultural and the provision of important data.

Further thanks to Bastiaen Boekelo who patiently taught me the basics in programming in R.

Lastly, I thank my examiner Ken Giller.

List of figures

List of Tables

List of Abbreviation

APSIM	Agricultural Production System Simulator
CMDT	Compagnie Malienne pour le Developpement des Textiles
fv	fodder variety
gv	grain variety
HRE	High resource endowed
HRE-LH	High resource endowed with large herd
LRE	Low resource endowed
NUE	Nitrogen use efficiency
MRE	Medium resource endowed
SI	Sustainable intensification
$PPP	US dollar purchasing power parity

Abstract

Rapid population growth in sub-Saharan Africa demands an intensification of the agricultural production. In addition, the key lessons drawn from the green revolution and low soil fertility in sub-Saharan Africa demand sustainable practices. An eminent solution is the concept of sustainable intensification (SI). The objective of SI is to increase agricultural production and simultaneously maintain or even increase the sustainability of the system at multiple domains. However, SI does not define specific pathways. Therefore, the aim of this study is to assess the state of representative farming systems in southern Mali in the near-term future (2027) based on biophysical and socio-economic trends in sub-Saharan Africa and thereby to identify promising pathways that enable SI. Accordingly, a model was developed to assess SI in the baseline situation and in six subsequent scenarios, based on incremental policy intervention and agricultural intensification strategies, for 411 smallholder farms in the ‘old cotton basin’ in southern Mali. The model checked for different SI indicators from four domains of sustainability. Under the assumption that intensification is the main objective of SI in sub-Saharan Africa three promising pathways were identified. Firstly, a successful promotion of family planning combined with the creation of job opportunities outside of agriculture reduced the pressure put by the rapid population growth on smallholder systems. Secondly, closing the yield gap up to 85 % of the water limited yield through different means distinctly intensified the system. However, trade-offs with the nitrogen use efficiency were identified. Lastly, the implementation of inventory credits for cereals increased the profitability but more importantly reduced farmers’ dependency on the cotton sector. Eventually, the research underlines that only a combination of multiple potential pathways can truly enable SI.

1 Introduction

1.1 Challenges for Future Food Production Systems

In the near future food production will face various challenges (Garnett & Godfray, 2012; Godfray et al., 2010; Tilman et al., 2011). Firstly, the rapidly growing world population demands a rising food production. By the year of 2050 the world population will exceed nine billion people (Godfray et al., 2010). Secondly, the demand for higher quality food like animal and processed products will expand because, worldwide, people's average income is increasing (Garnett & Godfray, 2012). The production of these products requires more natural resources than the production of a simple plant-based diet (Pretty, 2008). Thirdly, ensuring food security for the poor, especially in developing countries, requires a substantial local increase in food production. These three challenges demand a significant increase in food production until 2050 of about 70 % to 100 % (Godfray et al., 2010). Yet this rising demand for food must be fulfilled in an environmentally and socially sustainable approach which is the fourth challenge.

The competition for resources such as land, water, inputs and energy will rise (Garnett et al., 2013; Godfray et al., 2010). The negative effects of climate change and industrial input-focussed agriculture demand a rethinking of the current food production (Struik & Kuyper, 2017; Wezel et al., 2015). Godfray and Garnett (2014) describe food production as both *"the agent and the victim of environmental harms"*. Climate change, increase in greenhouse gas emissions, nutrient runoff, soil degradation, water shortage and biodiversity loss directly harm agriculture. The impact of climate change on agriculture can especially be seen in developing countries (Garnett et al., 2013). However, agriculture also strongly contributes to climate change as well as the other above listed issues (Howden et al., 2007; Garnett et al., 2013; Godfray et al., 2010; Struik & Kuyper, 2017).

This especially concerns the countries of sub-Saharan Africa which will face more often episodes of food insecurity based on the growing population and results of climate change (Smith et al., 2017). The population growth in sub-Saharan Africa is the highest globally (Tobergte & Curtis, 2013). Thus, the per capita availability of domestically produced food stagnated for 50 years in sub-Saharan Africa, even though, the agricultural production improved (Pretty, Toulmin, & Williams, 2011).

1.2 Agricultural Production and Challenges in Southern Mali

The rural areas of the Koutiala region in the south of Mali are characterised by smallholder farming systems like in most areas of sub-Saharan Africa (Harris & Orr, 2014; Kaya, et al., 2000). Farmers either consume self-produced crops or sell them on local markets, where additional food is bought as well. In Mali about half of the population lives below the poverty line of <1.9 US dollars per day (Jolliffe & Prydz, 2016). Farmers' low income per capita limits the opportunity of farmers to directly invest capital to prevent the degradation of the soils (van der Pol & Traore, 1993). Soil degradation is a major challenge for rural farming in southern Mali (Droppelmann et al., 2017). The increasing degradation of the soils is based on several events in the past. When cotton was introduced as main cash crop it was firstly intensively subsidised. From 1982 to 1986, however, the subsidies were reduced and thus farmers maintained the cotton production through extensification (Falconnier et al., 2015). Thereby, most arable land in the area was used and fallow periods were reduced. The farming systems turned into permanent cultures. Next to cotton, cereals are the most cultivated crops, namely maize, sorghum and millet. Nitrogen fixing legumes like cowpea and groundnut, however, are often just grown on small areas (Defoer et al., 1998). The use of inputs is often limited due to the widespread poverty (van der Pol & Traore, 1993). Additionally, a lack of adequate infrastructure in the rural areas limits the use of fertilizers. With the result that farmers tend to mine their soils (Defoer et al., 1998).

Thus, there is a need for sustainable intensification on the currently used farming areas to increase the productivity (Falconnier et al., 2016).

1.3 Sustainable Intensification

Sustainable intensification (SI) is an approach which is often discussed in literature (Garnett et al., 2013; Godfray et al., 2010; Smith et al., 2017; Struik & Kuyper, 2017; Tilman et al., 2011). The idea of SI for the future development of global agricultural practices was formed after recognition of the lack of social and environmental sustainability of the green revolution and industrial agriculture (Pretty, 1997). This is also reflected in the currently most commonly used definition of SI made by Pretty in 2008: *"Intensification using natural, social and human capital assets, combined with the use of best available technologies and inputs (best genotypes and best ecological management) that minimize or eliminate harm to the environment."* (Pretty, 2008; Wezel et al., 2015). For further clarification, Garnett et al. (2013) defined four underlying premises of SI:

1. *The need to increase production.*
   This aspect has been discussed already in section 1.1.

2. *Increased production through intensification and not through expansion.*
   Several authors showed that the expansion of cropping area has a more severe impact on the environment than intensification of the already used areas (Godfray, Garnett 2014; Phalan et al., 2011). An expansion of the cropping area may serve the aim to increase food production (Pretty et al., 2011) or the aim to decrease environmental impact of farming. The latter describes the concept of the land sharing which reduces the environmental impact of farming locally through de-intensification. This, however, creates a demand to use more farming area in total (Green et al., 2005). In sub-Saharan Africa, this is not possible due to area limitations (Jayne et al., 2014). Through expansion of cropping area, especially at the cost of forests, grass- or wetlands, many nitrogen, phosphorus and carbon stocks are eliminated and greenhouse gases are released into the atmosphere. Furthermore, biodiversity and ecosystem services of the area are substantially reduced (Godfray et al., 2014; Phalan et al., 2011). The alternative land sparing concept, which entails intensification on smaller cropping areas, enables preservation or even creation of natural habitats (Phalan et al., 2011). Thus, SI includes intensification as an approach for increasing food production (Garnett et al., 2013).

3. *Radical rethinking of food production to increase sustainability as well as food insecurity.*
   This premise states that reaching food security as well as environmental and social sustainability demand the same attention (Garnett et al., 2013). Therefore, often the existing farming system must to be radically reorganised. However, the approach to reach SI must be done in a context and location-specific way. For example, in highly developed areas the focus of SI needs to be more on the sustainability aspect than on intensification (Loos et al., 2014; Wezel et al., 2015). In sub-Saharan Africa the focus would lie more on intensification to overcome the challenge of achieving food security (Godfray et al., 2010).

4. *SI is an objective that is achievable through different approaches which are not clearly defined.*
   As described in the third premise, the approach is dependent on the circumstances in which SI takes place (Musumba et al., 2017). Often used methods are integrated pest management, the utilisation of legumes and mechanisation (Pretty, 1997; Tilman et al., 2011). However, SI is not restricted in its approaches as long as they are sustainable in their context and location (Garnett et al., 2013).

1.4 SI Indicators in Five Domains

The common way to assess SI is by applying a multi criteria assessment approach with SI indicators in different domains (Smith et al., 2017; Musumba et al., 2017). SI indicators are a way to measure a performance of a farming practice according to a certain objective.

(Musumba et al., 2017). They can be categorised in five different domains of SI: **productivity**, **economic sustainability**, **human well-being**, **environmental sustainability** and **social sustainability** (Smith et al., 2017; Musumba et al., 2017).

In this research productivity and economic sustainability are understood as domains which mostly describe intensification while the other domains focus more on the aspect of sustainability.

Snapp et al. (2018) emphasised that during the assessment of SI a holistic view is important. This entails the importance of using indicators from each domain to successfully assess SI. Otherwise, trade-offs between sustainability and intensification can hardly be identified (Smith et al., 2017). It is crucial for a successful SI assessment to choose the right indicators. Still the choice of indicators which are used is highly subjective (Olde et al., 2017; Marinus et al., 2016). Currently, indicators from the productivity and the environmental sustainability domain are mostly used in literature (Smith et al., 2017). Besides the classification of indicators in the different domains of SI they can be classified also on spatial scales: field, farm, household and landscape level (Musumba et al., 2017). For further information on how indicators are identified see 2.5.

1.5 Project Background

The research is part of the project: "**Pathways to agro-ecological intensification of mixed crop-livestock systems in southern Mali"**. In the following, the background of the project and the already conducted steps by Falconnier et al. (2015 – 2018) will be described.

In Mali cotton plays a key role as a source of income for most smallholder farmers. Malian farmers got institutional support for cotton production starting in 1970. From 1994 until 2010 farms of the Koutiala region, which lies in the cotton zone, were observed, classified and their farm development monitored. This timespan was divided in two periods. First, a favourable period for the cotton production until 2004 was observed. This period was followed by an unfavourable period until 2010 during which the institutional support collapsed.

Farms in the area were classified according to their resource endowment in four groups: High Resource Endowed farms with Large Herds (HRE-LH), High Resource Endowed farms (HRE), Medium Resource Endowed farms (MRE) and Low Resource Endowed farms (LRE). For further information about the classification criteria see (Table 10, Appendix). Furthermore, it was assessed how the farms developed during the observed period (Falconnier et al., 2015).

Subsequently, SI methods were tested on field plots over three years, from 2012 until 2014. These options included the introduction of cereal legume intercropping systems

(Falconnier et al., 2016). After these three years of research, the plot trials are continued until today but their results are not considered in this research.

Then, in a cyclic learning process it was attempted to redesign the farms from 2013 - 2015 to enhance farm productivity (Falconnier et al., 2017).

Lastly, a scenario analysis with five different scenarios for 2027 was conducted. The scenarios were based on future trends in socio-economic and biophysical conditions. This included marginalisation, a business as usual and two improvement scenarios. The main goal was to support the process of poverty reduction in the Koutiala region. For the scenario analysis two indicators were tested, namely food self-sufficiency and income per capita (Falconnier et al., 2018).

1.6 Scenario Analysis for Exploring Future Sustainability Pathways

To assess changes in SI over time it is necessary to check at different time points. Mostly, in literature this is done by comparing the current state of SI to the past. Common SI assessments also compare the sustainability of two farming practices. Checking multiple domains of SI ex-ante for future time points is a relatively new, but demanded approach as they assess prior to the implementation the sustainability of potential policy changes (Sadok et al., 2008; van Ittersum et al., 2008). A scenario analysis can create different future time points under certain assumptions (Schwartz, 1996 according to Duinker & Greig, 2007). Beside Falconnier et al. (2018) who used a scenario analysis, Swart et al. (2004) also advocate the utilisation of scenario analysis in the sustainability sciences.

Firstly, it needs to be defined what exactly a scenario is. In literature many definitions exist (Duinker & Greig, 2007). In this research the definition from Porter & Millar (1985) was used: *“A scenario is an internally consistent view of what the future might turn out to be - not a forecast, but one possible future outcome.”* This definition was chosen above all others because it clearly stresses the fact that a scenario is not an exact picture of the future. According to Duinker and Greig (2007), Schwartz (1996) stated that a scenario analysis is most efficient if different realistic scenarios are used and compared. It is important that the scenarios can be significantly distinguished from each other. Mostly, scenarios are used as a tool that supports the decision-making process (Swart et al., 2004).

1.7 Research Objectives and Questions

As already mentioned, a holistic view is important for SI (Snapp et al., 2018). A comparison of more indicators from different domains of SI (see section 1.4) enables a better understanding of the trade-offs of sustainability and intensification (Smith et al., 2017). Thus, the aim of this research is to extend the scenario analysis of Falconnier et al. (2018) and expand the model in order to include all SI domains based on a larger number of indicators. The research aims for the following objectives:

1. Develop a methodology to assess various domains of sustainable intensification for future scenarios.
    a. Define a formula for each identified indicator.
    b. Scale the indicators in a way that indicators with different units are comparable with each other.
2. Extend the scenario analysis of Falconnier et al. (2018).
    a. Extend the impact of the interventions so that they tackle all domains of SI.
    b. Add further scenarios.
3. Define the current and future sustainability based on all SI domains of the smallholder farms in the Koutiala region in southern Mali.
    a. Identify and assess trade-offs of SI.
4. Evaluate the effectiveness of the combination of SI assessment and scenario analysis as an approach for understanding the sustainability of farming systems in the future.

The following research question and sub-research questions can be derived from the objectives:

What pathways can enable sustainable intensification for farmers of Koutiala region in the near-term future?

1. How do the farms perform in the different domains of sustainable intensification based on the chosen indicators at both time points?
2. What are the key external and internal drivers that have the strongest impact on the sustainable intensification of the farms?
3. What are the trade-offs or interactions between the indicators and between the domains?

2 Materials and methods

The aim of the methodology is to design an approach for assessing the performance of the farms currently and in the potential near-term future in terms of SI. The methodology involves the main following steps: (1) Description of the farming systems and the definition of the main components; (2) Creation of a farm system model; (3) Identification of powerful indicators; (4) Scenario adaptation and creation; (5) Estimation of the current state of SI and the impact of the scenarios; (6) Aggregation of the results in a way that shows trade-offs and pathways to enhance farm performance in terms of SI.

2.1 Area Description

The Koutiala region is located in southern Mali (Figure 1). It is part of the ‘old cotton basin’. Farming systems of the villages Nampossela (99 farms) and Sirakélé (312 farms) were evaluated. The rainy season is from May until October and the rainfall fluctuates from 500 mm to 1200 mm. The population density reaches 70 people km^{-2} (Phoomthaisong and Toomsan, 2003, Soumaré et al., 2008).

Farmers grow mostly cotton and groundnut as cash crops and sorghum and millet as food crops. Maize serves as both a food and a cash crop. The harvest period is from September until November. Next to cropping, livestock is kept by farmers to store income, provide draught power for timely crop management and income generation by selling milk or animals. If the cattle herd is too large to feed during the dry period, farmers practise transhumance (for further information see 2.4.2). The most frequently kept animals are cattle, oxen, sheep and goats. Manure, mineral fertilizer and oxen as draught power are used to intensify the cropping systems.

A farm household regularly consists of the family with the head of the household including his brothers, sons, their wives and their children. After the death of the head of the household, the oldest son becomes the head (Jonckers & Colleyn, 1974). The household continues working together as a unit as the land is still owned by the multiple household members. This counters the subdivision of land that usually results in a decrease in land per capita (Falconnier et al., 2015). If the brothers of the head of the household disagree on the share of the income from cash crops, the land is split up and each member gets a share of the land.

Figure 1: Map of Mali; the Koutiala region is marked with a red circle (Google Maps).

2.2 Datasets and Data Collection

The farm characteristics of the current status were obtained from the surveys of the "Compagnie Malienne pour le Developpement des Textiles" (CMDT), which include all cotton farmers in the villages. The CMDT is a state agency that monitors the cotton production by conducting two different surveys, consisting of simple non-digital questionnaires. The first survey covers the area planning of cotton farmers for the following year in order to hand out input subsidies and hence is conducted annually. The second survey covers the farm characteristics to monitor the evolution of cotton producing households over time. This survey is not conducted annually.

For this research, data from the second type of survey from 2013 were used, as it was the most recent available digitalised data set. Only households with complete responses and sufficient data quality were included that allowed the calculations of the indicators, resulting in 99 responses from Nampossela and 312 from Sirakélé.

Based on crop area (ha), number of workers, herd size and number of draught tools the farms were classified in four different farm types following the typology of Falconnier et al. (2015). The distribution of the different farm types was 10 %, 21 %, 52 % and 17 % for HRE-LH, HRE, MRE and LRE farms respectively in Nampossela and 9 %, 22 %, 58 % and 11 % in

Sirakélé. In the analysis the farms were pooled. Differences between the villages were not identified.

2.3 Description of the Scenario Analysis by Falconnier et al. (2018) and the Extension Process

The scenarios of this research were based on the study by Falconnier et al. (2018), in which five scenarios were created for 2027. The scenarios were formed based on a combination of incremental policy interventions (P0 - P4) (Table 11, Appendix) and agricultural intensification strategies (A0 - A3) (Table 12, Appendix). Policy interventions, agricultural intensification strategies and scenarios (S0 – S4) built by Falconnier et al. (2018) are respectively summarised in the sections 2.3.1, 2.3.2 and 2.3.3. For more explicit information on the original scenarios see Falconnier et al. (2018).

The policy interventions were created based on observations of current trends in the local and global agricultural context. Additionally, policy solutions for current problems and current policy plans described in the literature were taken into account while building the policy interventions. The strategies for agricultural intensification were built based on promising agricultural technologies (Falconnier et al., 2018). In this research a similar methodology was used to extend the scenario analysis. The original scenarios mostly described interventions that affected the productivity and economic domains of SI. In this research further aspects were added to the interventions capture the corresponding effects on other domains of SI. This allowed to check a larger variety of indicators, thereby achieving a more holistic assessment. Additionally, another scenario was formed. The results of the scenario analysis extension are described in section 3.1.

2.3.1 Policy Interventions

The policy interventions were related to input and output prices as well as to socio-economic development and support to agriculture to close the yield gap.

2.3.1.1 Input and Output Prices

For the policy interventions (P1 – P4) the cotton price maintained the level of 2015 in the near future (Table 6) (Falconnier et al., 2015).

A negative policy intervention (P0) was formed based on the current low competitiveness of the Malian cotton sector on the world market and reduced fertilizer subsidies. It entailed lower producer prices (price the producers get for their products) for cotton and increased fertilizer prices (Table 6).

For the milk price optimistic policy trends were assumed in P2 – P4. High tariffs on imported agricultural goods (Dupraz & Postolle, 2013) and the preference of the local industry to use locally produced milk instead of milk powder (Aparisi et al., 2012), resulted in higher producer prices for milk (Table 6).

Furthermore, in the policy trends P2 – P4, a price reduction of cotton seed cake which can be used as livestock fodder was assumed (Table 6).

2.3.1.2 Socio-economic Development

In the policy interventions P3 – P4 the household size was reduced by two different means. Firstly, a successful governmental promotion of family planning in 2012 reduced the fertility rate of the households (Table 6). Secondly, the creation of job opportunities in urban areas by the government increased the rural to urban migration of young household members (Table 6).

2.3.1.3 Closing Yield Gap

Lastly, policy interventions that supported a closing of the yield gap to 85 % of the water limited yield for the cotton and the cereals were introduced in P4. Expanded fertilizer subsidies enabled the application of fertilizer to the point that nitrogen supply was no stress factor any more (Table 6). For cotton which is mainly harmed by the strong pressure of pests and diseases the use of Integrated Pest Management was supported as well (Table 6). Additionally, improved timeliness of crop management practices through small-scale mechanisation during crop establishment supported the competitiveness of cotton.

2.3.2 Agricultural Interventions

In the past decades the unfavourable cotton production conditions led farmers to replace cotton with sorghum (Falconnier et al., 2015). In the first agricultural intervention strategy (A0), a continuous decrease of the cotton area and a reduced fertilizer application on cotton were assumed (Table 6).

In A1 no changes in farm management were assumed (Table 6).

In A2 farmers started to intercrop maize with a fodder variety of cowpea (Table 6). This high quality fodder supported the introduction of stall feeding of lactating cows.

In the last agricultural intensification strategy (A3) practices that enabled farmers to close the yield gap were adopted. This involved increased application of fertilizer for cotton and cereals (Table 6) as well as the adoption of Integrated Pest Management and small-scale mechanisation for cotton.

2.3.3 Scenarios Created by Falconnier et al. (2018)

In the following, the existing scenarios created by Falconnier et al. (2018) are summarised.

The “Marginalisation” scenario S0 had a pessimistic view on the future and was formed by combining P0 and A0 (Figure 7). The scenario focussed on the marginalisation of the cotton production in the Koutiala region due to a reduction of input subsidies and a decrease in the cotton area.

In the “Business as usual” scenario S1 no changes were implemented by policy makers or farmers. It was formed through the combination of A1 and P1 (Figure 7). Only the household component was adjusted based on fertility and migration rates.

The “Dairy development” scenario S2 described the first optimistic view on the future. The combination of A2 and P2 formed this scenario. Promotions of the regional milk sector by policy interventions and diversification with legumes enabled an intensification of the livestock component of the farm.

The “Socio-economic development” scenario (S3) focused on the policy interventions (P3) in combination with A2 (Figure 7). S3 reduced the household size as it tackled the vast population growth in rural Mali through changes in fertility rates and rural to urban migration due to family planning and job creation.

The “Closing yield gap scenario” (S4) closed the yield gap of cotton and cereals. For this P4 and A3 were combined (Figure 7). This results in increased subsidies for fertilizer for cereals and Integrated Pest Management and small-scale mechanisation services for cotton.

2.4 Farm System Model

A simple model was developed calculating the various SI indicators to evaluate the baseline situation and demonstrate the impact of the different scenarios (Figure 3). This was based on interactions between three main farm components, the household, the cropland and the cattle herd component of each farm (Figure 3). Furthermore, in the depiction of the model it is indicated where the original scenarios directly interfere with the flows. The model was created in the programming language R. For running the model and graphical visualisation of the majority of results Rstudio1.1.456 was used. To calculate SI indicators, farm characteristics and external as well as internal drivers were used as inputs. The farm characteristics were derived from the surveys. External drivers are non-farm-specific variables and their values were based on literature research and rational and coherent assumptions established in the various scenarios. Internal drivers are farm specific variables. In the following section the different main components of the model are described, followed by a description of the model outputs.

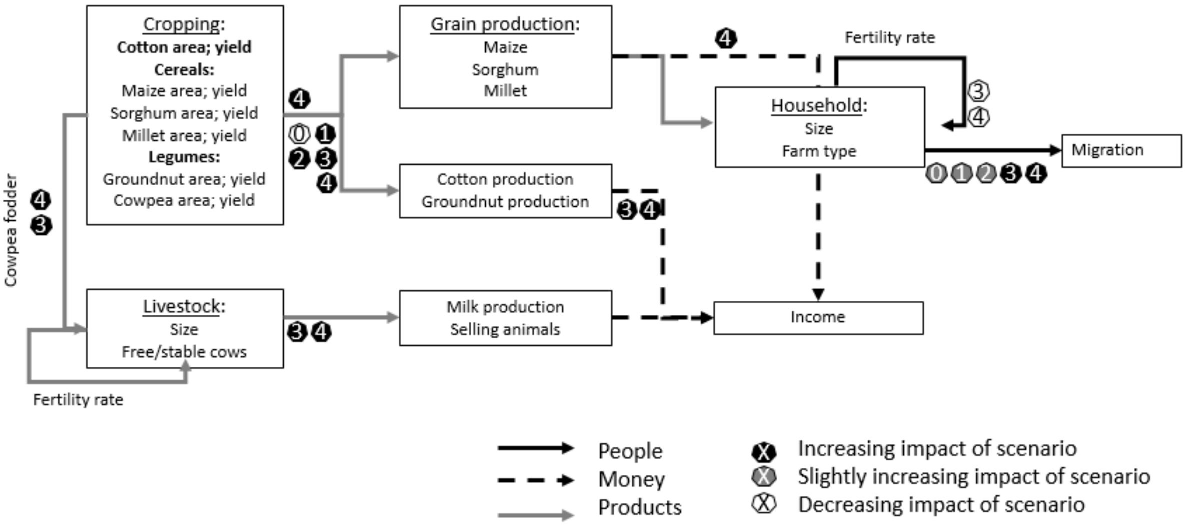


Figure 2: Simple model of a typical farm household from the Koutiala region and the different flows of people, money and products between the main household components independent from farm type (Falconnier et al., 2018). A hexagon indicates that a scenario described by the number in the hexagon, has a direct impact on a flow. The colour of the hexagon describes the nature of the effect on the scenario.

2.4.1 Household Component

The household component is affected by two major external drivers, fertility and rural to urban migration rates of household members. The baseline number of household members (*HHsize2013*) was taken from the CMDT. The household size in the near future (*HHsize2027*) for each household was estimated as follows:

$$HHsize_{2027} = HHsize_{2013} * \left(1 + (fertilityrate/100 - migrationrate/100)\right)^{2027-2013}$$

The current net fertility rate (birth minus death), calculated based on historical household data, is 3.4 %, similar for all farm types (Falconnier et al., 2018). The rural to urban migration rate, however, differed between the four farm types and was estimated as 0 %, 1.7 %, 1.2 % and 2.8 % for the HRE-LH, HRE, MRE and LRE farms in Nampossela (Falconnier et al., 2018). For the farms in Sirakélé the same migration rates per farm type were assumed.

The number of working household members was taken from the CMDT surveys. The number of workers in relation to the household size for each farm was kept constant for the near-term future.

2.4.2 Cropland Component

The cultivated area for each relevant crop, namely cotton, maize, sorghum, millet, groundnut and cowpea, of the baseline scenario was taken from the CMDT survey. The total cropping area, however, was estimated by summing the individual areas of all crops. This also includes the cropping area for crops that were not considered in this research because of their small spatial share or low abundance across the farms e. g. soybean and rice. While the individual cropping areas of the different crops were influenced by the scenarios, the total cropping area was kept constant, based on the assumption that no land separation will take place due to the common practice of leaving the household to the oldest son (see section 2.1). Farm size increases through e. g. acquisition were not considered.

The individual crop yields were taken from Falconnier et al. 2018 (Table 9, Appendix). For the individual yields it was checked what the limiting or reducing factors are. Based on this it was decided to show either annual variability or differences between farm types. Cotton yields are mainly affected by weeds and pests and thus yield variations through rainfall variability are not visible in the baseline situation, resulting in constant yields over all seasons (Table 9, Appendix). Due to this, farm management practices such as timely sowing, weeding and land preparation influence cotton most (Traore et al., 2013). As these practices are affected by the household's resource endowment, the cotton yields are dependent on the farm type (Falconnier et al., 2015). Maize yields, which were less affected by pests and weeds than cotton, are affected by rainfall variability and thus seasonal yields were simulated by using the Agricultural Production System Simulator (APSIM) (Keating et al., 2003). For this, weather data from the N'Tarla weather station over a period from 1965 - 1993 were used. Three years, however, are missing in the data set (1968, 1972, 1991), resulting in a total observation of 26 seasons. Differences in yield between the farm types occurred because LRE farmers are not able to invest in as much inputs (40 kg N ha^{-1}) for maize as the other farm types (60 kg N ha^{-1}). Sorghum, millet, cowpea and groundnut yields were found not to be affected by seasonal rainfall and did not differ for the farm types and were thus kept constant over all seasons and farms (Falconnier et al., 2018) (Table 9, Appendix).

A closed yield gap was defined as 80 % of the water-limited yield for cotton and the cereals, since the water-limited yield is not attainable in farm trials (Ittersum et al., 2013). For cotton the water-limited yield was reached by applying 90 kg ha^{-1} mineral fertilizer and 12.8 t dry matter manure ha^{-1} (Ripoche et al., 2015). For the cereals required N applications to reach 85 % of the water limited yield (maize 110 kg N ha^{-1}, millet and sorghum 150 kg N ha^{-1}) were identified through tests in APSIM (Table 9, Appendix).

2.4.3 Cattle Herd Component

The current herd size of each farm was taken from the CMDT survey. The net fertility rate of 10 % and the annual animal offtake were assumed to cancel each other out (Ba et al., 2011). Thus, the herd size of the farms in this research was assumed to stay stable. In the current situation the cows graze rangelands outside of the farm system from June to December and are fed with crop residues of maize, millet, groundnut, cowpea and sorghum from January until May. Leftover residues are assumed to be brought back on the fields. Feeding and guarding the cattle while grazing is done by one worker of the farm per day. During the night cattle are kept within a fenced area on the farm (Personal communication, Ousmane Sanogo, 2018). The amount of manure that is collected during free grazing was assumed to be 43 % of the total production, namely the amount cattle is excreting during the night (Rufino et al., 2006). This was assumed, as it is likely too time-consuming to collect the manure that is excreted during free grazing outside of the farm system. Thus, 57 % of the total manure is considered as a loss. The N-losses during storage are assumed to be 40 % in the baseline which corresponds to the high N losses from dairy cow manure in pit storage reported by Dong et al. (2006).

In the Koutiala region farmers with herds of more than 20 cattle practise transhumance. In this research the findings of Ba & Lesnoff (2011) were simplified. It was assumed that, except for 50 % of the milking cows, transhumance practising farmers give their herd to a herder for six months (March – September). The herders take the cattle to the north of Ivory Coast where more pasture is available due to a more suitable climate and lower population density (Ba & Lesnoff, 2011). Half of the lactating cows are left behind to cover the personal needs of the household. During transhumance brought away cattle were not treated as part of the farm system and thus their manure was neither considered as loss or input nor was income generated through the selling of their products.

The milk yield for the current practice and stall feeding as alternative livestock holding system was taken from de Ridder et al. (2015) and set to 65 kg per cow per year for free grazing and 226 kg per cow per year for stall-fed cows. The difference is based on the energy waste of the cows walking to grazing areas and through differences in fodder quality. 2.5 kg cowpea hay and 2 kg cotton seed cake per cow per day was assumed to be required as high quality feed to realise stall feeding (de Ridder et al., 2015). Based on the availability of high-quality fodder the amount of stall-fed cows were estimated. In herds with more than 23 cattle 34 % of the cows were assumed to be lactating at any point in time. In herds below this threshold only 22 % of the cows are lactating at a time (Ba & Lesnoff, 2011).

2.5 Sustainable Intensification Indicators

The indicator identification was based on a framework created by Florin et al. (2014) and further used in Marinus et al. (2018) (Figure 3). The framework describes the hierarchical structure of principles (referred as domains in this research), criteria and indicators (Figure 3). A single domain has different criteria and a criterion has several indicators.

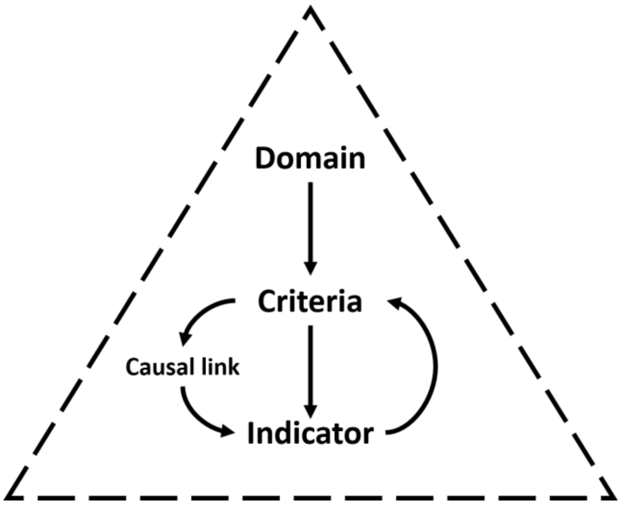


Figure 3: Hierarchical framework of principles, criteria and indicators selected (Marinus et al., 2018).

In order to expand the scenario analysis to all SI domains the indicator selection was made in consideration of the listed selection criteria from Olde et al. (2017) (Table 13, Appendix). The selection further depended on the relevance of the indicators for the context and the data availability for the calculation (Table 1).

To enhance their acceptance by other scientists and the comparability with other studies in literature, mainly indicators also listed by Smith et al. (2017) were chosen (Table 1). Indicators were calculated for both time points, baseline and near-term future, for each farm for a period of one year, except for labour intensity, which was calculated for ten fortnights of the agricultural season, and farm income per capita, which was converted to a daily basis. The impact of rainfall variability on the different indicators was also explored by estimating the coefficient of variation for each scenario with the following formula:

$$CV = \frac{\sigma}{\bar{x}}$$

With:

- CV Coefficient of variation
- σ Standard deviation
- $\bar{x}$ Average

Indicators that are not affected by the weather were kept constant over all 26 years.

Table 1: List of all indicators and pathways of their identification based on the hierarchical framework of principles, criteria and indicators by Florins et al., 2014 (Marinus et al., 2018).

Domain	**Criteria**	**Causal links to: the scenarios**	**Indicator**	**Unit**
Productivity	Increased farm productivity	… affect the crop and milk as well as manure production. Thereby the amount of self-produced biological inputs	Self-produced fodder	kg year^{-1}
			Self-produced manure	kg year^{-1}
		… increase the cropping and livestock intensification	Legumes production	kg year^{-1}
			Cereals production	kg year^{-1}
			Cotton production	kg year^{-1}
			Animal production	kg milk year^{-1}
Economic	Reduced poverty	… enhance on farm production and enable off-farm revenue streams which result in increased remittances	Income per capita	\$PPP capita^{-1} day^{-1}
		… change the income and the labour demand	Labour productivity	\$PPP man-day^{-1}
Environmental	Raised soil fertility	… change the nitrogen application and farming practices	Nitrogen balance	kg ha^{-1} year^{-1}
	Improved resource use efficiency	… implement new farming techniques	Nitrogen use efficiency	ratio
Human well-being	Enhanced human health	… will affect the income and food production	Nutritional self-sufficiency	ratio
			Labour intensity	ratio

2.5.1 Indicator Scaling

Scaling allows a comparison of indicators with different units and enables the identification of trade-offs of the indicators with radar charts. The indicators were scaled on a range from 0 to 10, where 10 displays an optimal performance. The feasible maximum and minimum for each indicator was identified for all farms over all scenarios. If not indicated differently, the indicator's best performing farm having on average over all seasons the highest performance of an indicator was taken to obtain X_{max}. It was assumed that this is the best attainable performance of farms in the Koutiala region. For X_{min} the worst performing farm was considered.

The scaling was based on the difference between the maximum and minimum value as follows:

$$X' = \left(\frac{X - X_{min}}{X_{max} - X_{min}}\right) * 10$$

In the following the chosen indicators for each domain are described and it is explained why a specific indicator was chosen. Afterwards the calculation is given and the individual scaling described. All used variables can be found in Table 8 (Appendix). All formulas for the indicator calculations are numbered for improved orientation in Table 8.

2.5.2 Productivity

2.5.2.1 Self-produced Biological Inputs

Increased production of biological inputs as manure for the cropping system as well as fodder for the livestock production has a positive effect on the intensification of farms in the Koutiala region (Falconnier et al., 2017). Furthermore, the on-farm production of inputs is important to describe the effect of S2 - S4, which directly affect the production of high-quality fodder and the intensification of the livestock systems (Table 6).

Calculation

The amount of manure that is produced as biological input for the crops was represented by the amount of available nitrogen that can be applied on the cropping area (*MANNav* in kg N year^{-1}). Only cattle manure was taken into account. The percentage of N in cattle manure (*Nmanure* in %) was set to 1.5 % (Rufino et al., 2007). The production rate (*pm* in kg DM cow^{-1} day^{-1}) was kept constant for all seasons and farm types at 1.22 kg DM cow^{-1} day^{-1} (de Ridder et al., 2015). The percentage of the total produced manure that is collected (*cm)* was set to 43 % for free grazing cows and assumed to be 100 % for stall-fed cows. During transhumance only on-farm cattle were considered. The following formula describes the amount of N that is available to be spread on the cropland.

$$MANNav = pm * Nmanure/100 * 365 * \ n_{cattle} * (cm/100) * \ (1 - sl/100) \tag{1}$$

With:

- n_{cattle} Number of cattle per farm
- sl N lost during pit storage [%] (Dong et al., 2006)

Originating from the cropland component harvested residues function as input for the livestock component. Therefore, the total amount of produced edible crop residues were calculated (FOD_{tot} in kg). This excludes the residues of cotton which are burned after the harvest. Differences in feed quality between the crops are not taken into account for this indicator. Thus, just the total residual biomass was calculated.

The fodder yield (FOD_i in kg ha^{-1}) per crop was estimated with the following formula.

$$FOD_i = ((1 - HI_i/100)/(HI_i/100)) * Yield_i * 1000 \qquad (2)$$

With:

- HI_i	Harvest index of crop i [%]
- $Yield_i$	Grain yield of crop i [tons ha^{-1}]

The total fodder production per farm ($FODtot$ in kg ha^{-1} year$^{-1)}$ was estimated as follows.

$$FOD_{tot} = \sum_i (A_i * FOD_i) \qquad (3)$$

With:

- A_i	Area under crop i [ha]

Scaling

As X_{min} and X_{max} indicator values from the worst and best performing farm were chosen. For this indicator best performance was defined as the largest production of organic inputs.

2.5.2.2 Total Production

In the literature, yield is the most commonly used indicator for the productivity domain in SI assessments (Smith et al., 2017). It describes the state of intensification in kg per hectare for crop production and in kg milk per TLU for livestock production. In this research, however, the total production per farm ($GRAIN_{tot\,i}$ in kg year^{-1}) was estimated since it was desired to evaluate also the impact of changes in the cropping area of the various crops. Almost every scenario tested was expected to have a direct or indirect impact on grain production.

Calculation

The calculation of the total crop production for each farm type was repeated for all 26 potential seasons to display the effects of rainfall variability. The total amount of produced grain was estimated for each crop as follows:

$$GRAIN_{tot\,i} = \sum_i (A_i * \ Yield_i * 1000) \qquad (4)$$

The milk production ($MILK_{tot}$ in kg year^{-1}), however, was assumed not to be affected by differences in rainfall as the yield of cowpea was not influenced by the different seasons and the effect of rainfall variability on pasture went beyond the scope of this research. Thus, the total milk production was kept constant over all potential weather data. During transhumance only on farm cattle (n_{cattle}) was considered in the calculation of the milk production. The total amount of produced milk was estimated with the following formula.

$$MILK_{tot} = n_{cattle} * lc/100 * Myield \qquad (5)$$

With:

- lc Proportion of herd lactating [%]
- $Myield$ Milk yield [kg cow^{-1} year^{-1}] (De Ridder et al., 2015)

Scaling

As X_{min} and X_{max} indicator values from the worst and best performing farm were chosen. Like for the previous indicator best performance was defined as highest production.

2.5.3 Economic Sustainability

2.5.3.1 Farm Income per Capita

The generated income is the most commonly used indicator of the economic domain for SI assessments (Smith et al., 2017). It gives an indication of the value of the agricultural production but also describes its profitability. For this research the net farm income per capita was calculated and compared between the farm types. Farm income per capita was given in US dollars purchasing power parity ($PPP) to allow a comparison with the international poverty line of 1.9 $PPP day^{-1} person^{-1} (Jolliffe & Prydz, 2016).

Calculation

The considered revenue streams were the sale of crop and livestock products and remittances from former household members. For this the own consumption of the household was not deducted and thus the total produced value was estimated. As expenses the production costs of the grain (seed and input costs, costs tractor services and for LRE farmers additionally renting ox labour), costs for livestock husbandry (fodder costs and vaccination costs) and depreciation costs of the machines and animals were considered (Falconnier et al., 2018).

The used fertilizer urea and NPK contained 46 % N and 15 % N and cost 65 $PPP and 74 $PPP per sack (50 kg) respectively (project intern survey). For each application rate two sacks of NPK fertilizer per hectare were used to cover the need for P and K and the rest of the required mineral N originates from urea as it is cheaper.

Grain prices for the years of 2018/2019 were obtained during market surveys done by other project members (Table 14, Appendix).

All scenarios affected the income per capita via increases in productivity or direct effects on the prices for produced goods or for inputs.

The final net income per capita (*Inc per capita* in \$PPP day^{-1}) was estimated with the following formula (Formula 6).

$$Inc\ per\ capita = \frac{(Crop\ inc + Lv\ inc + r + se - Depr\ costs)}{HHmembers * 365} \qquad (6)$$

With:

r	Remittances from migrated household members [\$PPP]
se	Income generated of the household members through self-employment (sale of on-farm produced non-agricultural goods) [\$PPP]

The estimations for variables in the income per capita calculation which require a more detailed description are listed in the following.

The production cost for each crop (i) per hectare was calculated as follows.

$$cost_i = (SD_i * seed_{costs\ i} + Finput_i) \qquad (7)$$

With:

$cost_i$	Costs for crop i [\$PPP ha^{-1}] (Market survey 2018)
SD_i	Sowing density of crop i [kg ha^{-1}]
$seed_{cost\ i}$	Costs for seeds of crop i [\$PPP kg^{-1}]
$Finput_i$	Costs for fertilizer for crop i [\$PPP ha^{-1}]

For LRE households who do not own draught animals, the costs related to the crop additionally include labour costs for three pairs of oxen of 23 \$PPP per ha^{-1}.

The total net income generated by the cropland component (*Crop inc* in \$PPP year^{-1}) was estimated as follows.

$$Crop\ inc = \sum_i((Yield_i * A_i * price_i) - (A_i * cost_i)) \ (8)$$

With:

$price_i$	Crop specific price farmers get for their products [\$PPP kg^{-1}] (Falconnier et al., 2018)

Income generated by farming activities in the cattle herd component was estimated with the following formula. The price per kilogram milk was dependent on the scenarios.

$$Lv\ inc = \Big((MILKtot\ *\ price_M)\ + (n_{cattle}\ * offtakerate/100\ *\ price_{anim}) - \Big((costs_{fd} + costs_v)\ *\ n_{cattle}\Big)\Big) \qquad (9)$$

With:

$Lv\ inc$	Income of livestock husbandry after subtracting expenses [\$PPP]
$price_M$	Milk price [\$PPP kg^{-1}] (Falconnier et al., 2018)
$costs_{fd}$	Costs for additionally bought fodder [\$PPP livestock $head^{-1}$ $year^{-1}$] (Falconnier et al., 2018)
$costs_v$	Costs for vaccination [\$PPP livestock $head^{-1}$ $year^{-1}$] (Falconnier et al., 2018)
$offtakerate$	Rate by which farmers sell animals to keep the herd size constant [%]
$price_{anim}$	Price for selling an animal [\$PPP $animal^{-1}$]

The depreciation costs for the different machines (plough, weeder, sowing machines and carts) and oxen were calculated.

$$Depr\ costs = \sum_m (n_{machines\ and\ ox} * bprice_{machines\ an\ ox} / ls_{machines\ and\ ox}\ ) \qquad (10)$$

With:

$Depr\ costs$	Total costs due to depreciation of machines and animals [\$PPP $year^{-1}$]
$bprice$	Buying price [\$PPP]
$n_{machines\ and\ ox}$	Amount of owned specific drawn equipment (plough, weeder, sowing machine, cart) and oxen
$ls_{achines\ and\ ox}$	Live span of the specific drawn equipment [year]

Scaling

As X_{min} and X_{max} indicator values from the worst and best performing farm were chosen. The highest per capita farm income was defined as best performance.

2.5.3.2 Labour Productivity

While labour intensity describes the workload of the household members in the farm labour, labour productivity relates it to the generated income. It describes the efficiency in which labour is converted into outputs (Tittonell et al., 2007). The considered output was the total income generated and the work input was assumed to be the demanded work. Both are affected by the scenarios.

Calculation

The labour productivity (*Labour prod* in \$PPP man-day^{-1}) was averaged over the whole year. An observation of the different fortnights respectively was not possible since many income streams are generated at the same time point, i.e. at the harvest. As a threshold for labour productivity the labour productivity of the minimum wage in Mali (185 \$PPP (https://wageindicator.org, last accessed 07/08/2019) to the predominant 40-hour workweek (160 hours per month) was taken.

$$Labour\ prod = \frac{tot.inc}{\sum_i (demand_i * A_i + demand_{animal})} \qquad (11)$$

With:

tot. inc	Total income generated after subtraction of expenses [\$PPP]
$demand_i$	Labour demand for crop i per year [man-days ha^{-1}]
$demand_{animal}$	Labour demand for the livestock husbandry per year [man-days]

Scaling

As X_{min} and X_{max} indicator values from the worst and best performing farm chosen while the highest indicator value describes the best performance.

2.5.4 Social

The domain of social sustainability was not covered in this research, even though, it was part of the initial research aim. To identify indicators of the social domain which have a sufficient level of significance for the observed system (de Olde et al., 2017), a deep understanding of the system is required. That is either built up by involvement of many stakeholders into the analysis or given after self-performed research. Both exceeds the scope of this research.

2.5.5 Environment

2.5.5.1 Partial Nitrogen Balance

Sub-Saharan African’s soils suffer degradation because of nutrient depletion. It is one of the main issues that need to be addressed to intensify sub-Saharan Africa’s agriculture (Droppelmann et al., 2017). This indicator states if N inputs balance with N outputs. Changes in the composition of the farming system or input and output flows by the different scenarios can affect the partial N balance and as N is often a limiting factor it is important to monitor these impacts.

Calculation

For the calculation of the partial N balance only flows that enter or leave the farm system were considered (Figure 1). Internal flows between the farm system components were not considered as the assessment was made at farm level. The household component was also not included to simplify the system. From this point on the partial N balance is simply called N balance.

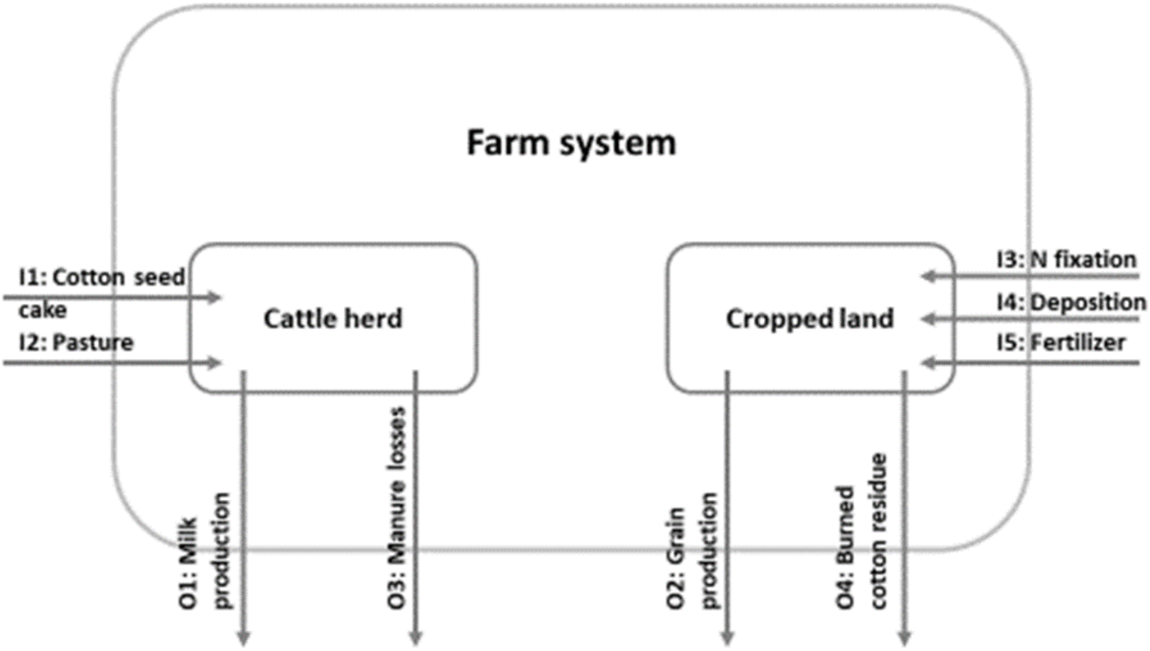


Figure 4: Depiction of the N balance and the considered N flows that enter (I1 - I5) or leave (O1 - O5) the cattle herd and cropped land components of the farming system and thereby the system itself. The household component is not included in the N balance.

Fertilizer application rates (I5) are defined based on the CMDT for the baseline or defined by the scenarios (Table 6). For the N fixation of legumes (I3) it was assumed that for cowpea and groundnut respectively 70 % and 64 % of the N in their above ground biomass originates from N fixation from the air (Sanginga et al., 2000; Phoomthaisong and Toomsan, 2003). The amount of nitrogen that is brought into the system by the grazing of the cattle (I2) was estimated by assuming that during the period of June until December all N in the manure originates from grazing on pasture. Nitrogen from pasture that ends up in milk or mass gain of the cattle was not considered. For the other months it was assumed that the cattle are fed with crop residues, which were not included in the N balance as they are self-produced, and bought cotton seed cakes, which were considered as an input. The amount of manure that is not collected was assumed to be 57 % of the total production during free grazing and 100 % during stall feeding. N losses during storage were assumed to be 40 % in the baseline. The observed systems are exposed to N deposition, including dry and wet deposition, (*depo* in kg ha^{-1} $year^{-1}$) of around 8 kg N ha^{-1} $year^{-1}$. The N contents of the different products that were multiplied with the produced quantities can be found in Table 2. The consumption of the household and the sales of products were both considered as outflows. As a threshold an N balance at the value 0 was taken as soil mining is a predominant risk in sub-Saharan Africa (see section 1.1). This threshold was defined as it is simple to understand. The definition of a more meaningful threshold that covered also the minimum amount of nitrogen that farmers want to get out of their system through the applied farm management systems lay beyond the scope of this research.

Table 2: Nitrogen content in percent of the various products considered in the N balance and proteins in general.

Product	N content (% of DM)	Reference
Maize grain	1.6	
Millet grain	2.1	
Sorghum grain	2.1	
Cowpea grain	3.8	Nijhof, 1987
Groundnut grain	4.5	
Cotton grain	1.9	
Cotton straw	1.5	
Cotton seed cake	5.12	Adegun and Aye, 2013
Milk	0.512	Agyemang et al., 1991
Manure	1.5	Rufino et al., 2006
Proteins	16	Maclean et al. (2003)

The N balance in kg N was then calculated with the following formula:

$$N\ balance = \left(\left(\textstyle\sum_i (fert_i * A_i)\right) + depo * A_t + Nfix + Tot.Ncsc + Npast\right) - (Ngrain + Nmilk + MANNloss + Tot.Nstraw_C\ loss) \quad (12)$$

With:

- $fert_i$ Crop specific amount of fertilizer applied [kg N ha^{-1}]
- A_t Total cropping area [ha]

In the following, the estimation of the different nitrogen inflows (I1 - I3, Figure 4) and nitrogen outflows (O1 - O5, Figure 4) which require a more detailed explanation are described (Formula 13 - 19). The formulas of the last two considered inflows I4 deposition and I5 fertilizer application will not be described in detail as they are included in the calculation of the N balance (Formula 14).

Materials and methods

The amount of N entering the system by buying cotton seed cakes (I1, *Tot.Ncsc* in kg year^{-1}) as fodder for stall-fed cows was estimated with the following formula (I1, Figure 4).

$$Tot.Ncsc = n_{sfcattle} * tsf * Fcsc * Ncsc/100 \qquad (13)$$

With:

$n_{sfcattle}$	Number of cattle that are stall-fed during the year
tsf	Number of days cattle is stall-fed [days year^{-1}]
Fcsc	Number of required cotton seed cake for a stall-fed cow [kg cow^{-1} day^{-1}]
Ncsc	N content of cotton seed cakes [%]

N entering the system through grazing of the cows on pasture (I2) outside of the farming system (*Npast* in kg year^{-1}) was estimated as follows (I2, Figure 4)

$$Npast = MANNav * tgr/100 \qquad (14)$$

With:

tgr	Proportion of the year the cows feed on pasture [%]

N entering the system through N fixation of legumes (*Nfix* in kg year^{-1}) (I3) was calculated as follows (I3, Figure 4).

$$Nfix = \sum_{leg}((Nfix_{leg}/100) * Cbiomass_{leg}) \qquad (15)$$

With:

$Nfix_{leg}$	N fixed of total N of each specific legumes [%] (Sanginga et al., 2010 and Phoomthaisong et al., 2003)
$Cbiomass_{leg}$	Above ground biomass produced for each legume [kg year^{-1}]

The N leaving the system through the consumption and sale of milk (*Nmilk* in kg year^{-1}) (O1) was calculated as follows (O1, Figure 4).

$$Nmilk = MILK_{tot} * milk\ prot/100 * Nprot/100 \qquad (16)$$

With:

milk prot	Protein content in milk [%]
Nprot	Average N content of proteins [%] (Maclean et al., 2003)

The second productive outflow is N that leaves the system through the consumption and sale of grains (*Ngrain* in kg year^{-1}) and it was estimated with the following formula (O2, Figure 4).

$$Ngrain = \sum_i GRAIN_i * (Ngrain_i/100) \qquad (17)$$

With:

$Ngrain_i$	N content in crop specific grain [%] (Table 2)

The amount of N that is lost due to losses of produced manure (*MANNloss* in kg) was estimated based on the amount of manure that was not collected and the amount of manure that was collected (*cm*) but lost during storage (*sl*) (O3, Figure 4).

$$MANNloss = (n_{cattle} * pm * 365 * Nmanure/100) * (\left(1 - \frac{cm}{100}\right) + (cm/100 * sl/100))$$

(18)

The amount of N that is lost due to combustion of cotton straw was calculated as follows (O4, Figure 4).

$$Tot.Nstraw_C\ loss = A_C * \left(\frac{1-\frac{HI_C}{100}}{\frac{HI_C}{100}}\right) * \ Yield_C * \ Nstraw_C/100 \qquad (19)$$

With:

$Tot.\ Nstraw_C\ loss$	N losses in burned cotton residues [kg year^{-1}]
A_C	Area cultivated with cotton [ha]
HI_C	Harvest index of cotton [%]
$Yield_C$	Cotton yield [kg ha^{-1}]
$Nstraw_C$	N content of cotton straw [%]

Scaling

As X_{min} and X_{max} indicator values from the worst and best performing farm were chosen. The worst performance is represented by the farm with the most negative N balance values and the highest performance by the farm with the most positive N balance.

2.5.5.2 Nitrogen Use Efficiency

In particular, in areas with degrading soils and low nutrient availability like in sub-Saharan Africa nutrient management is important (Droppelmann et al., 2017). The success of the efficient use of applied nitrogen in relation to productive N outputs can be described through nitrogen use efficiency (NUE). Every scenario has an impact on the application of inputs and crop yields of the different farm types, and hence on NUE.

Calculation

Materials and methods

For NUE, unlike for the N balance only profitable outputs were considered, namely nitrogen in grain and in produced milk. As inputs only the productive inputs were considered, namely fertilizers, cotton seed cakes, N fixation and N that originated from the pasture (Figure 5).

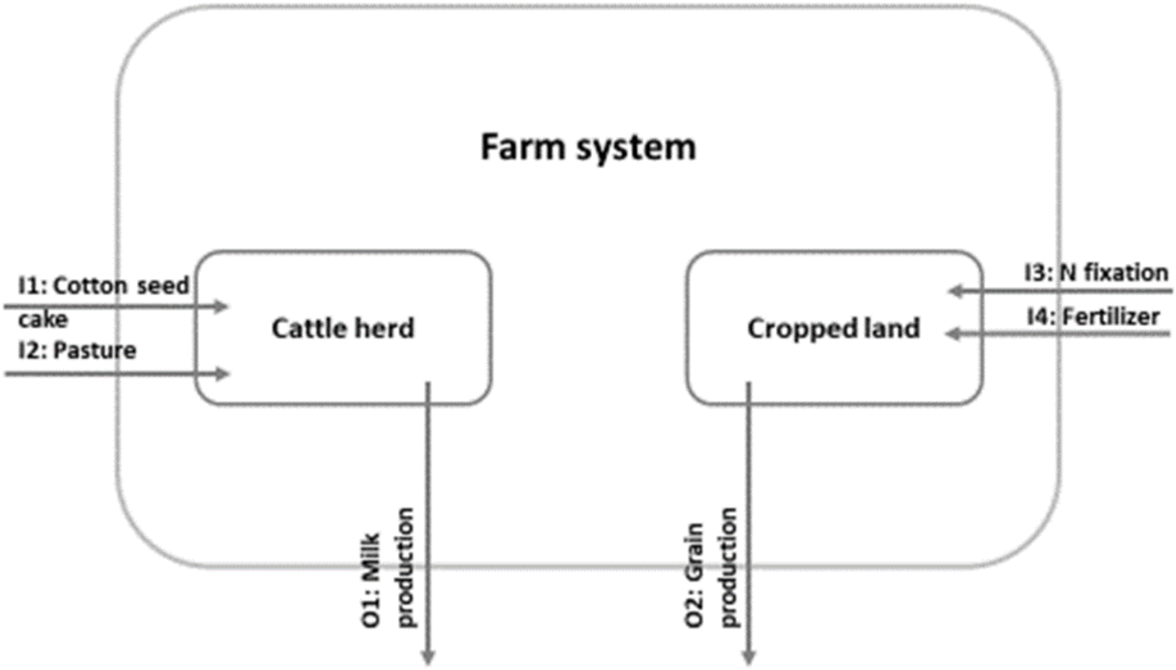


Figure 5: Depiction of the NUE and the productive N flows that enter (I1, I3 and I4) or leave (O1 and O2) the cattle herd and cropped land components of the farming system and thereby the system itself. The household component is not included in the NUE.

$$NUE = \frac{(Ngrain + Nmilk)}{((\sum_i (fert_i * A_i)) + NFix + Tot.Ncsc + Npast)} \qquad (20)$$

Scaling

NUE was scaled with the “traffic light indicator” scheme developed by Brentrup and Palliere (2010) (Marinus et al., 2018). However, it must be noted that this scaling does not allow differentiation between soil mining (NUE > 1) and a high risk for N losses (NUE < 0.5) which are both not desirable.

Table 3: Traffic light indicator scheme for the NUE indicator, assigning logical interpretations to outcomes range of the NUE to enable indicator scaling (European Expert Panel., 2015).

Interpretation	Nitrogen use efficiency
N mining	>1
Risk of N mining	0.9-1
Balanced N inputs	0.7-0.9
Risk of N losses	0.5-0.7

High risk of N losses	<0.5

Based on the classification in different ranges (Table 3) of the NUE a simple scaling was conducted following the pattern of Figure 6.

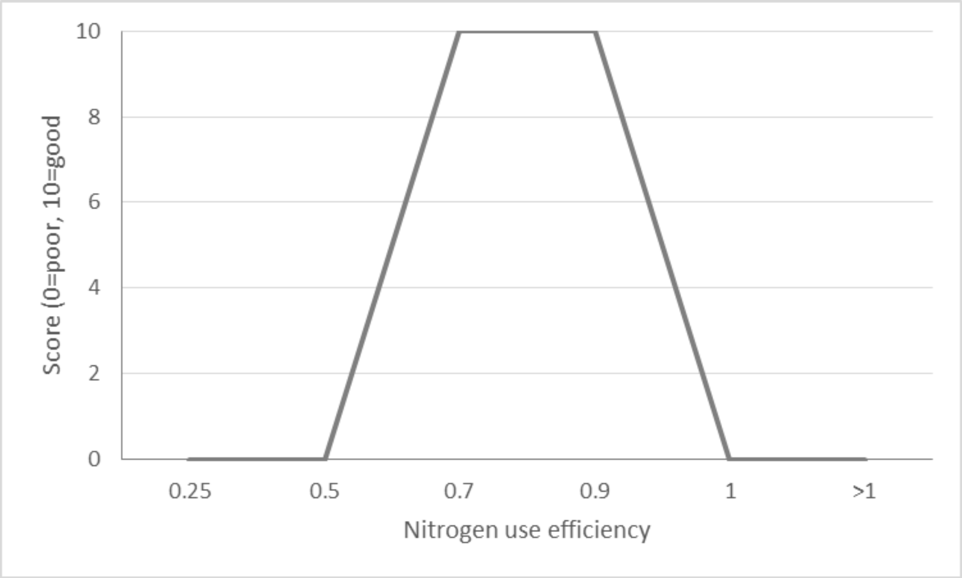


Figure 6: Scaling pattern of the NUE (Marinus et al., 2018).

2.5.6 Human well-being

2.5.6.1 Nutritional Food Self-sufficiency

In sub-Saharan Africa it can be a struggle to enhance the income without reducing the food self-sufficiency of smallholder farmers (Falconnier et al., 2017). Furthermore, being food self-sufficient as a farmer is very important in the Malian culture and thus a primary goal of farmers (Ollenburger et al., 2016). Thus, testing if changes in the different scenarios threaten or improve the food self-sufficiency is very important to point out impacts on the human well-being. Since Falconnier et al. (2018) already evaluated the food self-sufficiency in terms of energy supply, in this research the analysis was extended for specific nutrients, namely protein, iron and zinc and related to the daily requirements of an adult man (Table 4).

Table 4: Calories, protein, iron and zinc abundance per kilogram of on-farm produced products and the daily human need of the individual nutrient. The letters “a” to “e” after the values refer to the references listed below the table.

	Calories (kcal/kg)	**Protein (g/kg)**	**Iron (g/kg)**	**(g/kg)**
Maize	3500 [a]	87.5 [a]	0.01740 [a]	0.02240 [a]
Sorghum	3500 [a]	10.2 [a]	0.03360 [a]	0.01670 [a]
Millet	3500 [a]	11,0 [a]	0.08000 [a]	0.01680 [a]
Cowpea	3360 [a]	235 [a]	0.01381 [a]	0.00563 [a]
Groundnut	5670 [a]	258 [a]	0.04580 [a]	0.03270 [a]
Milk	1080 [a]	33.3 [a]	3,00000 [a]	0.00400 [b]
Beef	1530 [a]	311 [a]	0.02420 [a]	0.04930 [a]
Daily human need	2406 [c]	52,0 [d]	0.02400 [d]	0.09500 [e]

a: United States Department of Agriculture Agricultural Research Service USDA Food Composition Databases (https://ndb.nal.usda.gov, last accessed 02/06/2019)

b: American Dairy Science Association (https://www.journalofdairyscience.org, last accessed 02/06/2019)

c: Falconnier et al. (2018)

d: FAO (http://www.fao.org/docrep/U5900t/u5900t03.htm, last accessed 16/02/2019)

e: U.S. Department of Health & Human Services (https://ods.od.nih.gov/, last accessed 16/02/2019)

Calculation

The food self-sufficiency for energy intake and different nutrients were averaged over the year and expressed per day. For this the produced amounts were related to the daily demand of the specific nutrient of the household members.

First the demand for each nutrient per day was estimated with the following formula.

$$tot.Nutr\ demand_{nd} = Nutr\ demand_{nd}/1000 * HHmember \qquad (21)$$

With:

- $Tot.\ Nutr\ demand_{nd}$ Daily household demand of each observed nutrient [kg day^{-1}]
- $Nutr\ demand_{nd}$ Daily individual demand of each observed nutrient [g person^{-1} day^{-1}]

In the second step the total produced amount of each nutrient per day was estimated. For this the total amount of produced grain and animal products were considered. The meat production was based on the off-take rate, assuming that these cattle are potentially eaten. For this an average body weight of a N’dama cattle of 300 kg and a dressing percentage of 40 % was assumed (Agyemang, 1992).

$$tot.Nutr\ prod_{nd} = \frac{((Yield_i * A_i * grain\ nutr_i + MILKtot * nutr_{milk})\ /\ 1000 + n_{cattle} * offtakerate\ /\ 100 * c.weight * dressing\ /\ 100 * nutr_{meat})}{365}$$

(22)

With:

• *Tot. Nutr prod*$_{nd}$	Total amount produced of each observed nutrient [kg day^{-1}]
• *grain nutr*$_i$	Nutrient content of the specific crop grain [g kg^{-1}]
• *nutr*$_{milk}$	Nutrient content of milk [g kg^{-1}]
• *nutr*$_{meat}$	Nutrient content of meat [g kg^{-1}]
• *c.weight*	Average body weight of N'dama cow [kg]
• *dressing*	Proportion of cattle body weight that is turned into beef [%]

The final calculation of the nutrient self-sufficiency for each nutrient related the demand to the production (23). Values above one suggest that the farms are self-sufficient and below one that they are not. Thus, the value of one is defined as the threshold which the farms need to surpass to be sustainable.

$$nutr\ selfsuff_n = \frac{tot.Nutr\ prod_{nd}}{tot.Nutr\ demand_{nd}} \qquad (23)$$

Scaling

As X_{min} and X_{max} indicator values from the worst and best performing farm were chosen. The worst performance is defined as the lowest values and the best performance is described by the highest value.

2.5.6.2 Labour Intensity

Labour intensity describes the intensity of the farming system by relating the labour demand with the actual available labour. Changes in mechanisation or changes in the crop land distribution affect labour intensity.

Calculation

The calculation of the labour intensity per day was based on a labour calendar which describes the labour demand in man-days per hectare for the main operations per crop (Table 16, Appendix). In the calendar agricultural tasks are assigned to the fortnights in which they are usually completed (Falconnier, 2009). For the livestock component one man-day per day was assumed for the baseline since one person takes care of the herd during grazing, irrespective of the herd size (Personal communication, Ousmane Sanogo, 2018). The indicator was calculated by dividing the total labour demand on each farm with the number of workers (each representing one man-day). This was done for each observed fortnight. Only fortnights in which the cropping system demands labour were considered since the workload of the livestock system alone is negligible. The threshold for labour intensity is one; a labour intensity above one indicates a labour shortage.

The labour intensity per day of individual fortnights was estimated with the following formula.

$$Labour\ int_f = \frac{\left(\frac{\left(\sum_{if}\left(demand_{if} * A_i\right) + animal_f\right)}{days_f}\right)}{n_{worker}} \qquad (24)$$

With:

• $Labour\ int_f$	Labour intensity per day of individual fortnights
• $demand_{if}$	Labour demand for crop i and fortnight [man-days ha^{-1}]
• $animal_f$	Labour demand per fortnight for livestock holding [man-days]
• $days_f$	Number of days per fortnight
• n_{worker}	Number of workers per farm

Scaling

For the scaling the labour intensity will be averaged over all observed fortnights per year. As X_{min} and X_{max} indicator values from the worst and best performing farm were chosen. The performance in labour intensity is high if the indicator values are low.

2.6 Data Analysis

Different sets of indicators were chosen to answer the different research questions (Table 18, Appendix).

For the first research question, targeting the farms' performance in the different indicators and eventually the different domains at both time points, only indicators with a defined threshold signalling (un)satisfactory performance were analysed (Table 18, Appendix). These indicators were averaged over all seasons and presented for all farms. Additionally, for the over the farm types averaged farm income per capita was the seasonal variation shown to capture whether seasonal yield variation determines that the farms are below or above the poverty line. Furthermore, the scaled performance of the farms in the different domains was estimated by averaging the scaled performance of all indicators for each domain respectively.

For the second research question, identifying the key external and internal drivers, only indicators with a defined threshold were used as they already include the outcomes of the indicators of the productivity domain (Table 18, Appendix). As a comprehensive sensitivity analysis on each external and internal driver lay beyond the time limit of the research, more summarising potential determinants were identified and checked for. These are intermediate results of the indicator calculation process and themselves influenced by farm characteristics and several drivers. For the farm income per capita these are the different revenue streams, for labour productivity the total farm income, for N balance and NUE the considered inputs and for the nutrient self-sufficiency the produced food from the crop and livestock component of the farms (Table 5). Inputs costs and N outflows of the system were not considered as potential determinants due to time limitations.

Table 5: Potential determinants identified per indicator.

Indicator	Potential determinant	Unit
Farm income per capita	Income through sales of cotton, maize, sorghum, millet, groundnut, cowpea, livestock products, fodder surplus and income through self-employment	$PPP year^{-1}
Labour productivity	Total farm income	$PPP year^{-1}
N balance	Inflow of N through I1 cotton seed cakes, I2 pasture, I3 N fixation, I4 deposition, I5 fertilizer	kg year^{-1}
Nitrogen use efficiency	Inflow of N through I1 cotton seed cakes, I2 pasture,I3 N fixation, I5 fertilizer	kg year^{-1}
Calorie, protein, iron and zinc self-sufficiency	Total production of cotton, maize, sorghum, millet, groundnut, cowpea, milk	kg year^{-1}

The potential determinants were averaged over all farms and seasons and then plotted against farm performance in the related indicator for all scenarios. Furthermore, the sum of all potential determinants was calculated for each indicator respectively. To assess the individual impact of each potential determinant on the related indicator performance, the individual share of each potential determinant was estimated. Labour intensity was excluded as the key drivers were directly identified by evaluating in which fortnights labour intensity is the highest and how performance changes with different household sizes.

For the third research questions, the identification of the indicator and domain trade-offs, all scaled indicators were considered to capture all relations (Table 18, Appendix). The scaled indicators and eventually the scaled domains were plotted against each other to identify trade-offs.

3 Results

In the first parts of the results the extension of the scenario analysis of Falconnier et al. (2018) is described. In the second part the results of the SI assessment for both time points are displayed.

3.1 Extension of the Scenario Analysis of Falconnier et al. (2018)

3.1.1 Addition of a Policy Intervention

One new policy intervention was added. A policy trend affecting the producer prices for cereals is based on the introduction of inventory credits. Producer prices in Mali for sorghum and millet are highly volatile during the year. The ample availability during the post-harvest season and after high yielding “good seasons” strongly influences the small markets in sub-Saharan Africa (Sanders & Shapiro, 2006). Hence, shortly after the harvest, the prices are at their annual low while increasing distinctly over the rest of the year. A potential way of tackling this is to give farmers the opportunity to store their cereals and avoid selling their products during the saturated post-harvest period but during periods with a higher demand for cereals. In Niger farmer organisations act as warehouse operators and collect and buy cereal grains from farmers at low post-harvest prices and store the grain. Later the organisations sell the higher valued grain for a higher price and share the additional earnings with the farmers. The organisations are supported by financial inputs of NGOs or governmental projects (Sanders & Shapiro, 2006). Based on that, a policy trend P5 was formed, which takes over the concept of inventory credits. It is assumed that farmers get the highest potential prices of the year, while for all other policy interventions the lowest producer prices during the harvest period were used (Table 6). The policy intervention is placed in sixth place to minimize the trade-off between consumption and sale of cereals.

3.1.2 Adaptation of the Agricultural Intervention Strategies

No additional agricultural intervention strategies were formed in this research. Still their impact was extended to other domains of SI.

The strategy A2 was extended to explicitly measure the impact of improved livestock management on the environmental domain. For the stall-fed cattle it was assumed that 100 % of the manure is collected, whereas for the remaining cattle the standard collection rate of 43 % was assumed. Additionally, it was assumed that farmers also improved their pit storage to reduce N storage losses. The upgrade of the pit storage was assumed to reduce the N loss during storage from 40 % to 20 %, the lower range of observed N losses in pit storage reported by Dong et al. (2014) (Table 6). Despite the intensification of the cattle herd component no changes in labour demand were assumed in A2, as no indication for this was found in literature and personal communication with Ousmane Sanogo (2019). Hence, a man-day per day was assumed for cattle husbandry, including both husbandry practices. Additionally, due to increased fodder availability it was assumed that farmers are not dependent on transhumance anymore. Thus, from A2 onwards all cattle stay on farm all around the year. Thus, produced N inputs and N outputs as well as revenue streams of all owned cattle throughout the year were included in the SI assessment.

In A3 small-scale mechanisation on cotton was introduced. It promoted timely crop management but also reduced the labour demand as agricultural tasks are outsourced to external services. It was assumed that these tasks are performed during crop establishment since small-scale tractors can mostly be used during this season (Baudron et al., 2015). This included the transport and spread of organic manure and ploughing of the cotton fields. The household’s labour demand for these tasks was thus eliminated. Next to fertilizer for cotton and maize also the other cereals received fertilizer in A3 (Table 6), resulting in increased labour demand for these crops. For the application of fertilizer to the cropping area of sorghum and millet the same amount of man-days ha^{-1} as for maize was assumed.

3.1.3 Cereal Market Development Scenario

Based on the addition of P5 a new scenario (S5) was formed by combining P5 with A3 (Figure 7). In this scenario farmers’ income increased by getting better prices on the markets through inventory credits.

Results

Table 6: Key variables and their quantification in the current (2013) and future (2027) situation for hypothetical policy interventions (P0 ‑ P5) and hypothetical changes in agricultural practices (A0 - A3). In this research added or adapted aspects are written in bold font.

Hypothetical policy interventions	Key variables	Trend	2013	2027	Reference used to build the trend
No input/output subsidy for cotton production (P0)	Price paid to farmer for cotton (fcfa/kg)	Decrease	250	183	Coulibaly et al. (2015)
	Cost of fertilizer bag for cotton (fcfa/kg)	Increase	12500	17500	Coulibaly et al. (2015)
Input/output subsidy for cotton production (P1 – P4)	Price paid to farmer for cotton (fcfa/kg)	No change	250	250	Village survey data
	Cost of fertilizer for cotton (fcfa/kg)	No change	12500	12500	Village survey data
No input/output subsidy for milk production (P0 - P1)	Price paid to farmer for milk (fcfa/kg)	No change	250	250	Village survey data
	Cost of cotton seed cake (fcfa/kg)	Decrease	170	50	Village survey data
Input/output subsidy for milk production (P2 – P4)	Price paid to farmer for milk (fcfa/kg)	Increase	250	400	Aparisi et al. (2012)
	Cost of concentrates (fcfa/kg)	Decrease	170	50	Kelly et al. (2010)
No family planning programs (P0 – P2)	Net fertility rate (%)	Current rate	3.4	3.4	World Bank
Family planning programs (P3 - P4)	Net fertility rate (%)	Lower rate	2.2	2.2	Ministère de la santé et de l'hygiène publique (2014)
Limited job creation outside agriculture (P0 – P2)	Rural urban migration (HRE-LH, HRE, MRE, LRE) (%)	Current rates	0; 1.7; 1.2; 2.8	0; 1.7; 1.2; 2.8	SEP data
Important job creation outside agriculture (P3 - P4)	Rural urban migration (HRE-LH, HRE, MRE, LRE) (%)	Higher rates	2.8; 2.8; 2.8; 2.8	2.8; 2.8; 2.8; 2.8	African Development Bank (2012)
Integrated pest management programs for cotton production (P4)	Existence of the programs	-	No programs	Programs in place	Silvie et al. (2013)
Incentive subsidy for the development of private small-scale (P4) Mechanization services (P4)	Existence of the subsidy	-	No subsidy	Subsidy	de la Croix et al. (2011) and Baudron et al. (2015)
Fertilizer subsidy for sorghum and millet (P4)	Cost of fertilizer for sorghum and millet (fcfa/kg)	Decrease	17500	12500	Coulibaly et al. (2015)
Organisations provide the facilities to store cereals and give inventory credits (P5)	**Prices for, maize, sorghum and millet (fcfa/kg)**	**Increase**	**94; 110; 135**	**122; 160; 215**	**Sanders and Shapiro (2006)**

Hypothetical change in agricultural practices	Key variables	Trend	2013	2027	Reference used to build the trend
Decreasing cotton cultivation (A0)	Cotton share of cropland (HRE-LH, HRE, MRE, LRE) (%)	Decrease	31; 32; 21; 24	22; 11; 5; 8	Falconnier et al. (2015)
	N input on cotton, maize, sorghum, millet (kg/ha)	**Decrease**	**43; 60; 0; 0**	**43; 40; 0; 0**	**Falconnier et al. (2015)**
No change in farmer practices (A1)	Cotton share of cropland (HRE-LH, HRE, MRE, LRE) (%)	No change	31; 32; 21; 24	31; 32; 21; 24	Falconnier et al. (2018)
	N input on cotton, maize, sorghum, millet (kg/ha)	**No change**	**43; 60; 0; 0**	**43; 60; 0; 0**	**Falconnier et al. (2018)**
	Percent maize intercropped with cowpea (%)	**No change**	**0**	**0**	**Falconnier et al. (2018)**
	Small-scale mechanisation for cotton operations	No change	0	0	Falconnier et al. (2018)
	Percent cows in stall feeding	No change	0	0	Falconnier et al. (2018)
	Integrated Pest Management on cotton	No change	No	No	Falconnier et al. (2018)
	Practice of transhumance if cattle herd exceeds 20	**No change**	**Yes**	**Yes**	**Ba & Lesnoff (2011)**
Diversification with legumes (A2 - A3)	Percent maize intercropped with cowpea (%)	Increase	0	100	Falconnier et al. (2016)
Intensification of livestock production (A2 - A3)	Percent cows in stall feeding (%)	Increase	0	0 - 100; depending on fodder availability	De Ridder et al. (2015)
	Manure collection in current farmer practice and stall feeding (%)	**Stable**	**43, 100**	**43, 100**	**Rufino et al. (2005)**
	Percent in manure storage losses during stall feeding (%)	**Decrease**	**40**	**20**	**Dong et al. (2014)**
	Practice of transhumance if cattle herd exceeds 20	**-**	**Yes**	**No**	**Assumed in this study**
Closing yield gap (A3)	**N input on cotton, maize, sorghum and millet (kg/ha)**	**Increase**	**43; 60; 0; 0**	**90; 110; 150; 150**	**Falconnier et al. (2018)**
	Integrated Pest Management on cotton	Increase	No	Yes	Silvie et al. (2013)
	Small-scale mechanisation for cotton operations	**Increase**	**No**	**Yes**	**Baudron et al. (2015) and de la Croix et al. (2011)**

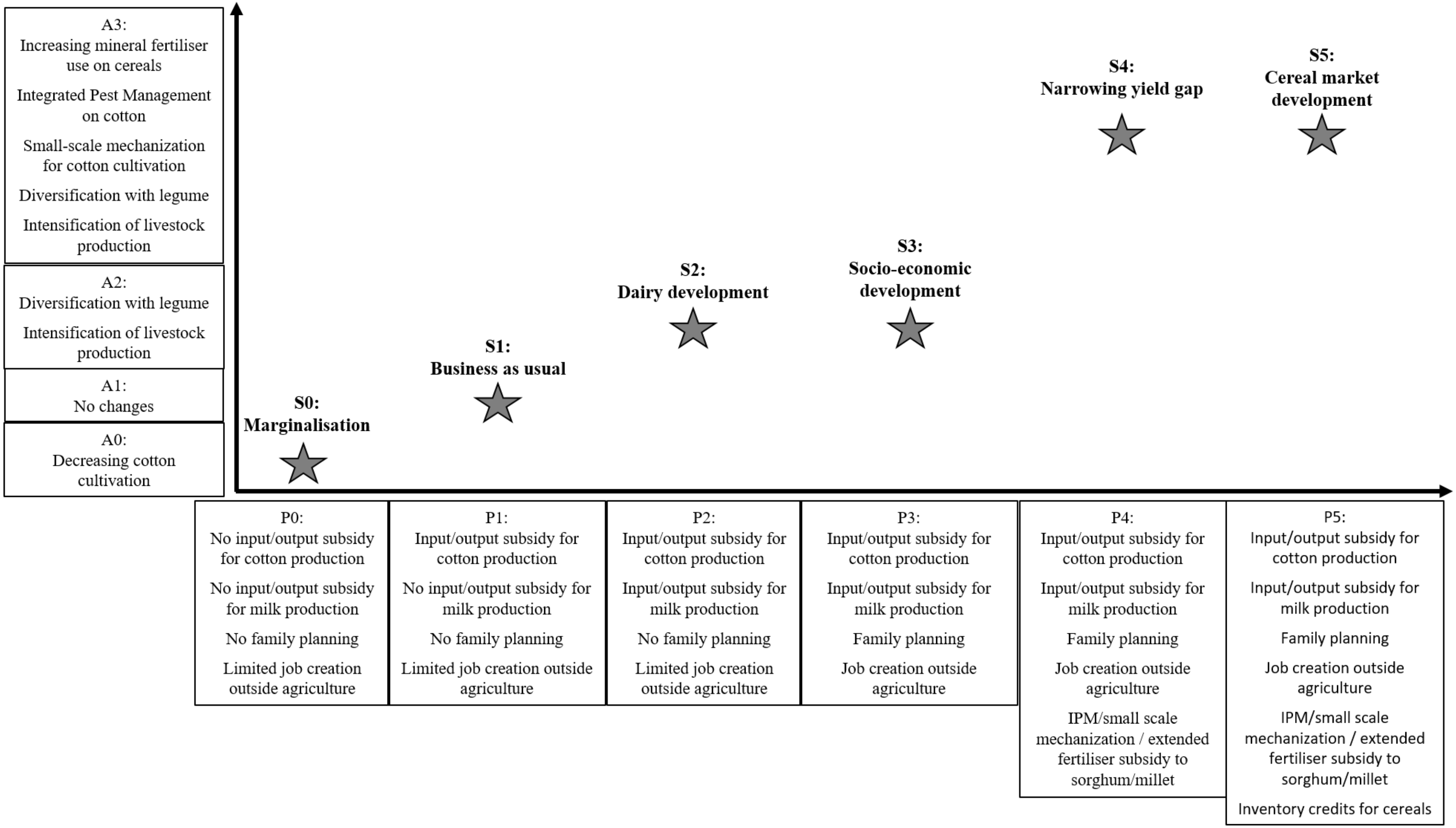


Figure 7: Six future scenarios based on six incremental policy interventions along the X axis and four agricultural intensification strategies along the Y axis; abbreviations: Integrated Pest Management (IPM).

3.2 Sustainable Intensification Assessment and Trade-off Identification

First, the results of the individual indicators for the different time points and scenarios will be described targeting the first research question. The second part relates to the second research question to identify the key external and internal drivers. Lastly, result of the trade-off analysis with the scaled indicators are presented targeting the third research question.

Before going further into the results of the SI assessment, it must be mentioned that a distinct variation of the composition of the farm types occurred in the different scenarios (Table 7). 54 MRE farms became HRE farms in S0, S1 and S2, due to increased household size (Figure 8) and number of active workers. From S3 and in the following scenarios only 74 of the original 89 HRE farms continued being HRE, the rest became MRE farms. These changes impacted the indicator results of both farm types. No other shifts in farm types and other farm type discriminating parameters were observed (Table 7).

Table 7: Distribution of the farms over the farm types for the baseline situation and each scenario (S0 – S5). n(farms) = 411.

Scenario	**HRE-LH**	**HRE**	**MRE**	**LRE**
Baseline	37	89	234	51
S0	37	143	180	51
S1	37	143	180	51
S2	37	143	180	51
S3	37	74	249	51
S4	37	74	249	51
S5	37	74	249	51

The household size increased in (S0 – S2) the most for the HRE-LH farmers who have a migration rate of 0 % (Table 6). For the other farm types the increase was less severe due to their larger rural to urban migration rates (Table 6). Due to family planning and an increased migration rate (Table 6) in S3 the household sizes in S3 – S5 were comparable to the one in the baseline situation (Figure 8).

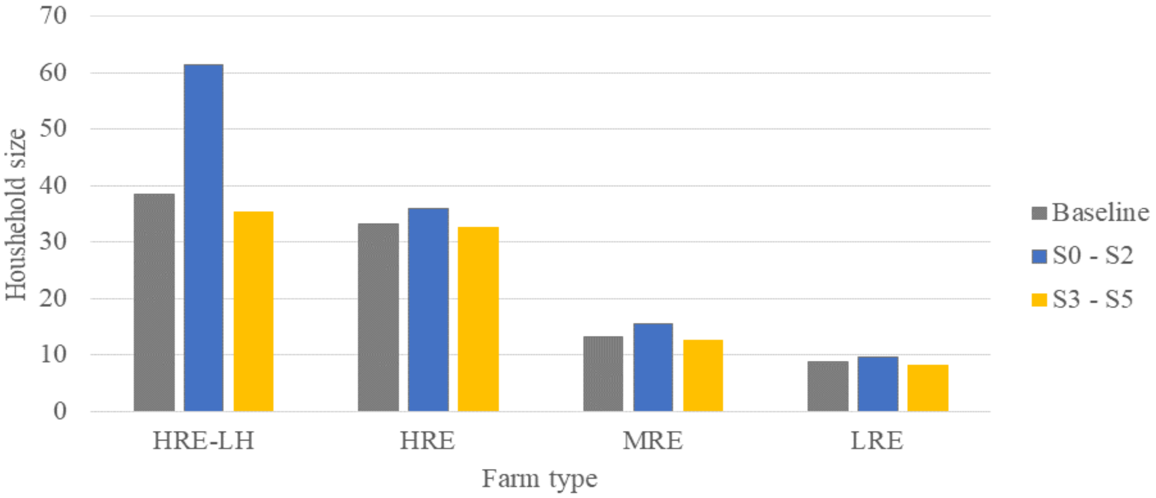


Figure 8: Average household size presented for farm types as well as for the baseline situation and for each scenario (S0 – S5); n(farms) = 411.

3.2.1 Farm Performance According to Different Indicators

In the baseline situation the majority of the farms (64 %) had a farm income per capita below the poverty line (Figure 9). In the “Marginalisation” scenario (S0) the income per capita decreased and 92 % of all farms were below the poverty line. A similar trend was observed in the scenario “Business as usual" (S1): the farms also performed worse than in the baseline situation with 86 % of the farms living below the poverty line. The “crop-livestock integration” (S2) showed with 35 % almost the same number of farms above the poverty line as in the baseline situation (Figure 9). The majority of the households (62 %) were lifted out of poverty in scenario S3 due to family planning and job creation outside agriculture. By narrowing the yield gap (S4) 84 % of the households performed above the poverty. Inventory credits (S5) raised the number of farms above the poverty line to 95 % (Figure 9).

The NUE and labour intensity of the farms will be described individually and thus not included in Figure 9.

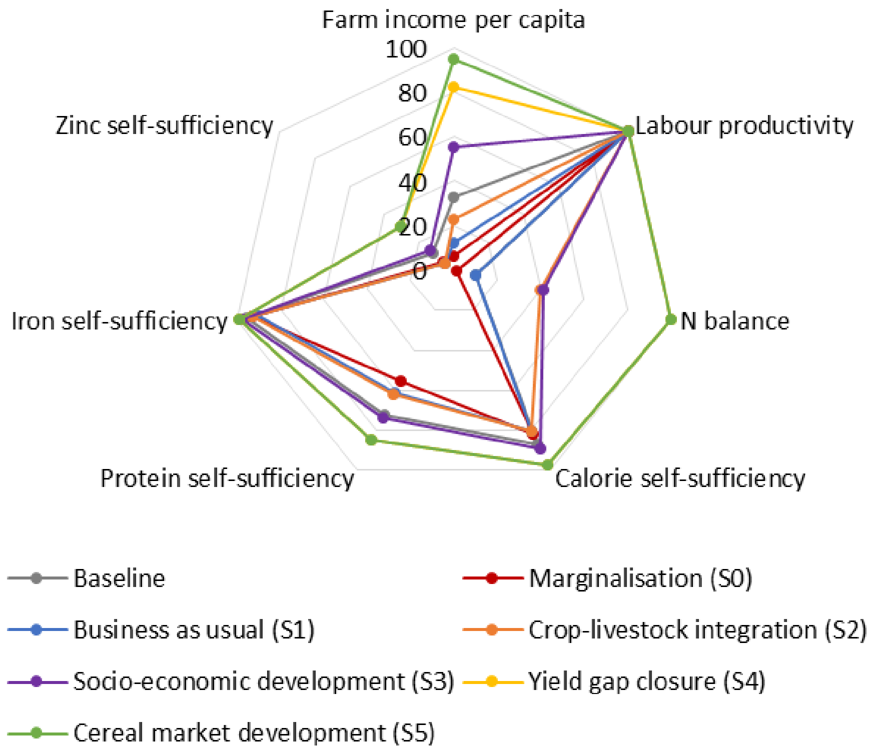


Figure 9: Percentage of farms that perform, averaged over all seasons, above the indicator specific threshold presented for the baseline situation and for each scenario as well as for various indicators. n(farms) = 411

Furthermore, it must be indicated that the between-farm variation of the indicator farm income per capita increased in S4 and S5 (Figure 12).

Looking at farm income per capita of the farm types over the seasons the described tendency was confirmed generally (Figure 10). But it has to be noticed that the positive effects of the last three scenarios were the highest for the group of MRE and HRE-LH farmers since the LRE and HRE farms lay still below the poverty line in most seasons in S3.

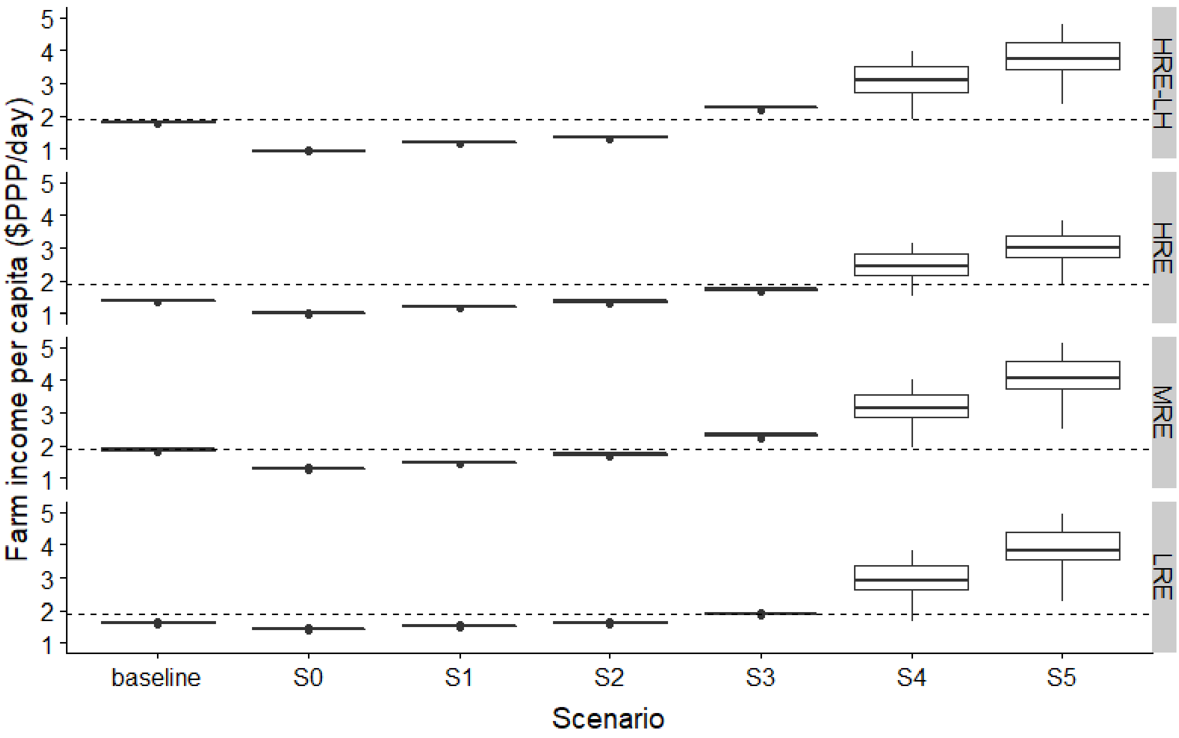


Figure 10: Boxplots showing farm income per capita in all seasons presented for farm types as well as for the baseline situation and for each scenario (S0 – S5). n(farms) = 411; n(seasons) = 26; abbreviations: US dollar purchasing power parity ($PPP).

All farms in all scenarios performed above the threshold of labour productivity in Mali (Figure 9). The labour productivity was similar in the baseline situation and the first two scenarios. From S3 onwards labour productivity and the variation between the farms rose with each subsequent scenario (Figure 12).

Scenario S1, S3 and S5 did not impact the N flows and therefore, they were not considered in the depiction of the seasonal variation of the N balance (Figure 11). In the baseline situation and in S0 and S2 the N balance of the farms was on average around zero (Figure 11 and Figure 12). However, the majority of the farms lay slightly below this threshold in all scenarios (Figure 11 and Figure 12). Due to increased fertilizer inputs in S4, the N balance of all farms increased distinctly and 100 % of the farms had a positive N balance (Figure 9, Figure 11 and Figure 12). Furthermore, the variation between the different farms increased (Figure 12). Looking at the N balance expressed per hectare little difference between the farm types can be observed (Figure 22, Appendix).

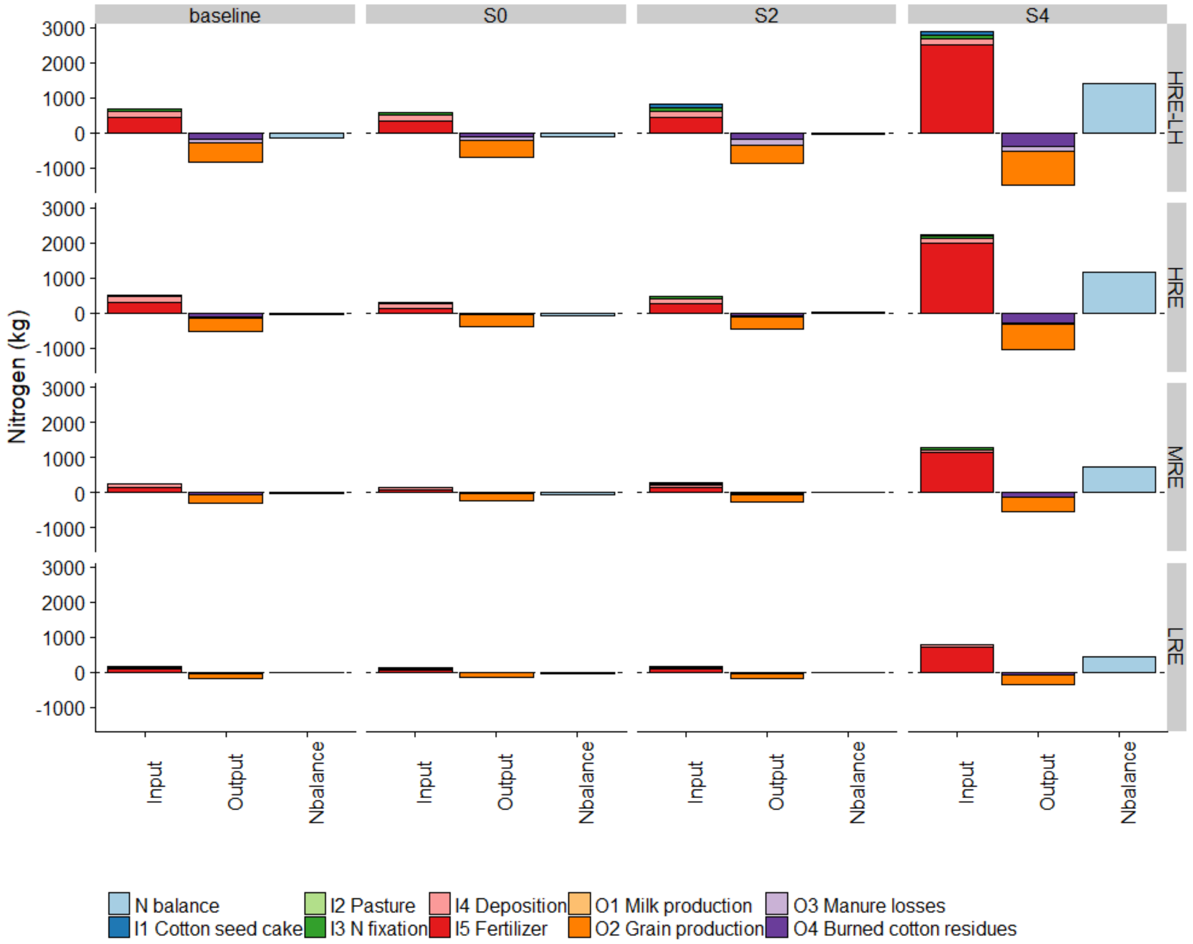


Figure 11: Amount of nitrogen (kg) entering the system through N inflows (I1 – I5) and leaving the system through N outflows (O1 – O4) and the resulting N balance presented for the baseline situation and for the relevant scenarios (S0; S2 and S4). The values of every farm are averaged over all seasons and then within farm types.

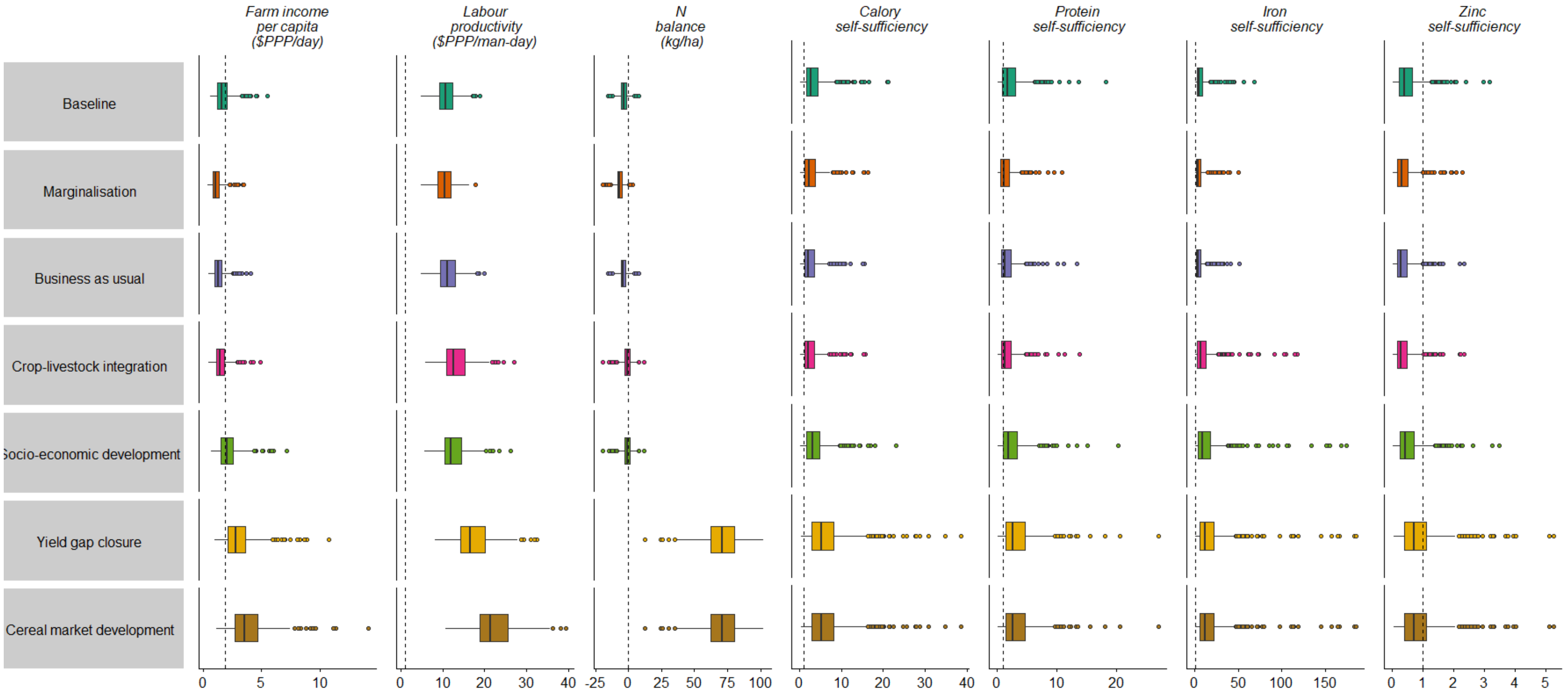


Figure 12: Boxplots showing the performance of various indicators of all farms presented for the baseline situation and for each scenario. The boxplot is formed with the average of every farm over all seasons. The dotted lines visualise the indicator-specific thresholds that indicate a satisfactory performance. n(farms) = 411; abbreviations: US dollar purchasing power parity ($PPP).

Scenario S1, S3 and S5 did not impact the N flows and therefore, they were not considered referring to the NUE. In the baseline and the considered scenarios the majority of the farms had a NUE outside of the desired range of 0.5 and 0.9 (Figure 13). In the baseline situation the NUE of most farms was outside the desired range (94 %) with the exception of some farms (Figure 13 a). In S2 more farms 77 % had a NUE in the desired range (Figure 13 c). In S0 all farms have a NUE higher than the 0.9 and thus risk soil mining (Figure 13 b). On the contrary in S4 the NUE of all farms shifted below 0.5 and they thus risk an inefficient use of N inputs (Figure 13 d).

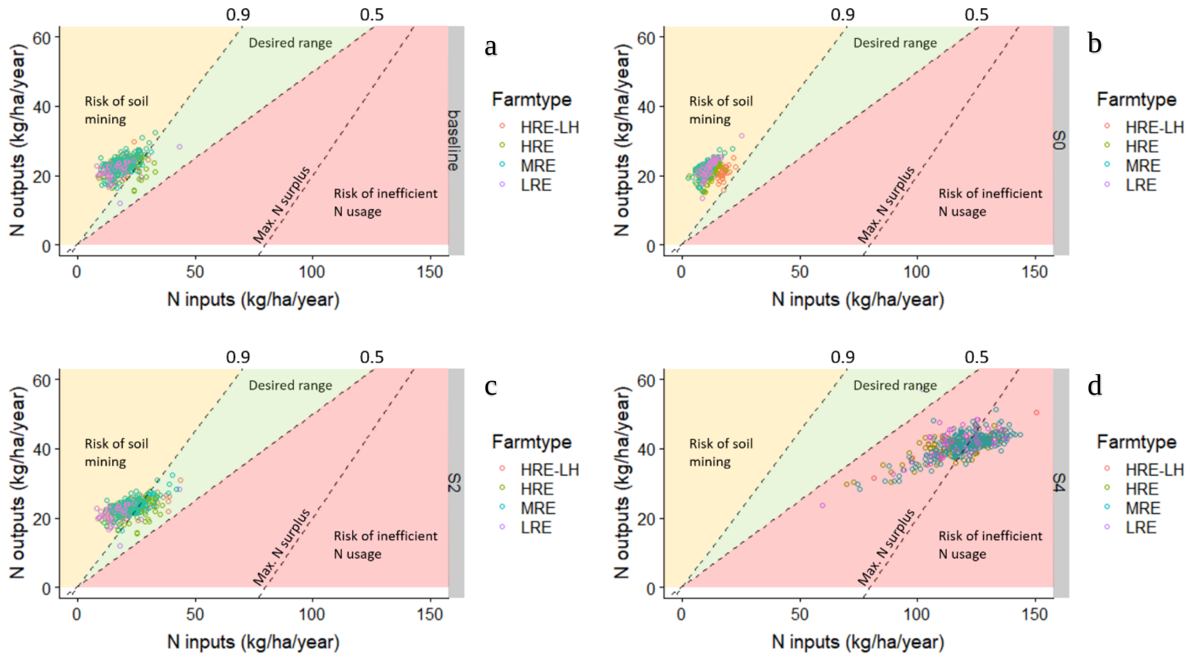


Figure 13: Distribution of the nitrogen use efficiency (NUE) of all farms averaged over all seasons and presented for farm types as well as for the baseline situation and for relevant scenarios (S1, S0, S2 and S4 plotted in the NUE scheme based on EU Nitrogen Expert Panel (2015). The scheme indicates if farms have a desired NUE, risk of soil mining or tend to use the applied nitrogen inefficiently. n(farms) = 411.

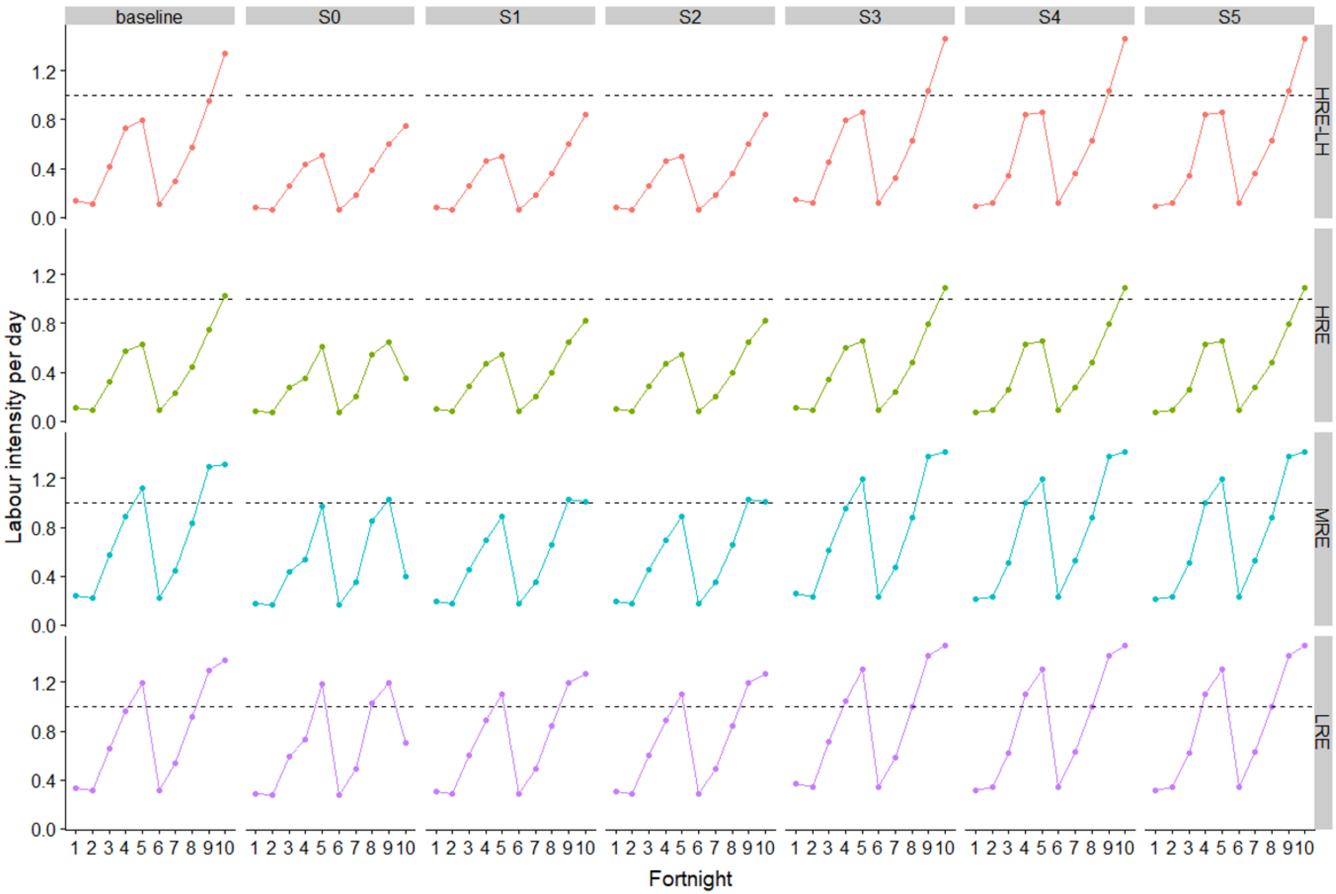
The majority of the farms (97 %) was calorie self-sufficient in the baseline situation (Figure 9). This was the same in S0, S1, S2 and S3 with respectively 94 %, 94 %, 94 % and 98 % of the farms being self-sufficient. In the last two scenarios 100 % of the farms were lifted far above the threshold of calorie self-sufficiency. No major differences between the farm types in calorie self-sufficiency were detected and seasonal rainfall has only a minor impact on the indicator as it did not put them below the threshold (Figure 23, Appendix). This can be stated for all indicators of nutrient self-sufficiency.

In the baseline situation 86 % of the farms were protein self-sufficient (Figure 9). In S0 the proportion of protein self-sufficient farms reduced to 60 %. In S1 and S2 this rose slightly to roughly 70 % again. In S3 with the reduction of household members 91 % were self-sufficient from S4 onwards 97 %.

All farms were iron self-sufficient in all scenarios (Figure 9). From S2 until S4 iron self-sufficiency even rose.

In the baseline situation and S0 - S3 none or almost no farms were zinc self-sufficient (Figure 9). In S4 and S5 the proportion of zinc self-sufficient farms increased but remained unsatisfactory at 18 % (Figure 9).

The variation between the farms in nutrient self-sufficiency performance increased with the subsequent scenarios (Figure 12).



Fortnight	Time window of the year	Crop management activities
1	15/04 - 30/04	Manure transport
2	15/05 - 31/05	Manure spreading
3	01/60 - 15/06	Hand sowing Ploughing Herbicide application
4	15/06 - 30/06	Ploughing Herbicide application Hand sowing Fertilizer application 1
5	01/07 - 15/07	Draught weeding 1 Draught weeding 2
6	15/07 - 31/07	Fertilizer application 2
7	15/08 - 31/08	Fertilizer application 2 Insecticide application Ridging
8	01/09 - 15/09	Harvest
9	01/10 - 15/10	Harvest
10	01/11 - 15/11	Harvest

Figure 14: Visualisation of the labour intensity per day for the observed fortnights presented for farm types as well as for the baseline situation and for each scenario (S0 – S5). The single values of every farm are averaged over all season and then within farm types. The dotted line marks the threshold of labour shortages. n(farms) = 411.

The labour intensity per day of individual fortnights was highly variable over the year (Figure 14). The farm labour was most intense in the fortnights 4, 5, 8, 9 and 10, in which farmers execute ploughing, weeding and harvesting. In the most other considered fortnights the labour intensity of the farms was low and lay far below the threshold when resources and workload equal each other. The LRE farms had the highest labour intensity throughout all fortnights compared to the other farm types (Figure 14). The MRE farms faced the second-largest labour intensity, followed by the HRE-LH. The HRE only had a labour intensity below one in the baseline situation and in S3 – S5 of the 10th fortnight (Figure 14).

In the baseline only the LRE and MRE farms experienced labour shortages in other fortnights (5 and 9) than fortnight 10. S0 in general reduced the labour intensity. The area exchange of cotton and sorghum distinctly reduced the labour intensity for all farm types in fortnight 10. However, it created a period of labour shortage for the LRE farms during fortnight 8. In S1 and S2 the farms faced the lowest labour intensity. LRE farms faced labour shortage in fortnight 5, 9 and 10 and MRE farms in fortnight 9 and 10. The other farm types faced no labour shortages. In S3, after the introduction of family planning and the creation of jobs outside of agriculture, the labour intensity rose. LRE faced labour shortages in fortnight 4, 5, 8, 9 and 10, MRE in fortnight 5, 9 and 10, HRE only in fortnight 10 and HRE-LH in fortnight 9 and 10. In S4 and S5 no major changes in labour intensity were observed compared to S3, except for the MRE which also faced additional labour shortage in fortnight 4 (Figure 14).

In order to give an impression of changes per domains under the influence of the different scenarios, the overall performance of all observed indicators per domain was calculated (Figure 15). Besides the baseline the subsequent scenarios (S2 – S5) performed better in most domains, except for in the environmental domain in which S2 and S3 performed best. Within the individual domains the scenarios differed in their relative increase in performance. The highest performance increase was found in the economic domain (Figure 15).

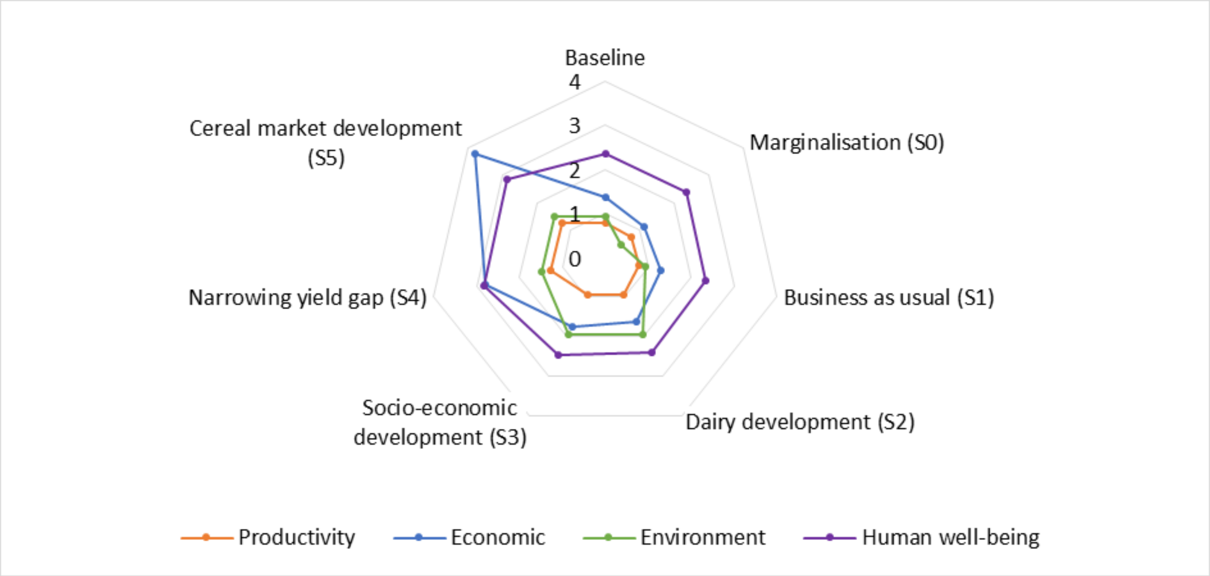


Figure 15: Performance of different domains presented for the baseline situation and for each scenario (S0 – S5) as average across the related scaled indicators. The indicators of each domain have been averaged over all farms and all seasons. These average values are scaled with the worst performance at 0 and best performance at 10, n(farms) = 411.

3.2.2 Key External and Internal Drivers

A higher income generated through cotton sales (Figure 16 a), cereal sales and the sale of livestock products, including spare fodder, had a positive impact on farm income per capita. One exception was the increase in sorghum income in S0 which resulted in a decreased farm income per capita. The higher income generated through sorghum sales in S0 was not able to outweigh the reduction in cotton and maize production (Figure 12). Looking at the proportion of the income generated in the most scenarios (baseline, S1 - S5) cotton was the crop that most strongly determines income (Table 19, Appendix). The share of cotton income of the total income from various sources was strongly affected by the changes in S0 and S4. The amount of income generated through on-farm self-employment had a negative relation with farm income per capita. This again indicated that an increase in household size negatively affected the indicator since the income generated through self-employment only increased with a rising number of household members.

Between total farm income and labour productivity existed a strong positive relation (Figure 16 c). But the changes in labour input between the scenarios were so little that this was much less defining.

Fertilizer as input was by far the potential determinant with the biggest share of all N inflows for both the N balance and NUE (Table 20, Appendix; results not shown). Looking at the effect of an increase of the other potential determinants very little impact on the N balance was observed. Fertilizer, however, showed a very strong positive relation with the N balance (Figure 16 b). A similar pattern was shown in the NUE (results not shown). A high increase in fertilizer application resulted in a strong increase in NUE.

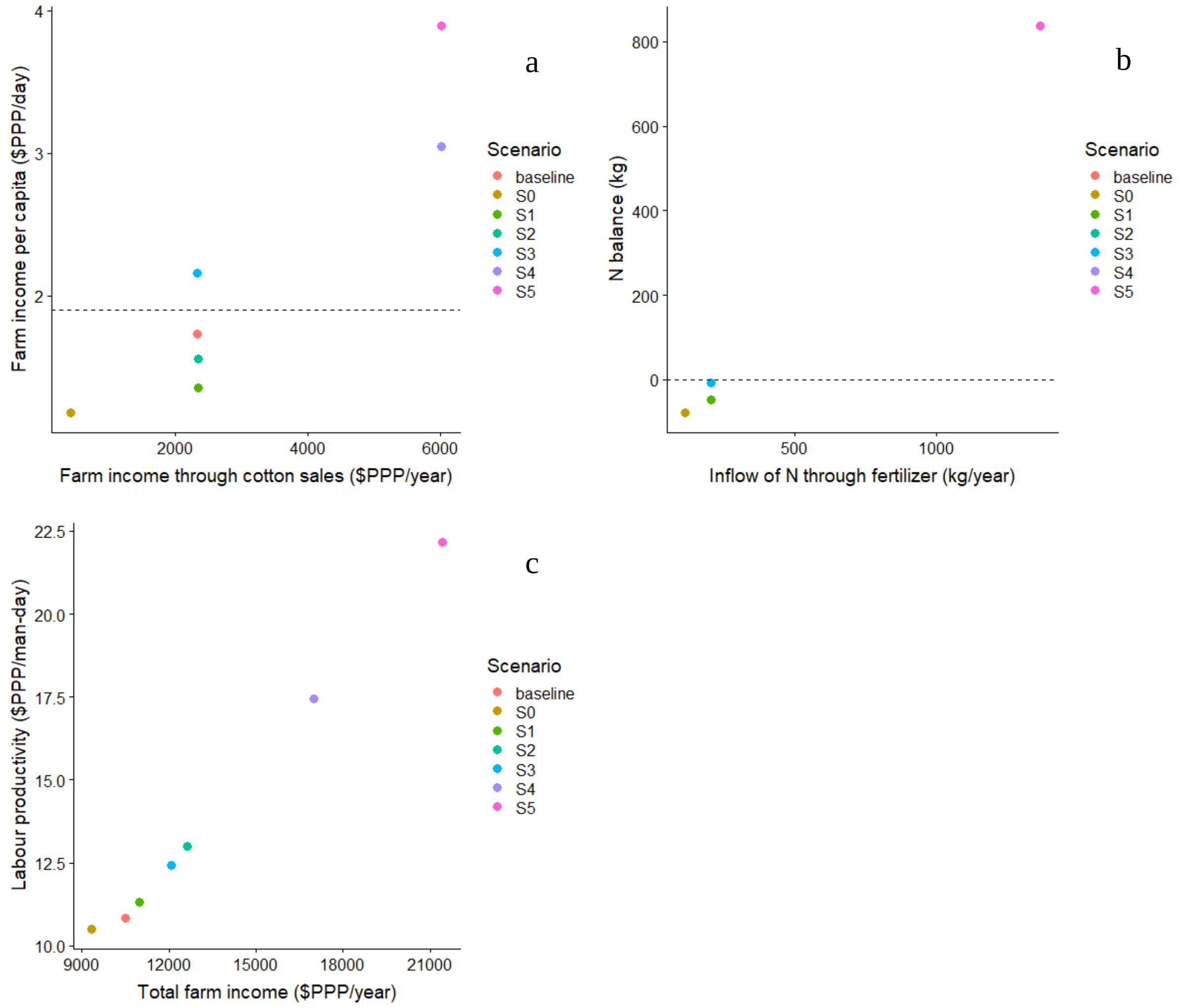


Figure 16: Scatterplots: a) determinant “Income through cotton sales” against indicator “Farm income per capita”; b): determinant “Inflow of N through fertilizer” against indicator “N-balance”; c): determinant “Total farm income” against indicator “Labour productivity”. The coloured dots show the respective values of determinants and indicators for the baseline situation and for each scenario (S0 – S5). The single values are averaged over all seasons and all farms. Dotted lines indicate indicator specific thresholds. n(farms) = 411; abbreviations: US dollar purchasing power parity ($PPP).

The production of the different kinds of cereals and milk had a positive relation with the self-sufficiency of all nutrients. However, it must be said that the crop grains have much bigger share of the total production and thus are defining nutrient self-sufficiency, with maize being the most dominant food crop (Table 21, Appendix).

Additional to the potential determinants, the farm characteristic household size was identified as highly influential. Multiple indicators showed a clear relationship to it (Figure 17 a - c). A higher number of household members had a negative effect on farm income per capita, labour intensity and calorie self-sufficiency.

For labour intensity the amount of household members was most defining as it decreases with a higher number of household members (Figure 17 b). Changes in the household size throughout the different scenarios were shaping the performance of all farm types (Figure 14). In general, farm types with bigger household sizes and lower migration rates (HRE-LH, HRE and MRE) exhibited bigger changes in the household size. Among the different tasks the labour demand during the ploughing, weeding and harvest periods had the largest impact on labour intensity since labour intensity was the highest in these periods (Figure 14).

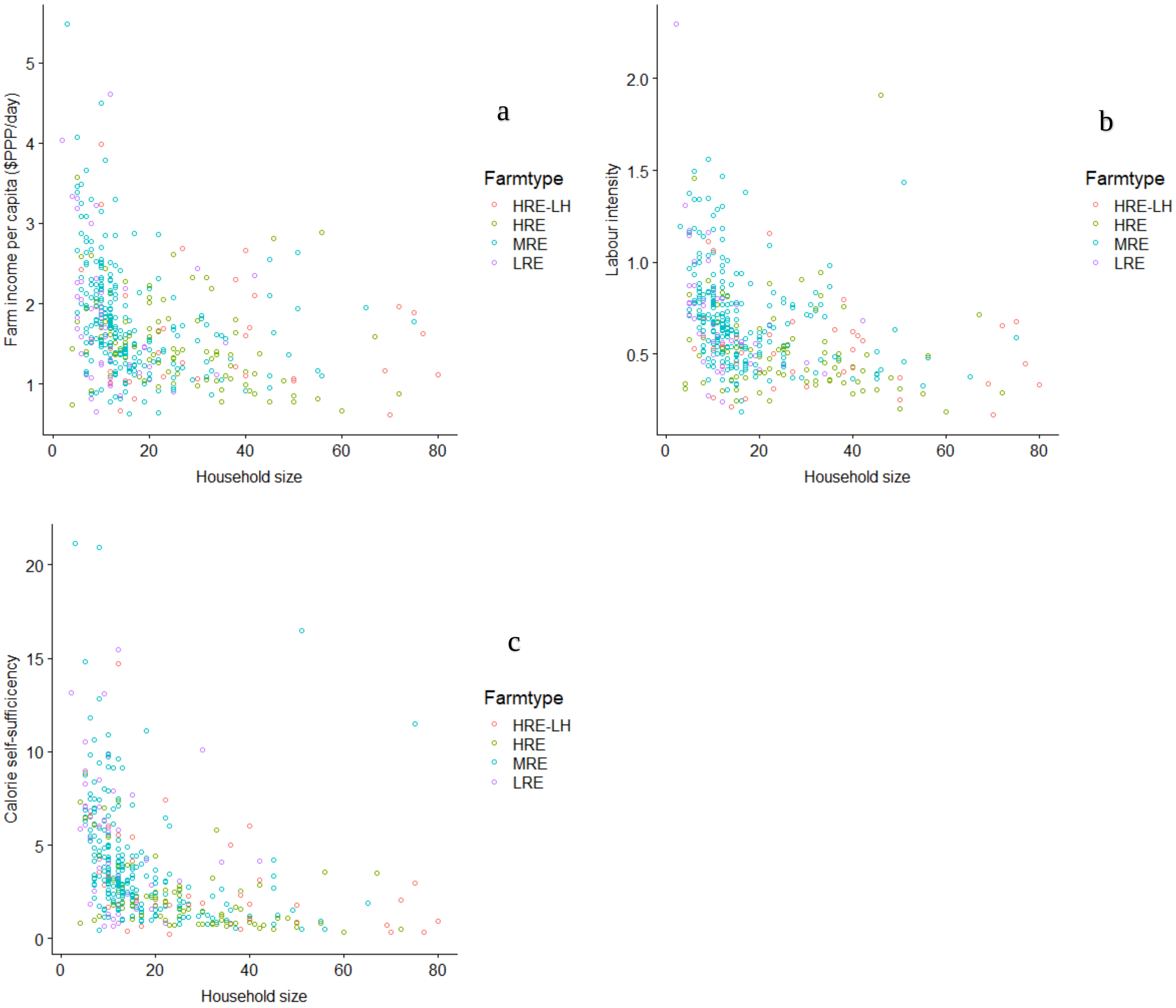


Figure 17: Scatterplots of the household size in the baseline situation against different indicators a): ”Farm income per capita”, b): “Labour intensity” and c): “Calorie self-sufficiency” (representing all nutrients). The dots show the respective values for all farms and the colour refers to the farm type. n(farms) = 411; abbreviations: US dollar purchasing power parity ($PPP).

3.2.3 Trade-off Analysis

For all indicators the best performing farm was usually identified in S5 (Figure 18). The indicators NUE and labour intensity were an exception to this result. Compared to the best performing farm, most farms performed poorly in most of the indicators. The indicator labour intensity in which the farms in all scenarios perform relatively similar to the best performing farm was an exception to this observation (Figure 18).

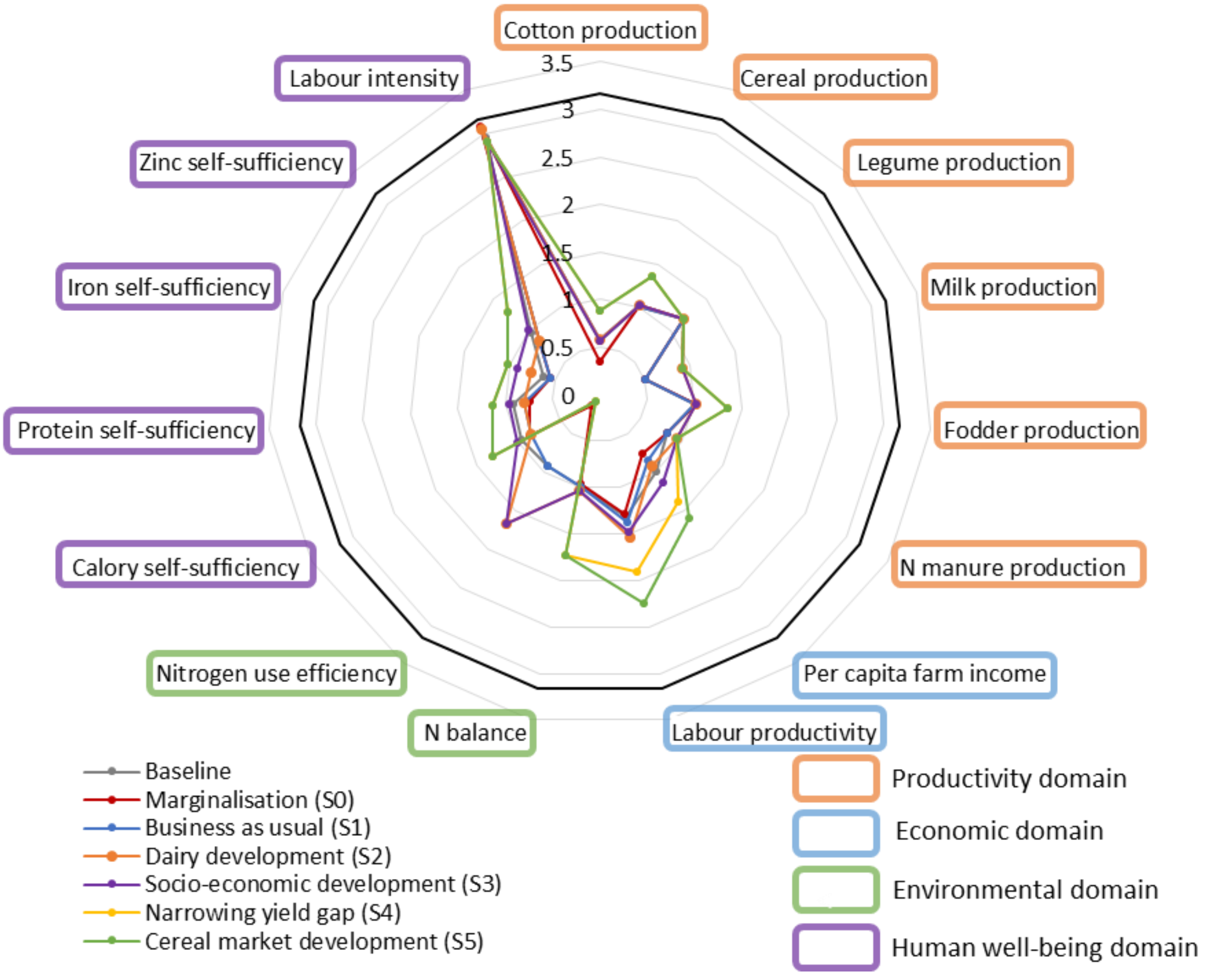


Figure 18: Scaled performance of various indicators presented for the baseline situation and for each scenario (S0 – S5). The various indicators were previously averaged over all seasons and all farms and the square root was calculated. The black line indicates the scaled optimal performance for all indicators (3.16). The square root was taken to improve the depiction of the data. n(farms = 411.

One observed trade-off was between the two indicators of the environmental domain (Figure 19). In the baseline situation and S0 - S3 the farms either mined their systems or lay within the desired range of NUE. The N balance in these scenarios was mostly below the threshold. Contrastingly, in S4 and S5 the N balance increased distinctly above the threshold for all farm types, whereas the NUE lay under the desired range for all farm types.

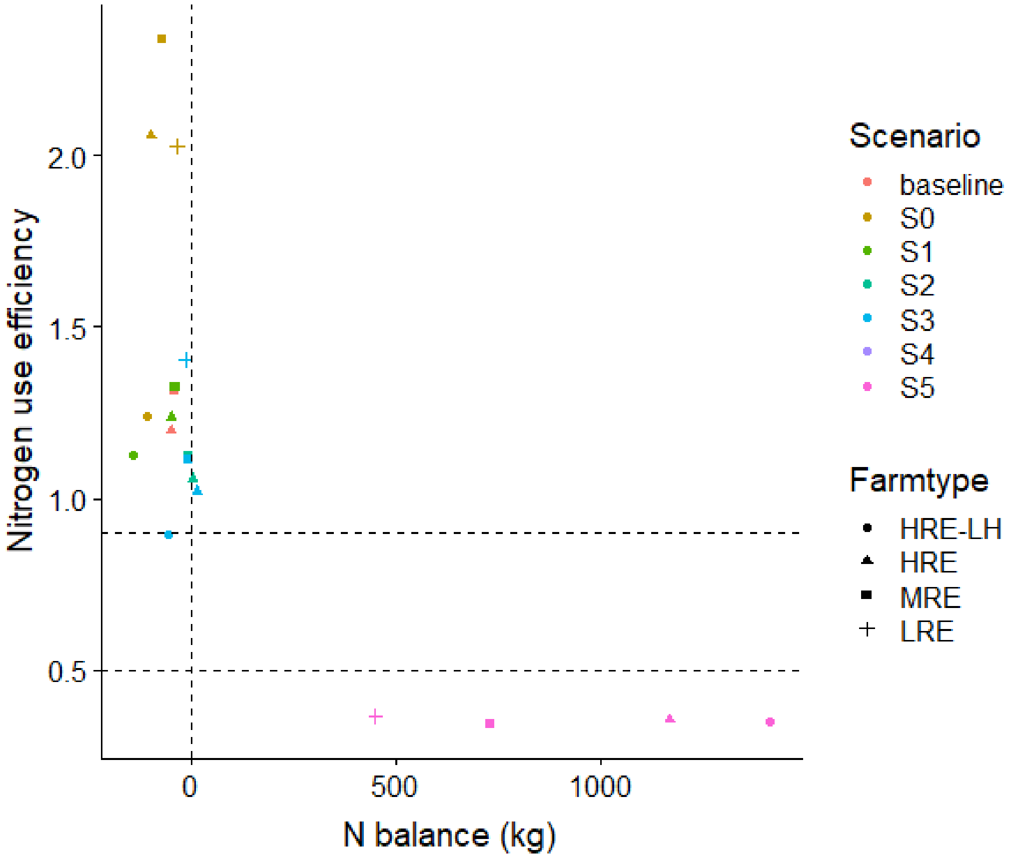


Figure 19: Scatterplots of the N balance against the nitrogen use efficiency presented for farm types as well as for the baseline situation and for each scenario. Dotted lines indicate indicator specific thresholds. n(farms) = 411; abbreviations: High resource endowed large herd (HRE-LH); High resource endowed (HRE); Medium resource endowed (MRE) and Low resource endowed (LRE).

Other direct trade-offs between individual indicators were not identified. Nevertheless, differently from the other indicators, a reduction of labour intensity was positively related with an increasing household size as more workers were available (Figure 17 b). Thus, there was an indirect trade-off between labour intensity and other by the household size affected indicators. A reduction of the household size to improve e. g. farm income per capita resulted in an increase in labour intensity. In general the existing relations between the indicators were most commonly positive (results not shown).

Another type of trade-off was identified between an increase in productivity and the variation in the seasonal performance of the farms. The variation in the indicators farm income per capita, labour productivity, NUE, calorie self-sufficiency, protein self-sufficiency, iron self-sufficiency and zinc self-sufficiency showed a similar pattern over the different scenarios. The variation was slightly smaller in S0 compared to the baseline situation, S1, S2 and S3. In S4 and S5, the variation increased distinctly (Figure 20). This pattern was not observed for the N balance. The variation of the N balance differed more between the different scenarios compared to other indicators, especially for the MRE and HRE farms. Additionally, the variation of the

N balance in S4 and S5 was lower than in the other scenarios which was deviant to the previously listed indicators (Figure 20). Furthermore, it must be noted that the general variation in the baseline situation and the different scenarios for the N balance was higher than for the other indicators.

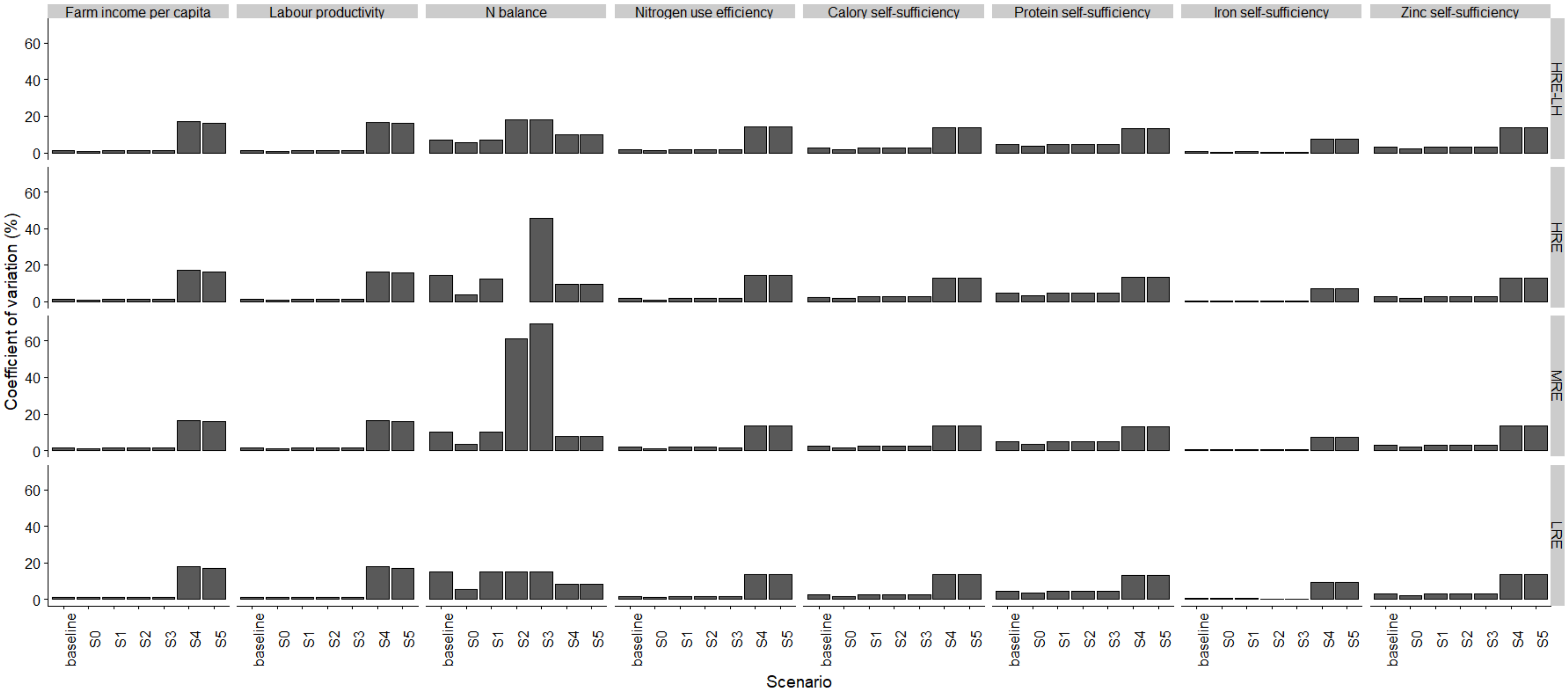


Figure 20: Coefficient of variation of the individual indicators presented for farm types as well as for the baseline and for each scenario (S0 – S5). The coefficient of variation is calculated for every farm over the seasons and then averaged per farm type. The coefficient of variation for N balance of HRE farms in S2 was 158 % and not shown. n(farms) = 411

Results

It was observed that all relations and trade-offs between the different domains originated from changes in the productivity domain. Overall the indicators of the productivity domain showed a simple positive relation with the indicators of the economic and human well-being domain and thus shaped their performance. The performance of the most indicators of the economic and human well-being domain was increasing with increasing productivity (Figure 24, Appendix). The relation between the productivity and the environmental domain, however, was more complex (Figure 21). In the baseline situation and first scenarios (S0 - S3) the relation was still positive. In S0 a decrease in productivity resulted in a strong reduction in performance of indicators of the environmental domain. Until S3 increases in productivity also resulted in an increase in performance of the N balance and NUE. From S4 onwards, however, trade-off between the indicators of the different domains were observed (Figure 21). The increase in performance of the scaled N balance was not able to outweigh the reduction in performance of the scaled NUE (Figure 18). This resulted in a decrease of the environmental performance of the system. The relation of the economic domain to the environmental domain was similar to the relation of the productivity domain and the environmental domain due to the close relation between the productivity and economic domain (Figure 24, Appendix). Further logical relations or trade-offs between the domains were not identified.

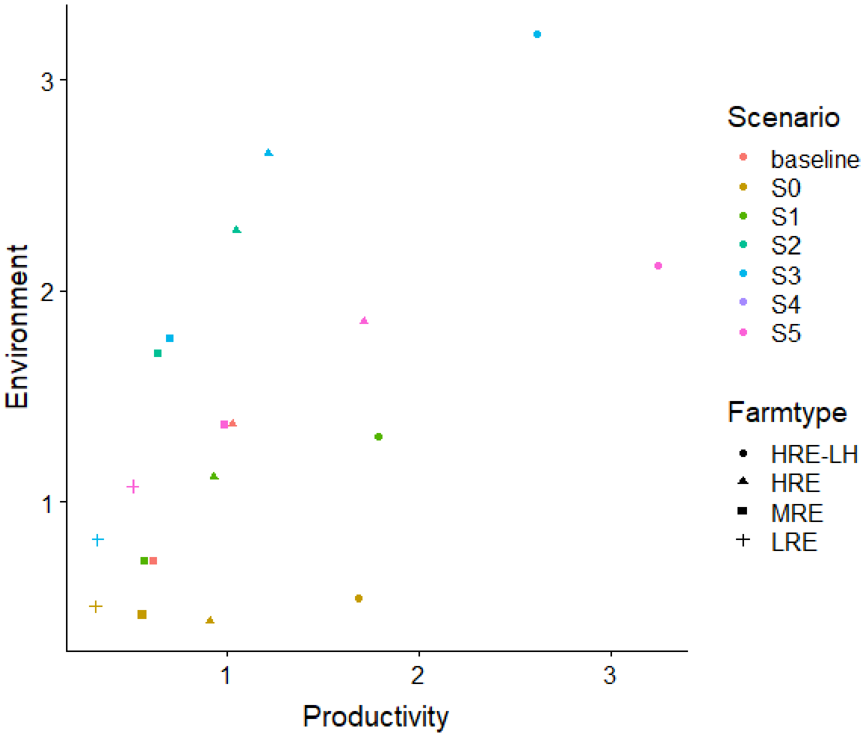


Figure 21: Scatterplots of the scaled productivity domain against the scaled environmental domain presented for farm types as well as for the baseline situation and for each scenario (S0 = S5), the data of the baseline and S1 – S5 can overlap. n(farms) = 411.

4 Discussion

4.1 Contribution of the Research Design

The combination of a SI assessment and a scenario analysis was proven to be a useful decision-making tool. In the reviewed literature the SI in potential futures is rarely assessed. The general aim of SI assessments is to create decision-making guidelines for sustainably intensifying farming systems on all domains. However, it is discussed whether these SI assessments give accurate guidelines or simply create awareness of the system's problems (Garnett & Godfray, 2012; Kiker et al., 2005). Often further steps and studies are required to establish actual guidelines for the future pathways. This calls for a more accurate and faster approach that proposes and evaluates potential guidelines directly; e.g. through ex-ante studies in the form of scenario analyses (Sadok et al., 2008). However, many present scenario analyses in sustainability studies do not match with the wide scope of SI assessments. Often, only one or a limited number of domains are captured e. g. productivity (Traore et al., 2013); productivity and environmental domain (Bruun et al., 2006; Stoorvogel et al., 2004); or productivity and economic domain (Roxburgh & Rodriguez, 2016). In particular, the social and human well-being domain is often lacking in the reviewed literature.

The inclusion of an extensive SI assessment in a scenario analysis can solve both problems. The approach tests the future state of sustainability under the assumptions of the scenarios which allow decision makers to assess the sustainability of numerous guidelines. Thus, actual potential guidelines that may enable SI are identified during the scenario creation. The extensive SI assessment and the capturing of system component interactions through modelling (Ollenburger et al., 2016) allow decision makers to gain the required foresight and understanding about the outcomes of their potential decision (Lancker & Nijkamp, 2012).

After evaluating the trade-offs, advantages and disadvantages regarding SI, the most promising pathways can be identified. Thereby, more accurate guidelines can be formed in a quicker ex-ante approach. This approach is suitable for current decision-making processes, which demand a radical rethinking of the food production in a short amount of time (Sadok et al., 2008; Struik & Kuyper, 2017).

Additionally, many SI constraints are institutional, namely poor conditions of the underlying institutional or market environments (Schut et al., 2016). Scenario analyses, taking policy interventions into account, can address institutional constraints directly and allow the testing of potential solutions in the different scenarios (Falconnier et al., 2018).

Furthermore, the pathways are evaluated based on the future structure of the farming system. This facilitates a more accurate backing for decision-making by considering the temporal dynamics in system transformation. The development of the system over time, e.g. in terms of household size, can impact the sustainability or efficiency of interventions (Falconnier et al., 2018).

4.2 Reflection on the Methodology

4.2.1 Limitations of the Indicator Identification Process

The main limiting factor during the indicator selection was the availability of data and one-sidedness of the available data. The CMDT survey and the on-farm trials were not designed to support a comprehensive, multi-dimensional SI assessment. The focus of both data sources was on farm aspects of the productivity domain, namely farm characteristics and yield data. This limited not only diversity and number of investigable indicators but eventually also the domains that could be considered.

The social domain in SI assessments covers social interactions between the members of the farming community, gender equity and equity of social groups, as well as conflict behaviour (Musumba et al., 2017). Instead of surveys, other methods are required to generate data on these topics, e. g. interviews, focus group discussion or participatory evaluations. Thus, an expedient inclusion of the social domain was not possible in this research. According to the definition of sustainable systems created by Smyth and Dumanski (1993) and adapted from Florin et al. (2014) that includes an *“economically and socially acceptable, stable production level (...)”*, not all aspects of sustainability were assessed in this research.

The applied framework from Florin et al. (2014) clarifies the indicator selection process by generating a holistic view on the system. It points out linkages of the chosen indicators with the systems criteria (outcomes) and the systems overarching domains. However, a main selection criteria is the system specific relevance of the indicators (de Olde et al., 2017). A participatory approach of identifying and engaging main stakeholders in the indicator identification process increases the system specific relevance for all domains and the set of chosen indicators, and reduces the level of subjectivity (Dale et al., 2015; Marinus et al., 2018). This approach was not feasible in this study. Only for farm income per capita and calorie self-sufficiency was farmers’ opinions taken into account by Falconnier et al. (2017). It is likely that a different set of researchers, including the main stakeholders, would have decided for a different set of indicators.

4.2.2 Indicator Scaling

With the exception of NUE, the scaling of the indicators was done by taking the best and the worst performing farm over all scenarios as references. It was assumed that the best performing farm is the actual best attainable performance for each individual indicator. However, it is questionable whether the best performance may serve as the desirable state for all other farms. This counts especially for indicators without a defined threshold (productivity domain).

Displayed in radar charts the scaled averaged farm performance is very low in comparison to the best performance. This might imply that most farms are not sustainable. The radar charts in this study only give an overview of changes in performance of the indicators in the different scenarios, as they do not refer to relatable values like the defined thresholds (Figure 18), e. g. the radar chart indicates a maximum performance for labour intensity and thereby implies sustainability. By relating the indicator to its (subjective) threshold, it was observed that actual sustainability was not reached (Figure 14).

4.2.3 Model Limitations

A model is a simplified representation of reality with the aim of showing only key aspects and eliminating unnecessary components. In this research, the model reflects the farming systems in the Koutiala region. The model specification for this region may have impacted the results of the SI assessment in two different ways.

Firstly, simplifications of farm characteristics or management practices like transhumance may impact the results. For example, the interventions of the dairy development scenario may not be strong enough to completely stop transhumance, as was assumed, which would reduce the real milk production distinctly for HRE-LH farmers compared to what was simulated.

Secondly, the simplification of the indicator calculations impacts the indicator outcomes. The actual state of sustainability in the different time points might distinctly vary from reality, so that conclusions need to be drawn carefully. A monitored positive N balance does not necessarily imply that farmers improved their N management since the implementation of overlooked outflows (N leaching, N runoff, volatilisation and denitrification) will reduce the N balance strongly (see 4.3.2). The same limitations apply for nutrient self-sufficiency in which the consumption of other grown crops or food from the gardens are not considered, so that the simulations may have underestimated actual self-sufficiency. Nevertheless, changes in performance from one scenario to the other were captured.

Another limitation of the model is based on the data availability of drivers. Where possible, the drivers were updated to the most recent available values, however, some drivers were taken from the research of Falconnier et al. (2018) and were based on older data sources. This may have resulted in inconsistencies or wrong depiction of the most recent policy trends, e. g. changes in migration rates and number of remittances.

The inputs of the model were based on trends or rational speculations. This implies a level of uncertainty in the model’s input data and, consequently, of the outputs. In this research this uncertainty was not captured. This could have been done by including a sensitivity analysis. Furthermore, the influence of the drivers can be tested more accurately by testing the variation of each driver individually.

4.3 Farm Performance in Different Domains

Results show that farms performed best in S5 for the majority of the indicators, except for the indicators NUE, labour intensity and with regards to the coefficient of variation (Figure 18). The results of the indicators farm income per capita and calorie self-sufficiency were already discussed in Falconnier et al. (2018) and their findings just briefly summarised here.

4.3.1 Economic Domain: Farm Income per Capita and Labour Productivity

Farm income per capita was mostly influenced by changes in the household size. Differences in scenario effects on farm income per capita between the farm types were related to their migration rate which overrode the changes in the farming practices starting in S2. Farm types with a higher migration rate (LRE and HRE) suffered a smaller decrease in farm income per capita in S0 – S2 (Table 6 and Figure 10). From S3 onwards, farm practices alone defined the differences in economic performance of the farm types.

Labour productivity was for all farms far above the estimated threshold which implies high productivity in the baseline situation and in each scenario (Figure 12). However, this might be misleading, as all processing steps of farm products after the field work were not included in the used labour calendar, even though they are also part of the value creation process. Hence, the actual labour demand was probably strongly underestimated. This issue also impacts the later discussed labour intensity. To assess the full labour productivity the labour calendar must be extended. It must also include transport/walking time to the fields as well as to the markets and the time it takes to process the goods. Additionally, labour productivity was estimated by dividing the total income by the total labour demand of the agricultural system per year. The labour demand, however, differs strongly between different fortnights. Thus, over the year the system seemed highly productive, while during the labour-intensive fortnights this may not be the case. Nevertheless, the increase of income in the subsequent scenarios was able to improve productivity (Figure 16 c). Changes in labour demand due to small-scale mechanisation (Table 6) were not strong enough to impact productivity.

4.3.2 Environmental Domain: N Balance and Nitrogen Use Efficiency

Results of the N balance and NUE showed that the farming systems mined their soils in the baseline situation (Figure 13), which is consistent with findings of Lassaletta et al. (2013) for sub-Saharan African soils. The intensification of the livestock system and diversification with legumes in S2 showed the expected, yet subtle, increase in performance for both indicators (Figure 12 and Figure 13). The total increase of crop available nitrogen through N fixation, increased available nitrogen in manure through the recycling of nitrogen contained in the imported cotton seed cakes and improved capturing and storage capabilities was also very minor (Figure 11), resulting in only small changes in the N balance and no distinct increase in N outputs.

On the one hand a high intensification through fertilizer applications in S4 and S5 boosted the N balance of the systems (Figure 11). On the other hand, it increased the risk of using N inputs inefficiently. This was represented in the trade-off between these two indicators and partly confirms the common presumption that a strong increase in intensification reduces the performance in the environmental domain.

In this study the residual straw of the crops was not considered as output in the NUE, as the straw stays within the system and NUE was estimated at farm level. The farms are probably more efficient in their N usage if the straw would be counted as a productive output, as in Marinus et al. (2016). Whether farms would then be efficiently using the nitrogen can not be said.

NUE decreased from S3 to S4 probably as a result of different reasons. Firstly, strong seasonal rainfalls during the wet season in the Koutiala region (https://de.climate-data.org/, last accessed 23/08/2019) can promote N leaching and N runoff (Schwenke, 2014; Sadras, 2002). Secondly, drought induced stress can affect plants also during the wet season, due to irregular rainfall patterns. Drought stress limits uptake of water soluble nutrients (Pessarakli, 1999). Not absorbed nitrogen was probably lost, resulting in a decreased NUE.

The actual performance of the farms in the N balance was lower than estimated. Nitrogen outflows like N leaching, runoff, volatilisation and denitrification were not considered. Hence, the performance of the environmental domain may be even lower, especially in the last two scenarios.

NUE and N balance may have differed for individual farm components or even fields based on the applied farming practice (Sadok et al., 2008). An observation on farm level does not consider efficiency rates on the individual fields. This is based on the diversity of crop management practices in smallholder systems of sub-Saharan Africa (Vanlauwe & Giller, 2006).

The N balance showed a different pattern in its coefficient of variation than the other indicators (Figure 20). In general the coefficient of variation was extremely high in S2 and S3 for MRE and HRE farmers. For the scenarios S0 - S3, the N balance lay around zero. By dividing the standard deviation by a mean around zero, the coefficient of variation for the N balance reached very high values. Thus, the coefficient of variation does not seem to be a useful tool to visualise the variation for the N balance.

4.3.3 Human Well-being: Nutritional Self-sufficiency and Labour Intensity

From S3 onwards human well-being improved distinctly (Figure 12). For the various nutrient self-sufficiency indicators household size was the dominating factor (Falconnier et al., 2018). It overrode the slightly increased production of livestock products and maize in S2; until S3 the migration rate defined performance differences between the farm types. The yield gap closure in S4 boosted the performance in calorie, protein and iron self-sufficiency (Figure 12).

In contrast, over all scenarios most farms were zinc insufficient. Zinc is hardly present in the considered crops which represented the majority of the produced and consumed food. It is most abundant in red meat, oysters and shellfish (Caulfield et al., 2004). In reality, zinc self-sufficiency may be even worse, even though, the potential consumption of zinc rich red meat was not considered. The problem of zinc self-insufficiency is not only based on low availability but also that the provided food intensifies the problem. A diet rich in staple food such as corn and rice increases the uptake of phytates, which are inhibitors for the uptake of zinc. This is the main reason for zinc deficiencies (Lönnerdal, 2000), although these effects were not simulated.

Labour intensity was an issue for most farm types (Figure 14). It is likely to be even higher than calculated, due to the above-discussed underestimation of the farming system's actual labour demand. Changes in labour demand were not high enough to impact labour intensity. Small-scale mechanisation in S4 did not tackle the most labour demanding periods and it concentrates only on cotton (Figure 14). Thus, to increase its impact, small-scale mechanisation must be extended to the most labour demanding tasks like weeding and harvesting. This would require an increasing number of task specific tools which need to be provided and likely would increase the price of renting services. Nevertheless, expenses of farmers in S4 on renting services (~ 79 \$PPP year^{-1} averaged over all farms) were low compared to their total income (~

21,946 \$PPP year^{-1} averaged over all farms). Hence, an increase in renting prices may not be a limiting factor.

Despite resulting in a generally lower labour intensity in S1 (Figure 14), an increasing household size did not seem like a viable solution for two reasons. Firstly, it did not tackle the strong variability of the labour intensity throughout the year. Secondly, the household size interfered negatively with many other indicators.

4.4 Pathways towards Sustainable Intensification

Key drivers influencing the state of sustainability in many domains were identified. This included the application of mineral fertilizer as well as the migration and fertility rate which both affected the household size.

All key drivers pointed in a negative direction in the marginalisation scenario. As a result, the performance of all domains was reduced (Figure 15). Thus, S0 underlined the importance of family planning and the application of mineral fertilizer. A reduction of the cotton production especially affected farm income per capita, as cotton was the main source of income for farmers due to its high price compared to other crops and given subsidies (Table 6 and Table 14, Appendix).

Even though all domains were affected by the changes in S2, the impact SI was very small (Figure 12). Diversification with legumes for an increased N fixation and increased use of organic N sources raised the NUE (Lassaletta et al., 2013). This impact, however, was insufficient for most farms (Figure 13 c). For the majority of the farms the cattle herd component was too small to translate interventions in S2 into a substantial SI improvement. Furthermore, to implement the changes from S2, innovations of the agricultural system are needed. Farmers would need to have more power to influence the policies regarding agriculture and the creation of a more favourable milk sector (Falconnier et al., 2018).

Combining S2 with the family planning of S3, the performance of all household size related indicators changed. The impact of the dairy development scenarios became more visible. Due to family planning, the household size was reduced in S3 again to a level comparable to the baseline situation. Nevertheless, several farms still performed below the threshold for various indicators, e. g. farm income per capita, N balance and zinc self-sufficiency (Figure 12). Referring to the definition of sustainable agricultural systems of Florin et al. (2014) dairy development alone is not able to create a sustainable farming system.

Scenario S4 implemented policy interventions and intensification strategies to close the yield gap. That improved the indicators of many domains, to even a maximum in some cases.

The environmental domain was an exception to this increase in performance, as on average its indicators reacted negatively on increased fertilizer applications; resulting in a low NUE (Figure 13). Yet it is questionable if the impact on the environment is solely negative as predicted in this study. Many other potential effects of fertilizer on the environmental domain were not tested in this research (Musumba et al., 2017). The soils in southern Mali are lacking organic matter (Ripoche et al., 2015). Fertilizer applications in sub-Saharan African soils can slightly promote the increase of organic matter, and thereby soil structure and biological activities (Vanlauwe & Giller, 2006). On the contrary, greenhouse gas emissions may increase, especially under irresponsible application management (Snyder et al., 2009). Nevertheless, the justification or rejection of such strong intensification of the system must be done under the awareness of the regional conditions (Loos et al., 2014; Musumba et al., 2017). Mali is one of the poorest countries in the world with a strong need to substantially increase food production in the future (http://worldpopulationreview.com/, last accessed 24/08/2019). The rise in fertilizer application lifted the majority of the farms out of poverty and ensured their food-security (Figure 12).

High fertilizer applications increased the variability of the crop production (Table 9, Appendix). There is a strong relation between the success of a fertilizer application and rainfall pattern (Ripoche et al., 2015). As rainfall is highly variable in the Koutiala region, the success of fertilizer applications was equally variable. Increased nutrient supply affected many different indicators as well, including farm income per capita, different nutrient self-sufficiencies and various other indicators (Figure 20). For different nutrient self-sufficiencies, excluding zinc which performed below the threshold of self-sufficiency, yield variations are less of a problem. The variations did not impact these indicators enough to endanger the households (Figure 23, Appendix). However, for farm income per capita, on average HRE-LH farms came very close to the poverty line and HRE farms even went below it in individual seasons due to the high yield variations (Figure 10). Whether this drawback strongly affected the pathway's sustainability can be questioned, as averaged over the season most farms were lifted out of poverty (Figure 12). However, it was found that smallholder farmers would actually sacrifice income if they can reduce the risk of going below the poverty line (Komarek et al., 2012; Rötter & Van Keulen, 1997). This raises the questions to which extent the intensification strategies in A3 would be adopted by farmers in the first place. The implementation of small-scale irrigation systems could reduce the dependency on rainfall (Dillon, 2011), however, the overall sustainability of small-scale irrigation in the Koutiala region needs to be tested.

Next to the increased input, small-scale mechanisation was introduced to close the yield gap of cotton. Nonetheless, it failed to reduce the labour intensity of the farming system during labour-intensive periods since the assumed mechanisation did not tackle these periods (Figure 14). The issue of labour shortage demands to either expand the mechanisation or to find other strategies like organising harvest groups. In the latter case, farmers can, with combined force, complete tasks more efficiently for one farm after another (Personal communication, Arouna Dissa, 2019).

Variation in indicator performance increased in S4 between the farms because differences in farm characteristics such as equal yield were considered for all farms. Farms with more resources (e. g. total cropped area) benefited more from the increased yields. As a result, despite a closing yield gap, inequality was promoted between farms (Figure 12). This may result in a loss of competitiveness of some left behind farms in S4.

After all, the aim of closing the yield gap within 15 years is highly ambitious (Falconnier et al., 2018). It requires a high number of investments from the public sector to facilitate the input subsidies. Fertilizer subsidies in sub-Saharan Africa come along with high costs which are often not paid back due to low return rates of the investments (Jayne & Rashid, 2013). Usually, the simple provision of fertilizer subsidies is not enough to obtain the aspired outcome, as trainings on correct application methods are equally important (Morris et al., 2007). Additionally, the water deficiency in rainfed farming systems reduces the effectivity of the fertilizer applications, and eventually the subsidies, as it limits the return of investment (Jayne & Rashid, 2013). Moreover, due to their large governmental investments, input subsidy programmes are often prone to the crowding-out effect. The government is driving down the economy through investments and creates an increased dependency on public financial support which is not sustainable in the long term (Jayne & Rashid, 2013). Lastly, fertilizer subsidies are often accompanied by high opportunity costs as investments are lacking in other areas, e. g. market development, agricultural research or transportation infrastructure (Marenya et al., 2012).

The implementation of the cereal market development scenario, S5, lifted on average all farm types in all seasons out of poverty through a substantial increase in farm income (Figure 10). However, to implement inventory credits, prerequisites must be matched. Firstly, a sufficient legal framework must exist or be created to support them. Secondly, an abundance of market insights is required for the warehouse operators to act profitably. Both conditions are likely to be sufficiently met in Mali, as programmes with inventory credits were already conducted (Coulter et al., 1995). Lastly, inventory credits demand reliable warehouse operators with good business skills to increase the total reliability of the system (Coulter et al., 1995).

If the system is set up correctly, the total potential increase of revenues and possibilities to sell cereals at advantageous time points of the year with increased value is beneficial for farmers. In particular, the option of selling cereals shortly before harvest is valuable, as in this period producer prices as well as the demand for cash are the highest. During August farmers need to spend money on educational purposes and thus have high costs (Personal communication. Arouna Dissa, 2019). Additionally, increased producer prices reduce the risk of farming (Rötter & Van Keulen, 1997) which was shown in a slightly decreased variation of farm income per capita (Figure 20).

Moreover, farmers are currently highly dependent on a strong cotton sector (Djouara et al., 2006). An increased potential of also growing cereals as cash crops can reduce farmers' dependency on cotton (Table 19, Appendix). Furthermore, income diversification reduces the risk of farming activities (Barrett, Reardon, & Webb, 2001). Nevertheless, inventory credits are not a long-term solution. With an increasing number of joining farmers, the yearly variation of the prices is eliminated. Thus, the increased return as a result of storage is eliminated, too (Sanders & Shapiro, 2006).

5 Conclusion

The state of SI of the whole population of two villages in the Koutiala region (411 smallholder farms) was estimated for the current time point and the near-term future. Therefore, six different scenarios were developed based on incremental policy interventions and agricultural intensification strategies and combined with a SI assessment. A full assessment of the sustainability, however, was not possible due to a one-sided database and thus only four out of five domains of sustainability were captured, being the economic, productivity, environmental and human well-being but not the social domain.

Making use of scenario analyses and models to test the future state of SI of farming system was proven to be a useful tool, which full potential was not exhausted in this research yet. The assessment showed that few trade-offs between intensification and sustainability have to be feared when implementing the tested interventions. By assuming intensification to be the main objective for agricultural smallholder system in sub-Saharan Africa three potential pathways that enable sustainable intensification were identified.

Firstly, family planning interventions reducing the household size improved the sustainability in many domains with very few trade-offs. As stressed in the literature and confirmed in this research the rapid population growth puts high pressure on the systems in sub-Saharan Africa and needs to be tackled.

Secondly, increasing productivity through mineral fertilizer subsidies also enables enhanced SI in many domains. The given trade-offs, especially with the environmental domain should be considered under the aspect of understanding the biggest needs of the system. Input subsidies for fertilizer seem to be a promising pathway to enable SI especially, if they are introduced with sufficient training and accompanied by other agronomical means and especially when the fertilizer is applied in sound quantities.

Lastly, give the relevance of the cotton sector and its current shrinkage, cereal inventory credits were also identified as a viable pathway. Even though, they only affect the economic domain directly and are most promising if combined with an improved productivity they increase income diversification and thereby mitigate the risk of farming and dependency on cotton. However, it must be stated that with increased intensification inequalities between the single farms are growing and next to the productivity also the variation in productivity rises.

Conclusion

This research confirmed that there is no silver bullet that single-handedly enables SI. Only the combination of different interventions is able to improve the sustainability of the households. Moreover, it showed that there is a need for multi-dimensional assessments to capture the full impact of proposed interventions.

6 Suggestions for Further Research

It was out of the scope of this study to identify and involve many different stakeholders, namely the farmers, the downstream industry and policy makers. We believe that an inclusion of the stakeholders would have increased the value of the outcomes substantially. Thus, we suggest that in future studies the stakeholders should be much more involved, especially in the indicator identification process.

Furthermore, we suggest to increase relevance of indicator outcomes by introducing additional steps in future-oriented SI assessments:

(1) The implementation of a study-specific data gathering process to gain more diverse data would allow to test for more indicators and thereby more aspects of SI could have been analysed. (2) Weighting the indicators in collaboration with the stakeholders would better match the, multi attribute utility theory, which is well-established in multi-criteria decision-making (Velasquez & Hester, 2013). Thereby, the outcome of the research and proposed pathways to enable SI will be more targeted to the needs of the people. (3) The implementation of a sensitivity analysis and/or the reduction of some indicators to field level could provide a deeper understanding of the key drivers and allow to draw more targeted conclusions on how to reach sustainable intensification.

Reference List

Adegun, M. K., & Aye, P. A. (2013). Growth performance and economic analysis of West African Dwarf Rams fed Moringa oleifera and cotton seed cake as protein supplements to Panicum maximum Department of Animal Production and Health Sciences. *American Journal of Food and Nutrition*, *3*(2), 58–63. https://doi.org/10.5251/ajfn.2013.3.2.58.63

Agyemang, K. (1992). Breed evaluation and characterisation work in the Gambia. In *African animal genetic resources: their characterisation,conservation and utilisation*. ILCA. Retrieved from https://hdl.handle.net/10568/2753

Agyemang, K., Dwinger, R. H., Grieve, A. S., & Bah, M. L. (1991). Milk Production Characteristics and Productivity of N'Dama Cattle Kept Under Village Management in The Gambia. *Journal of Dairy Science*, *74*(5), 1599–1608. https://doi.org/10.3168/jds.S0022-0302(91)78322-7

Aparisi, A. M., Diallo, F., & Bailié, J. (2012). *Analyse des incitation et penalisations pour le coton au Mali*. Rome. Retrieved from http://www.fao.org/3/a-at572f.pdf

Ba, A., & Lesnoff, M. (2011). Demographic dynamics and off-take of cattle herds in southern Mali. *Tropical Animal Health and Production*, *43*, 1101–1109. https://doi.org/10.1007/s11250-011-9808-2f

Barrett, C. B., Reardon, T., & Webb, P. (2001). Nonfarm income diversification and household livelihood strategies in rural Africa : concepts , dynamics , and policy implications. *Food Policy*, *26*, 315–331. https://doi.org/10.1016/S0306-9192(01)00014-8

Baudron, F., Sims, B., Justice, S., Kahan, D. G., Rose, R., Mkomwa, S., Mkomwa, S.; Kaumbutho, P.; Sariah, J.; Nazare, R. (2015). Re-examining appropriate mechanization in Eastern and Southern Africa : two-wheel tractors , conservation agriculture , and private sector involvement. *Food Security*, *7*, 889–904. https://doi.org/10.1007/s12571-015-0476-3
Bazile, D. (2008). Étude originale, *17*, 79–85.

Bruun, S., Lund, T., Christensen, T. H., Magid, J., & Jensen, L. S. (2006). Application of processed organic municipal solid waste on agricultural land - a scenario analysis. *Environmental Modeling and Assessment*, *11*, 251–265. https://doi.org/10.1007/s10666-005-9028-0

Caliskan, S., Caliskan, M. E., Arslan, M., & Arioglu, H. (2008). Effects of sowing date and growth duration on growth and yield of groundnut in a Mediterranean-type environment in Turkey. *Field Crops Research*, *105*, 131–140. https://doi.org/10.1016/j.fcr.2007.08.007

Coulibaly, J. Y., Sanders, J. H., Preckel, P. V, & Baker, T. G. (2015). Will cotton make a comeback in Mali ? *Agricultural Economics*, *46*, 53–67. https://doi.org/10.1111/agec.12140

Dale, V. H., Efroymson, R. A., Kline, K. L., & Oak Ridge. (2015). Review A framework for selecting indicators of bioenergy sustainability. *Biofuels Bioprod. & Bioref.*, *9*, 435–446. https://doi.org/10.1002/bbb.1562

de la Croix, K., Ferry, L., & Landy, F. (2011). De nouvelles dynamiques de mécanisation agricole : commerce, usages et spatialisation au sein de la région de Ségou (Mali) / New dynamics of farming mechanization: trade, uses and spatialization in the Ségou region (Mail). *Annales de Géographie*, *2*(678), 174–192. https://doi.org/10.3917/ag.678.0174.

de Olde, E. M., Moller, H., Marchand, F., McDowell, R. W., MacLeod, C. J., Sautier, M., Halloy, S.; Barber, A.; Benge, J.; Bocksteller, C.; Bokkers, E. A. M.; de Boer, I. J. M.;

Legun, K. A.; Le Quelle, I.; Merfield, C.; Oudshoorn, F. W.; Reid, J.; Schader, C.; Szymanski, E., Sørensen, C. A. G.; Whitehead, J.; Manhire, J. (2017). When experts disagree: the need to rethink indicator selection for assessing sustainability of agriculture. *Environment, Development and Sustainability*, *19*(4), 1327–1342. https://doi.org/10.1007/s10668-016-9803-x

de Ridder, N., Sanogo, O. M., Ru, M. C., Keulen, H. Van, & Giller, K. E. (2015). Milk : the new white gold ? Milk production options for smallholder farmers in Southern Mali. *Animal*, 1–9. https://doi.org/10.1017/S1751731115000178

Defoer, T., de Groote, H., Kanté, S., & Budelman, A. (1998). Participatory action research and quantitative analysis for nutrient management in southern Mali : a fruitful marriage ? *Agriculture, Ecosystems and Environment*, *71*, 215–228. https://doi.org/10.1016/S0167-8809(98)00142-X

Delon, C. (2014). Nitrogen emission and deposition budget in West and Central Africa. *Environmental Research Letters*, *9*. https://doi.org/10.1088/1748-9326/9/12/125002

Dillon, A. (2011). The effect of irrigation on poverty reduction, asset accumulation, and informal insurance: Evidence from Northern Mali. *World Development*, *39*(12), 2165–2175. https://doi.org/10.1016/j.worlddev.2011.04.006

Djouara, H., Bélières, J., & Kébé, D. (2006). Étude originale. *Cahiers Agricultures*, *15*(1), 64–71. Retrieved from http://revues.cirad.fr/index.php/cahiers-agricultures/article/view/30561

Droppelmann, K. J., Snapp, S. S., & Waddington, S. R. (2017). Sustainable intensification options for smallholder maize-based farming systems in sub-Saharan Africa. *Food Security*, *9*(1), 133–150. https://doi.org/10.1007/s12571-016-0636-0

Duinker, P. N., & Greig, L. A. (2007). Scenario analysis in environmental impact assessment: Improving explorations of the future. *Environmental Impact Assessment Review*, *27*(3), 206–219. https://doi.org/10.1016/j.eiar.2006.11.001

Dupraz, C. L., & Postolle, A. (2013). Food sovereignty and agricultural trade policy commitments: How much leeway do West African nations have? *Food Policy*, *38*, 115–125. https://doi.org/10.1016/j.foodpol.2012.11.005

Falconnier, G. N. (2009). *Diagnostic agraire des villages de Wacoro , Nangola et Belekou dans le secteur CMDT de Dioïla : Quelles évolutions possibles pour les exploitations agricoles familiales du Sud du Mali face à la crise de la filière cotonnière ?* Montpellier Supagro.

Falconnier, G. N., Descheemaeker, K., Mourik, T. A. V., & Giller, K. E. (2016). Unravelling the causes of variability in crop yields and treatment responses for better tailoring of options for sustainable intensification in southern Mali. *Field Crops Research*, *187*, 113–126. https://doi.org/10.1016/j.fcr.2015.12.015

Falconnier, G. N., Descheemaeker, K., Traore, B., Bayoko, A., & Giller, K. E. (2018). Agricultural intensification and policy interventions: Exploring plausible futures for smallholder farmers in Southern Mali. *Land Use Policy*, *70*, 623–634. https://doi.org/10.1016/j.landusepol.2017.10.044

Falconnier, G. N., Descheemaeker, K., Van Mourik, T. A., Adam, M., Sogoba, B., & Giller, K. E. (2017). Co-learning cycles to support the design of innovative farm systems in southern Mali. *European Journal of Agronomy*, *89*, 61–74. https://doi.org/10.1016/j.eja.2017.06.008

Falconnier, G. N., Descheemaeker, K., Van Mourik, T. A., Sanogo, O. M., & Giller, K. E. (2015). Understanding farm trajectories and development pathways: Two decades of change in southern Mali. *Agricultural Systems*, *139*, 210–222. https://doi.org/10.1016/j.agsy.2015.07.005

Florin, M. J., van de Ven, G. W. J., & van Ittersum, M. K. (2014). What drives sustainable biofuels? A review of indicator assessments of biofuel production systems involving smallholder farmers. *Environmental Science and Policy*, *37*, 142–157. https://doi.org/10.1016/j.envsci.2013.09.012

Garnett, T., Appleby, M. C., Balmford, A., Bateman, I. J., Benton, T. G., Bloomer, P., Burlingame, B.; Dawkins, M.; Dolan, L.; Fraser, D.; Herrero, M.; Hoffmann, I.; Smith, P.; Thornton, P. K.; Toulmin, C.; Vermeulen, S. J.; Godfray, H. C. J. (2013). Sustainable intensification in agriculture: Premises and policies. *Science*, *341*(6141), 33–34. https://doi.org/10.1126/science.1234485

Garnett, T., & Godfray, C. (2012). *Sustainable intensification in agriculture. Navigating a course through competing food system priorities*. *Food Climate Research Network and the Oxford Martin Programme on the Future of Food*. https://doi.org/1835-0135

Godfray, H. C. J., Beddington, J. R.; Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J.; Robinson, S.; Thomas, S. M.; Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *Science (New York, N.Y.)*, *327*(5967), 812–818. https://doi.org/10.1126/science.1185383

Godfray, H. C. J., Garnett, T., & B, P. T. R. S. (2014). Food security and sustainable intensification Food security and sustainable intensification. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*. https://doi.org/10.1098/rstb.2012.0273 One

Green, R. E., Cornell, S. J., Scharlemann, J. P. W., & Balmford, A. (2005). Farming and the fate of wild nature. *Science (New York, N.Y.)*, *307*(5709), 550–555. https://doi.org/10.1126/science.1106049

Harris, D., & Orr, A. (2014). Is rainfed agriculture really a pathway from poverty? *Agricultural Systems*, *123*, 84–96. https://doi.org/10.1016/j.agsy.2013.09.005

Heuer, B., & Nadler, A. (2008). Physiological Parameters , Harvest Index and Yield of Deficient Irrigated Cotton Physiological Parameters , Harvest Index and Yield of Deficient Irrigated Cotton. *Journal of Crop Production*, *2*(2), 229–239. https://doi.org/10.1300/J144v02n02 09

Ittersum, M. K. Van, Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., & Hochman, Z. (2013). Field Crops Research Yield gap analysis with local to global relevance — A review. *Field Crops Research*, *143*, 4–17. https://doi.org/10.1016/j.fcr.2012.09.009

Jayne, T. S., Chamberlin, J., & Headey, D. D. (2014). Land pressures, the evolution of farming systems, and development strategies in Africa: A synthesis. *Food Policy*, *48*, 1–17. https://doi.org/10.1016/j.foodpol.2014.05.014

Jayne, T. S., & Rashid, S. (2013). Input subsidy programs in sub-Saharan Africa : a synthesis of recent evidence. *Agricultural Economics*, *44*, 547–562. https://doi.org/10.1111/agec.12073

Jolliffe, D., & Prydz, E. B. (2016). Estimating International Poverty Lines from Comparable National Thresholds. *The Journal of Economic Inequality*, *14*(2), 185–198. https://doi.org/10.1007/ s10888-016-9327-5

Jonckers, D., & Colleyn, J. (2018). La communauté familiale chez les Minyanka du Mali. *Journal de La Société Des Africanistes*, *1*, 43–52. https://doi.org/10.3406/jafr.1974.1748

Kaya, B., Hildebrand, P. E., & Nair, P. K. R. (2000). Modeling changes in farming systems with the adoption of improved fallows in southern Mali. *Agricultural Systems*, *66*, 51–68.

Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., Huth, N. I.; Hargreaves, J. N. G.; Meinke, H.; Hochman, Z.; McLean, G.; Verburg, K.; Snow, V.; Dimes, J.P.; Silburn, M.; Wang, E.; Brown, S.; Bristow, K. L.; Assesng, S.; Chapman, S.; McCown, R. L.; Freebairn, D. M.; Smith, C. J. (2003). An overview of APSIM , a model designed for farming systems simulation. *European Journal of Agronomy*, *18*, 267–288. https://doi.org/10.1016/S1161-0301(02)00108-9.

Kiker, G. A., Bridges, À. T. S., Varghese, À. A., & Thomas, P. (2005). Application of Multicriteria Decision Analysis in Environmental Decision Making. *Integrated Environmental Assessment and Management*, *1*(2), 95–108. https://doi.org/10.1897/IEAM_2004a-015.1

Komarek, A. M., Mcdonald, C. K., Bell, L. W., Whish, J. P. M., Robertson, M. J., Macleod, N. D., & Bellotti, W. D. (2012). Whole-farm effects of livestock intensification in smallholder systems in Gansu , China. *Agricultural Systems*, *109*, 16–24. https://doi.org/10.1016/j.agsy.2012.02.001

Lancker, E., Nijkamp, P., Lancker, E., & Nijkamp, P. (2012). Sustainable agriculture for Nepal. *Impact Assessment and Project Appraisal*, *18*(2), 111–124. https://doi.org/10.3152/147154600781767493

Lassaletta, L., Billen, G., Grizzetti, B., & Anglade, J. (2013). 50 year trends in nitrogen use efficiency of world cropping systems : the relationship between yield and nitrogen input to cropland. *Environmental Research Letters*, *9*. https://doi.org/10.1088/1748-9326/9/10/105011

Lönnerdal, B. (2000). Dietary Factors Influencing Zinc Absorption. *The Journal of Nutrition*, *130*(5), 1378–1383. https://doi.org/10.1093/jn/130.5.1378S

Loos, J., Abson, D. J., Chappell, M. J., Hanspach, J., Mikulcak, F., Tichit, M., & Fischer, J. (2014). Putting meaning back into “sustainable intensification.” *Frontiers in Ecology and the Environment*, *12*(6), 356–361. https://doi.org/10.1890/130157

Maclean, W. C., Harnly, J. M., Chen, J., Chevassus-Agnes, S., Gilani, G., Livesey, G., Mathioudakis, B.; Munoz De Chavez, M.; Devasconcellos, M. T.; Warwick, P. (2003). *Food energy – methods of analysis and conversion factors*.

Marenya, P., Nkonya, E., Xiong, W., Deustua, J., & Kato, E. (2012). Which policy would work better for improved soil fertility management in sub-Saharan Africa , fertilizer subsidies or carbon credits ? *Agricultural Systems*, *110*, 162–172. https://doi.org/10.1016/j.agsy.2012.04.004

Marinus, W., Ronner, E., Ven, G. W. J. Van De, Kanampiu, F., Adjei-nsiah, S., & Giller, Ken, E. (2018). The devil is in the detail! Sustainability assessment of African smallholder farming. In *Routledge Handbook of Sustainability Indicators* (pp. 427–450). https://doi.org/10.4324/9781315561103-28

Morris, M., Kelly, V. A., Kopicki, R. J., & Byerlee, D. (2007). *Fertilizer Use in African Agriculture: Lessons Learned and Good Practice Guidelines*. The World Bank. https://doi.org/10.1596/978-0-8213-6880-0

Muchow, R. C. (1989). Comparative Productivity of Maize , Sorghum and Pearl Millet in a Semi-Arid Tropical Environment II . Effect of Water Deficits. *Field Crops Research*, *20*, 207–219. https://doi.org/0378-4290/89/$03.05

Musumba, M., Grabowski, P., Palm, C., & Snapp, S. S. (2017). Guide for the Sustainable

Intensification Assessment Framework, 1–46. Retrieved from https://hdl.handle.net/10568/90523

Nijhof, K. (1987). *The Concentration of Macro-elements in Economic Products and Residues of (Sub)Tropical Field Crops*. *Centre for World Food Studies*.

Oenema, O., Brentrup, F., Lammel, J., Bascou, P., Billen, G., Dobermann, A., Erisman, J. W.; Garnett, T.; Hammel, M.; Haniotis, T.; Hillier, J.; Hoxha, A.; Jensen, L. S.; Oleszek, W.; Pallière, C.; Powlson, D.; Quemada, M.; Schulman, M.; Sutton, M. A.; Van Grinscen, H. J. M.; Winiwarter, W. (2016). Nitrogen Use Efficiency (NUE) - an indicator for the utilization of nitrogen in agriculture and food systems. *Encyclopedia of Applied Plant Sciences*, *1*, 278–284. https://doi.org/10.1016/B978-0-12-394807-6.00138-6

Ollenburger, M. H., Descheemaeker, K., Crane, T. A., Sanogo, O. M., & Giller, K. E. (2016). Waking the Sleeping Giant: Agricultural intensification, extensification or stagnation in Mali’s Guinea Savannah. *Agricultural Systems*, *148*, 58–70. https://doi.org/10.1016/j.agsy.2016.07.003

Pessarakli, M. (1999). *Handbook of Plant and Crop Stress*.

Phalan, B., Onial, M., Balmford, A., & Green, R. E. (2011). Reconciling Food Production and. *Science*, *333*, 1289–1291. https://doi.org/10.1126/science.1208742

Phoomthaisong, J., Toomsan, B., Limpinuntana, V., Cadisch, G., & Patanothai, A. (2003). Attributes affecting residual benefits of N 2 -fixing mungbean and groundnut cultivars. *Biology and Fertility of Soils*, *49*(1), 16–24. https://doi.org/10.1007/s00374-003-0676-4

Porter, M. E., & Millar, V. E. (1985). How information gives you competitive advantage. *Harvard Buiness Review*, *63*(4), 149. https://doi.org/10.1038/bdj.2007.481

Pretty, J. (2008). Agricultural sustainability: Concepts, principles and evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1491), 447–465. https://doi.org/10.1098/rstb.2007.2163

Pretty, J. N. (1997). The sustainable intensification of agriculture. *Natural Resources Forum*, *21*(4), 247–256. https://doi.org/10.1111/j.1477-8947.1997.tb00699.x

Pretty, J., Toulmin, C., & Williams, S. (2011). Sustainable intensification in African agriculture. *International Journal of Agricultural Sustainability*, *9*(1), 5–24. https://doi.org/10.3763/ijas.2010.0583

Ripoche, A., Crétenet, M., Corbeels, M., Affholder, F., Naudin, K., & Sissoko, F. (2015). Field Crops Research Cotton as an entry point for soil fertility maintenance and food crop productivity in savannah agroecosystems – Evidence from a long-term experiment in southern Mali. *Field Crops Research*, *177*, 37–48. https://doi.org/10.1016/j.fcr.2015.02.013

Rötter, R., & Van Keulen, H. (1997). Variations in yield response to fertilizer application in the tropics: II. Risks and opportunities for smallholders cultivating maize on Kenya’s arable land. *Agricultural Systems*, *53*(1), 69–95. https://doi.org/10.1016/S0308-521X(96)00037-6

Roxburgh, C. W., & Rodriguez, D. (2016). Ex-ante analysis of opportunities for the sustainable intensification of maize production in Mozambique. *Agricultural Systems*, *142*, 9–22. https://doi.org/10.1016/j.agsy.2015.10.010

Rufino, M. C., Rowe, E. C., Delve, R. J., & Giller, K. E. (2006). Nitrogen cycling efficiencies through resource-poor African crop – livestock systems. *Agriculture, Ecosystems and Environment*, *112*, 261–282. https://doi.org/10.1016/j.agee.2005.08.028

Sadok, W., Angevin, F., Jacques-éric, B., Christian, B., Bruno, C., Laurence, G., Raymond, R.; Thierry, D. (2008). Review article Ex ante assessment of the sustainability of alternative cropping systems : implications for using multi-criteria decision-aid methods . A review. *Agronomy for Sustainable Development, 28,* 163–174. https://doi.org/10.1051/agro:2007043

Sadras, V. (2002). Interaction between rainfall and nitrogen fertilisation of wheat in environments prone to terminal drought : economic and environmental risk analysis. *Field Crops Research, 77,* 201–215. https://doi.org/10.1016/S0378-4290(02)00083-7

Sanders, J. H., & Shapiro, B. (2006). Accelerate Technological Change in the Semiarid Zones : A Focus on Sub-Saharan Africa. *Agronomy*, *23*(2), 879–900.

Sanginga, N., Lyasse, O., & Singh, B. B. (2000). Phosphorus use efficiency and nitrogen balance of cowpea breeding lines in a low P soil of the derived savanna zone in West Africa. *Plant and Soil, 220,* 119–128. https://doi.org/10.1023/A:1004785720047

Schut, M., Asten, P. Van, Okafor, C., Hicintuka, C., Mapatano, S., Léon, N., Kagabo, D.; Muchunguzi, P.; Njukwe, E.; Dontsop-nguezet, P. M.; Sartas, M.; Vanlauwe, B. (2016). Sustainable intensi fi cation of agricultural systems in the Central African Highlands : The need for institutional innovation. *Agricultural Systems, 145,* 165–176. https://doi.org/10.1016/j.agsy.2016.03.005

Silvie, P. J., Renou, A., Vodounnon, S., Bonni, G., Adegnika, M. O., Héma, O., Prudent, P.; Sorèze, J.; Ochou, G. O.; Togola, M.; Badiane, D.; Ndour, A.; Akantetou, P. K.; Ayeva, B.; Brévault, T. (2013). Threshold-based interventions for cotton pest control in West Africa: What's up 10 years later? *Crop Protection, 43,* 157–165. https://doi.org/10.1016/j.cropro.2012.09.006

Smith, A., Snapp, S., Chikowo, R., Thorne, P., Bekunda, M., & Glover, J. (2017). Measuring sustainable intensification in smallholder agroecosystems: A review. *Global Food Security, 12,* 127–138. https://doi.org/10.1016/j.gfs.2016.11.002

Snapp, S. S., Grabowski, P., Chikowo, R., Smith, A., Anders, E., Sirrine, D., Chimonyo, V.; Bekunda, M. (2018). Maize yield and profitability tradeoffs with social, human and environmental performance: Is sustainable intensification feasible? *Agricultural Systems, 162,* 77–88. https://doi.org/10.1016/j.agsy.2018.01.012

Snyder, C. S., Bruulsema, T. W., Jensen, T. L., & Fixen, P. E. (2009). Review of greenhouse gas emissions from crop production systems and fertilizer management effects. *Agriculture, Ecosystems and Environment, 133,* 247–266. https://doi.org/10.1016/j.agee.2009.04.021

Stoorvogel, J. J., Antle, J. M., Crissman, C. C., & Bowen, W. (2004). The tradeoff analysis model : integrated bio-physical and economic modeling of agricultural production systems. *Agricultural Systems, 80,* 43–66. https://doi.org/10.1016/j.agsy.2003.06.002

Struik, P. C., & Kuyper, T. W. (2017). Sustainable intensification in agriculture: the richer shade of green. A review. *Agronomy for Sustainable Development*, *37*(5). https://doi.org/10.1007/s13593-017-0445-7

Swart, R. J., Raskin, P., & Robinson, J. (2004). The problem of the future: Sustainability science and scenario analysis. *Global Environmental Change*, *14*(2), 137–146. https://doi.org/10.1016/j.gloenvcha.2003.10.002

Tilman, D., Balzer, C., Hill, J., & Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, *108*(50), 20260–20264. https://doi.org/10.1073/pnas.1116437108

Tittonell, P., Vanlauwe, B., de Ridder, N., & Giller, K. E. (2007). Heterogeneity of crop productivity and resource use efficiency within smallholder Kenyan farms : Soil fertility gradients or management intensity gradients ? *Agricultural Systems*, *94*, 376–390. https://doi.org/10.1016/j.agsy.2006.10.012

Tobergte, D. R., & Curtis, S. (2013). World Population Prospects The 2015 Revision Volume I: Comprehensive Tables. *Journal of Chemical Information and Modeling*, *53*(9), 1689–1699. https://doi.org/10.1017/CBO9781107415324.004

Traore, B. B., van Wijk, M. T., Descheemaeker, K., Corbeels, M., Rufino, M. C., & Giller, Ken, E. (2015). Climate Variability and Change in Southern Mali: Learning from Farmer Perceptions and On-farm Trials. *Experimental Agriculture*, *51*(4), 615–634. https://doi.org/10.1017/S0014479714000507

Traore, B., Corbeels, M., Wijk, M. T. Van, Rufino, M. C., & Giller, K. E. (2013). Effects of climate variability and climate change on crop production in southern Mali. *European Journal of Agronomy*, *49*, 115–125. https://doi.org/10.1016/j.eja.2013.04.004

van der Pol, F., & Traore, B. (1993). Soil nutrient depletion by agricultural production in Southern Mali. *Fertilizer Research*, *36*(1), 79–90. https://doi.org/10.1007/BF00749951

van Ittersum, M. K., Ewert, F., Heckelei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I.; Brouwer, Fl.; Donatelli, M.; Flichman, G.; Olsson, L.; Rizzoli, A. E.; van der Wal, T.; Wien, J. E.; Wolf, J. (2008). Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS). *Agricultural Systems*, *96*(1–3), 150–165. https://doi.org/10.1016/j.agsy.2007.07.009

Vanlauwe, B., & Giller, K. E. (2006). Popular myths around soil fertility management in sub-Saharan Africa. *Agriculture, Ecosystems and Environment*, *116*, 34–46. https://doi.org/10.1016/j.agee.2006.03.016

Velasquez, M., & Hester, P. T. (2013). An analysis of multi-criteria decision making methods. *International Journal of Operations Research*, *10*(2), 56–66. https://doi.org/10.1007/978-3-319-12586-2

Wani, S. P., Albrizio, R., & Vajja, N. R. (2012). Sorghum. In *Crop yield response to water* (p. 10).

Wezel, A., Soboksa, G., McClelland, S., Delespesse, F., & Boissau, A. (2015). The blurred boundaries of ecological, sustainable, and agroecological intensification: a review. *Agronomy for Sustainable Development*, *35*(4), 1283–1295. https://doi.org/10.1007/s13593-015-0333-y

Worku, M., & Zelleke, H. (2007). East African Journal of Sciences (2007) Advances in Improving Harvest Index and Grain Yield of Maize in Ethiopia. *East Aftican Journal of Science*, *1*(2), 112–119. https://doi.org/10.4314/eajsci.v1i2.40349

Appendix

Table 8: Listing of all variables, assigned to their indicator and formula, their descriptions# and unit. The variables are nor repeatedly listed if they are already explained in another indicator or formula.

Indicator	Formula	Variable	Description	Unit
Self-produced biological inputs	*1*	*MANNav*	Manure that can be spread on the cropland	kg N year^{-1}
		pm	Manure excreted	kg DM cow^{-1} day^{-1}
		Nmanure	N in manure	%
		n_{cattle}	Number of cattle per farm	cattle
		cm	Collected manure of total manure produced	%
		sl	N lost during pit storage	%
Total production	*2*	FOD_i	Fodder produced from crop	kg ha^{-1}
		HI_i	Harvest index of crop i	%
		$Yield_i$	Yield of crop i	tons ha^{-1}
	3	FOD_{tot}	Total produced fodder	kg year^{-1}
		A_i	Area under crop i	ha
	4	$GRAIN_{tot\ i}$	Total produced grain of crop i	kg year^{-1}
	5	$MILKt_{tot}$	Total Milk produced	kg year^{-1}
		lc	Proportion of herd lactating	%
		Myield	Milk yield	kg cow^{-1} year^{-1}
Income per capita	*6*	*Inc per capita*	Income per capita generated after subtracting all expenses	[$PPP
		r	Remittances from migrated household members	$PPP
		se	Income generated of the household members through self-employment	$PPP
	7	$cost_i$	Costs for crop i per hectare	$PPP ha^{-1}
		SD_i	Sowing density of crop i	kg ha^{-1}
		$seed_{costs\ i}$	Costs for seeds of crop i	$PPP kg^{-1}
		$Finput_i$	Costs for fertilizer for crop i	$PPP
	8	*Crop inc*	Income of crop production after subtracting crop specific expenses	$PPP
		$price_i$	Crop specific price farmers get for their products	$PPP kg^{-1}

		$cost_i$	Crop specific costs per hectare	\$PPP ha^{-1}
	9	*Lv inc*	Income of livestock husbandry after subtracting expenses	\$PPP
		$price_M$	Milk price	\$PPP kg^{-1}
		$costs_{fd}$	Costs for additionally bought fodder per livestock head	\$PPP live-stock head^{-1} year^{-1}
		$costs_v$	Costs for vaccination per livestock head	\$PPP live-stock head^{-1} year^{-1}
		offtakerate	Rate in which farmers sell animals to keep the herd size constant	%
		$price_{anim}$	Price for selling an animal	[\$PPP animal^{-1}
	10	*Depr costs*	Total costs due to depreciation of machines and animals	\$PPP
		$n_{machines}$	Amount of owned specific drawn equipment (plough, weeder, sowing machine, cart) and oxen	machines
		bprice	Buying price	\$PPP
		$ls_{machines}$	Live span of the specific drawn equipment	month
Labour productivity	11	*Labour prod*	Total income generated per workload	\$PPP (men-day)$^{-1}$
		tot. inc	Total income generated after subtraction of expenses	\$PPP year^{-1}
		$demand_i$	Labour demand for crop i per year	man-days
		animal	Labour demand for the live-stock husbandry per year	man-days
Nitrogen balance	12	*N balance*	Balance of N in- and outputs	kg ha^{-1} year^{-1}
		$Fert_i$	Crop specific amount of fertilizer applied	kg ha^{-1} year^{-1}
		depo	N Deposition	kg ha^{-1} year^{-1}
		A_t	Total cropping area	ha
	13	*Tot. Ncsc*	Total amount of N in all bought cotton seed cakes per year	kg year^{-1}
		$n_{sfcattle}$	Amount of cattle that are stall-fed during per year	stall-fed cattle
		tsf	Amount of days cattle is stall-fed per year	days year^{-1}
		Fcsc	Number of required cotton seed cake for a stall-fed cow	kg day^{-1}

		Ncsc	N content of cotton seed cakes	%
	14	*Npast*	Total N that is introduced into the farm system through grazing on pasture	kg
		tgr	Proportion of the year the cows feed on pasture	day day^{-1}
	15	*.Nfix*	Total N fixed by legumes	kg
		$Nfix_{leg}$	N fixed of total N of each specific legumes	%
		$Cbiomass_{leg}$	Total biomass produced for each legume	kg year^{-1}
	16	*Nmilk*	Total N in produced milk	kg year^{-1}
		milk prot	Protein content in milk	%
		Nprot	Average N content of proteins	%
	17	*Ngrain*	Total N harvested grain of all crops	kg
		$Ngrain_i$	N in crop specific grain	%
	18	*MANNloss*	N lost due to losses of produced manure during storage or missed uptake	kg year^{-1}
	19	*Tot.* $Nstraw_C$ *loss*	N losses in burned cotton residues	kg year^{-1}
		A_C	Area cultivated with cotton	ha
		HI_C	Harvest index of cotton	%
		$Yield_C$	Cotton yield	kg ha^{-1}
		$Nstraw_C$	N content of cotton straw	%
Nutrient use efficiency	*20*	NUE	Efficiency of nutrient use	kg kg^{-1}
Nutritional self-sufficiency	*21*	*Tot. Nutr* $demand_{nd}$	Daily household demand of each observed nutrient	kg day^{-1}
		Nutr $demand_{nd}$	Daily individual demand of each observed nutrient	g person^{-1} day^{-1}
	22	*Tot. Nutr* $prod_{nd}$	Total amount produced of each observed nutrient averaged of the year per day	kg day^{-1}
		grain $nutr_i$	Nutrient content of the specific crop grain	g kg-1
		$nutr_{milk}$	Nutrient content of milk	g kg^{-1}
		$nutr_{meat}$	Nutrient content of meat	g kg^{-1}
		c.weight	Average body weight of N’dama cow	kg

		dressing	Proportion of cattle body weight that is turned into beef	%
	23	$nutr\ selfsuff_n$	Daily self-sufficiency of the farm for each individual observed nutrient	g g^{-1}
Labour intensity	*24*	$Labour\ int_f$	Labour intensity per day of individual fortnights f	worker (demanded human labour)$^{-1}$
		$demand_{if}$	Labour demand for crop i and fortnight f	man-days ha^{-1}
		$animal_f$	Labour demand per fortnight for livestock holding	man-days
		$days_f$	Number of days per fortnight	days
		n_{worker}	Number of worker per farm	worker

Table 9: Yield in the baseline situation, averaged over the seasons and 85% of water limited yield. Abbreviations: HRE-LH: High Resource Endowed farms with Large Herds, HRE: High Resource Endowed farms, MRE: Medium Resource Endowed farms, LRE: Low Resource Endowed farms (standard deviation in brackets). The letters “a” to “i” after the values refer to the references listed below the table.

Crop	**Average yield (kg ha^{-1}) in the baseline situation**				**85 % of potential water limited yield (kg ha^{-1})**			
	HRE-LH	**HRE**	**MRE**	**LRE**	**HRE-LH**	**HRE**	**MRE**	**LRE**
Cotton	1050 [a]	940 [a]	910 [a]	750 [a]	2220 (± 599) [d]	2220 (± 599) [d]	2220 (± 599) [d]	2220 (± 599) [d]
Maize	3480 (± 190) [e]	3480 (± 190) [e]	3480 (± 190) [e]	2700 (± 125) [f]	4630 (± 680) [h]	4630 (± 680) [h]	4630 (± 680) [h]	4630 (± 680) [h]
Maize in maize/ cowpea intercropping	3654 (± 190) [g]	3654 (± 190) [g]	3654 (± 190) [g]	2835 (± 125) [g]	4860 (± 680) [g]	4860 (± 680) [g]	4860 (± 680) [g]	4860 (± 680) [g]
Sorghum	1030 [b]	1030 [b]	1030 [b]	1030 [b]	2060 (± 320) [i]	2060 (± 320) [i]	2060 (± 320) [i]	2060 (± 320) [i]
Millet	850 [c]	850 [c]	850 [c]	850 [c]	1730 (± 510) [i]	1730 (± 510) [i]	1730 (± 510) [i]	1730 (± 510) [i]
Groundnut	530 [b]	530 [b]	530 [b]	530 [b]	-	-	-	-
Cowpea grain	150 [b]	150 [b]	150 [b]	150 [b]	-	-	-	-
Cowpea fodder in maize/cowpea intercropping	1380 [b]	1380 [b]	1380 [b]	1380 [b]	-	-	-	-

a: Falconnier et al. (2015)

b: Falconnier et al. (2016)

c: Traore et al. (2015)

d: Ripoche et al. (2015)

e: APSIM simulation with a fertilizer application of 60 kg N ha^{-1} (Falconnier et al., 2018)

f: APSIM simulation with a fertilizer application of 40 kg N ha^{-1} (Falconnier et al., 2018)

g: APSIM simulated maize yield multiplied by 1.08, i. e. the maize partial land equivalent ratio for intercropping when grown after cotton (Falconnier et al., 2016)

h: APSIM simulation with a fertilizer application of 110 kg N ha^{-1} (Falconnier et al., 2018)

i: APSIM simulation with a fertilizer application of 150 kg N ha^{-1} (Falconnier et al., 2018)

Table 10: Criteria for farm classification of farms from the Koutiala region (Falconnier et al., 2015).

Farm type	**Average worker (n)**	**Average cropping area (ha)**	**Average draught tools (n)**	**Average herd size (TLU)**
HRE-LH	28	17	4	46
HRE	18	12	4	8
MRE	7	8	5	6
LRE	5	3	1	2

Table 11: Five potential future policy interventions of the Malian government and their effect on input and output prices, socio-economic development and on use of practices to closing the yield gap of smallholder farm households in the Koutiala region (based on Falconnier et al. (2018)).

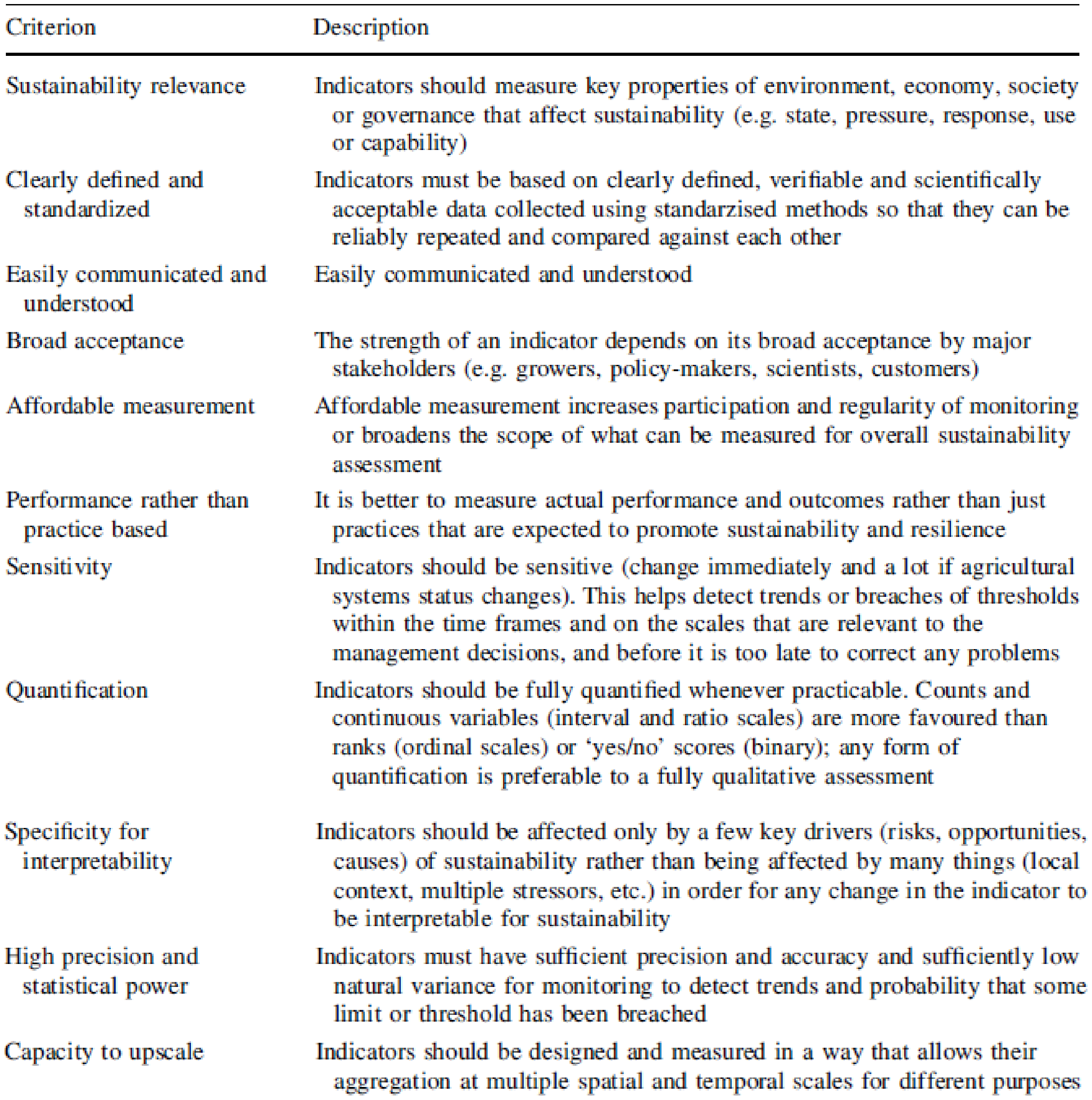
Intervention	**Effects on input and output prices**	**Effect on socio-economic development**	**Effect on use of practices to closing the yield gap**
P0	Steady decline in cotton prices + Steady increase in mineral fertilizer prices because of dropped fertilizer subsidies	None	None
P1	Cotton price maintained on the level of 2011 - 2015 + Mineral fertilizer prices maintained on the level of 2011-2015 high + Increased demand and thus increased prices for locally produced milk because of rising prices for milk powder	None	None
P2	Like P1 + Cotton seed cake price at low price level of 2003	None	None
P3	Like P2	35 % decrease in fertility rate to 2.2 % because of an increase of the rate of contraceptive use from 9.9 % to at least 15 % from 2015 – 2018 + 2.8 % of rural to urban migration through training programmes for young entrepreneurs in urban sectors	None

P4	Like P3	Like P3	Increased usage of integrated pest management based on policy intervention + Small-scale mechanisation improved land preparation, sowing and weeding of cotton + Subsidies for fertilizer use for sorghum and millet
P5	Like P4 + Cereal prices increased to the maximum in potential price in the year	Like P4	Like P4

Table 12: Four potential future agricultural intensification strategies of smallholder farm households in the Koutiala region (based on Falconnier et al. (2018)).

Agricultural intensification	Impact on farm management
A0	Reduction of cotton area specific for all farm types (HRE-LH 30 %; HRE 66 %; MRE 75 % and LRE 66 %). Reduction of mineral fertilizer use for all farm types to LRE level
A1	No effect
A2	Adopting farming practices from prior research (Falconnier et al., 2016, 2017): Maize/cowpea intercropping (diversification with legumes) + Stall feeding of lactating cows (intensification) + Cowpea as fodder
A3	Increased mineral fertilizer usage for cereals up to 85 % of the potential yield + Small-scale mechanisation for cotton production + Application of integrated pest management for cotton

Table 13: Possible criteria for selecting individual agricultural sustainability indicators, after Moller and MacLeod (2013) (source: (de Olde et al., 2017)).



Criterion	Description
Sustainability relevance	Indicators should measure key properties of environment, economy, society or governance that affect sustainability (e.g. state, pressure, response, use or capability)
Clearly defined and standardized	Indicators must be based on clearly defined, verifiable and scientifically acceptable data collected using standarzised methods so that they can be reliably repeated and compared against each other
Easily communicated and understood	Easily communicated and understood
Broad acceptance	The strength of an indicator depends on its broad acceptance by major stakeholders (e.g. growers, policy-makers, scientists, customers)
Affordable measurement	Affordable measurement increases participation and regularity of monitoring or broadens the scope of what can be measured for overall sustainability assessment
Performance rather than practice based	It is better to measure actual performance and outcomes rather than just practices that are expected to promote sustainability and resilience
Sensitivity	Indicators should be sensitive (change immediately and a lot if agricultural systems status changes). This helps detect trends or breaches of thresholds within the time frames and on the scales that are relevant to the management decisions, and before it is too late to correct any problems
Quantification	Indicators should be fully quantified whenever practicable. Counts and continuous variables (interval and ratio scales) are more favoured than ranks (ordinal scales) or 'yes/no' scores (binary); any form of quantification is preferable to a fully qualitative assessment
Specificity for interpretability	Indicators should be affected only by a few key drivers (risks, opportunities, causes) of sustainability rather than being affected by many things (local context, multiple stressors, etc.) in order for any change in the indicator to be interpretable for sustainability
High precision and statistical power	Indicators must have sufficient precision and accuracy and sufficiently low natural variance for monitoring to detect trends and probability that some limit or threshold has been breached
Capacity to upscale	Indicators should be designed and measured in a way that allows their aggregation at multiple spatial and temporal scales for different purposes

Table 14: Product prices farmers obtain at the local markets in $PPP in the different scenarios (crops per kg); abbreviations: fodder variety (fv), grain variety (gv) (internal project survey).

Product	**Baseline, S1, S2, S3, S4**	**S0**	**S5**
Cotton	1.16	0.85	1.16
Maize	0.43	0.43	0.55
Millet	0.62	0.62	0.99
Sorghum	0.51	0.51	0.74
Groundnut	0.92	0.97	0.97
Cowpea fv	1.5	1.5	1.5
Cowpea gv	1.5	1.5	1.5
Maize cowpea intercropping	0.48	0.48	0.48
Cattle	602	602	602

Table 15: Sowing quantities of the different crop seeds per hectare; abbreviations: fodder variety (fv), grain variety (gv) (Personal communication, Amadou Traore, 2019).

Crop	**Sowing quantity (kg/ha)**
Cotton	30
Maize	25
Millet	8
Sorghum	8
Cowpea gv	20
Cowpea fv	25
Groundnut	80

Table 16: Man-days required for the individual observed crops per year in the baseline situation and each scenario (Internal project survey).

Scenario	**Cotton**	**Maize**	**Millet**	**Sorghum**	**Groundnut**	**Cowpea**
Baseline - S3	78	42	33	42	16	54
S4 – S5	72	44	33	44	16	54

Table 17: Harvest index and the nitrogen content of both grain and straw of the observed crops in percent. The letters "a" to "g" after the values refer to the references listed below the table.

Crop	**Harvest index [%]**	**N in grain [%]**	**N in straw [%]**
Cotton	35.21 [a]	1.90 [g]	1.50 [g]
Maize	41.40 [b]	1.55 [g]	0.90 [g]
Sorghum	35.00 [c]	2.10 [g]	0.75 [g]
Millet	25.00 [d]	2.10 [g]	0.70 [g]
Groundnut	40.00 [e]	4.50 [g]	1.40 [g]
Cowpea	30.00 [f]	3.80 [g]	1.73 [g]

a: Heuer and Nadler (2008)

b: Worku and Zelleke (2007)

c: Steduto et al. (2012)

d: Muchow (1989)

e: Caliskan et al. (2008)

f: Falconnier et al. (2016)

g: Nijhof (1987)

Table 18: Overview of the used average, graphical representation method and indicators/domains to answer the individual research questions.

Research question	Used average	Graphical representation	Considered indicators/ domains
1. How do the farms perform in the different domains of sustainable intensification based on the chosen indicators at both time points?	Over all seasons	Boxplot	Farm income per capita, labour productivity, N balance, NUE, calorie self-sufficiency; protein self-sufficiency, iron self-sufficiency, zinc, self-sufficiency and labour intensity
	Over the farm types	Boxplot	Farm income per capita
	Over all seasons and farm types	Boxplot	N balance
	Over all seasons and farms	Radar chart	All domains
2. What are the key external and internal drivers that have the strongest impact on the sustainable intensification of the farms?	Over all seasons and farms	Scatterplot	Farm income per capita, labour productivity, N balance, NUE, calorie self-sufficiency; protein self-sufficiency, iron self-sufficiency, zinc, self-sufficiency and labour intensity
3. What are the trade-offs or interactions between the indicators and between the domains?	Over all seasons and farms	Radar chart	All indicators
	Over all seasons and farm types	Scatterplot	All indicators and all domains

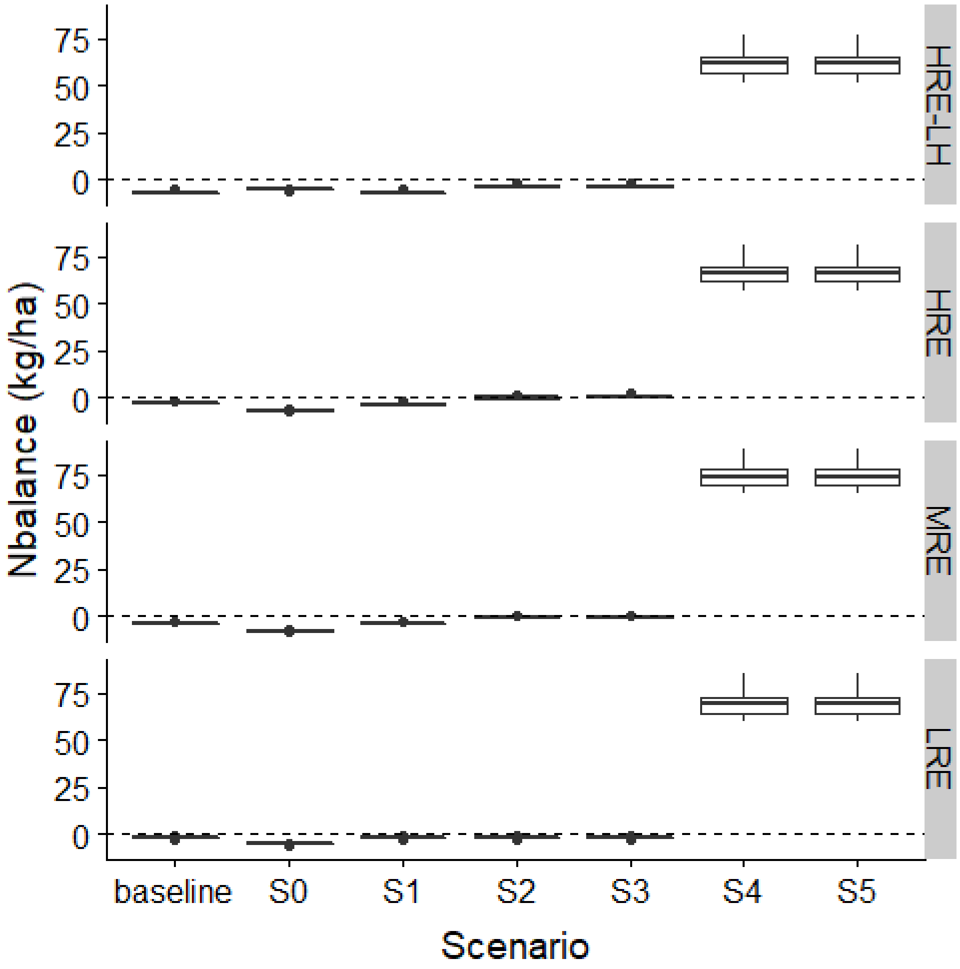


Figure 22: Boxplots showing the N balance in all seasons presented for farm types as well as for the baseline situation and for each scenario (S0 – S5). n(farms) = 411; n(seasons) = 26.

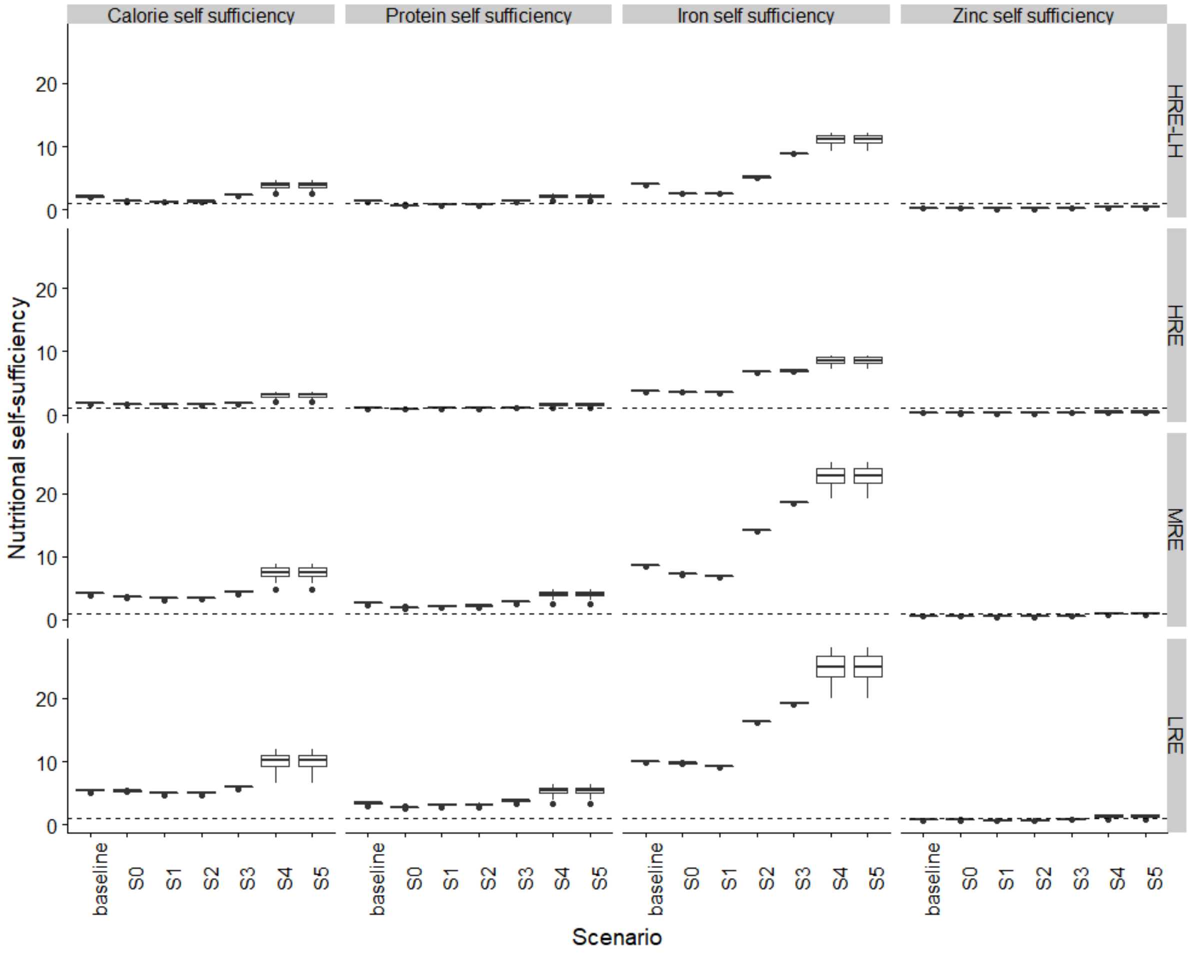


Figure 23: Boxplots showing the nutrient self-sufficiency of calories, protein, iron and zinc in all seasons presented for farm types as well as for the baseline situation and for each scenario (S0 – S5). n(farms) = 411; n(seasons) = 26.

Table 19: Individual share of each potential determinant of the sum of all potential determinants of farm income per capita for the baseline situation and each scenario in percent. The potential determinants were averaged over all farms. n(farms)=411.

Potential determinant	**Baseline (%)**	**S0 (%)**	**S1 (%)**	**S2 (%)**	**S3 (%)**	**S4 (%)**	**S5 (%)**
Cotton	21	4	20	15	16	31	25
Maize	16	14	1&	14	14	14	15
Sorghum	14	24	1$	10	11	9	13
Millet	15	17	15	11	12	10	17
Groundnut	11	13	11	8	9	7	5
Cowpea	≈0	≈0	≈0	≈0	≈0	≈0	≈0
Livestock	7	8	7	22	23	17	14
Fodder surplus	≈0	≈0	≈0	6	6	4	4
Remittances	5	6.4	6	4	4	3	2
Self-employment	9	14.2	12	9	7	5	34
Summ	100	100	100	100	100	100	100

Table 20 Individual share of each potential determinant of the sum of all potential determinants of the N balance for the baseline situation and each scenario in percent. The potential determinants were averaged over all farms. n(farms)=411.

Potential determinant	**Baseline (%)**	**S0 (%)**	**S1 (%)**	**S2 (%)**	**S3 (%)**	**S4 (%)**	**S5 (%)**
Cotton seed cake	≈0	≈0	≈0	5	5	1	1
Pasture	2	2	2	2	2	≈0	≈0
N fixation	7	10	7	13	13	3	3
Deposition	30	41	30	26	26	6	6
Fertilizer	62	47	62	55	55	89	89
Summ	100	100	100	100	100	100	100

Table 21: Individual share of each potential determinant of the sum of all potential determinants of calorie self-sufficiency (representing all nutrients) for the baseline situation and each scenario in percent. The potential determinants were averaged over all farms. n(farms)=411.

Potential determinant	**Baseline (%)**	**S0 (%)**	**S1 (%)**	**S2 (%)**	**S3 (%)**	**S4 (%)**	**S5 (%)**
Cotton	18	7	18	17	17	23	23
Maize	37	31	36	38	38	29	29
Sorghum	22	36	22	21	21	24	24
Millet	19	21	19	19	19	22	22
Groundnut	3	4	3	3	3	2	2
Cowpea	≈0	≈0	≈0	≈0	≈0	≈0	≈0
Milk	1	1	1	3	3	1	1
Summ	100	100	100	100	100	100	100

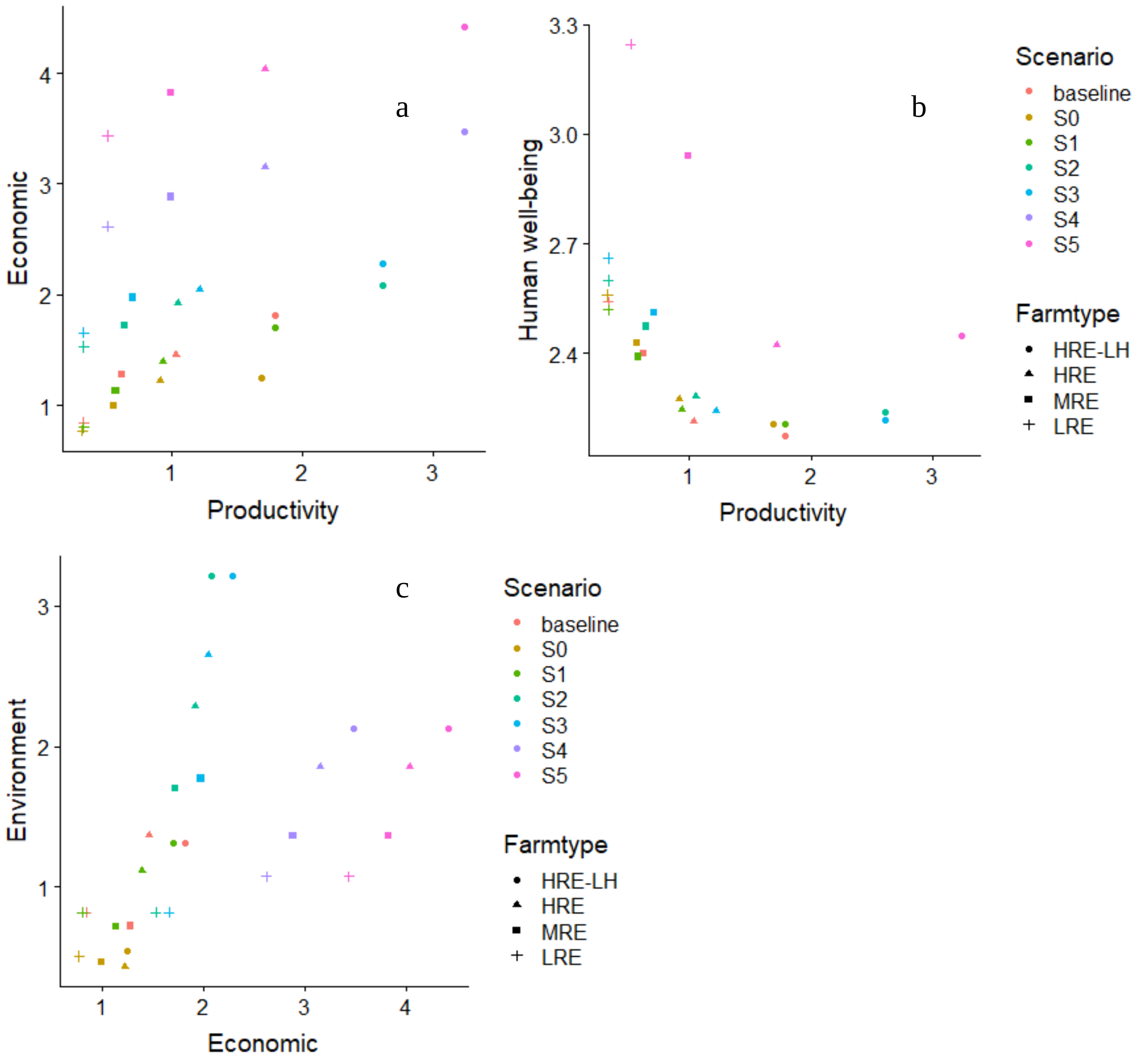


Figure 24: Scatterplots of a: the scaled economic domain against the scaled productivity domain, b: the scaled human well-being domain against the scaled productivity domain and c: the scaled environmental domain against the scaled economic domain in the baseline situation and the different scenarios (The data of the baseline and S1, S2 and S3 and S4 and S5 can overlap). The indicators are averaged over the different farm types. n(farms)=411.