Using yield gap analysis to give sustainable intensification local meaning

João Vasco Silva

João Vasco Silva
Using yield gap analysis to give sustainable intensification local meaning

Using yield gap analysis to give sustainable intensification local meaning

João Vasco Silva

Thesis committee

Promotor
Prof. Dr Martin K. van Ittersum
Personal chair at Plant Production Systems Group
Wageningen University & Research

Co-promotors
Prof. Dr Ken E. Giller
Professor of Plant Production Systems Group
Wageningen University & Research

Dr Pytrik Reidsma
Assistant professor, Plant Production Systems Group
Wageningen University & Research

Other members
Prof. Dr Paul C. Struik, Wageningen University & Research
Dr Patricio Grassini, University of Nebraska-Lincoln, USA
Dr Elke E. Stehfest, PBL Netherlands Environmental Assessment Agency, The Hague
Dr Argyris Kanellopoulos, Wageningen University & Research

This research was conducted under the auspices of the C.T. de Wit
Graduate School of Production Ecology & Resource Conservation

Using yield gap analysis to give sustainable intensification local meaning

João Vasco Silva

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Tuesday 12 December 2017
at 4 p.m. in the Aula.

João Vasco Silva
Using yield gap analysis to give sustainable intensification local meaning
376 pages.

PhD thesis, Wageningen University, Wageningen, NL (2017)
With references, with summaries in English, Dutch and Portuguese

ISBN 978-94-6343-714-1
DOI 10.18174/425752

Para os meus pais,
para a Susanne,
para a Marta.

Abstract

Yield gap analysis is useful to understand the relative contribution of growth-defining, -limiting and -reducing factors to actual yields. This is traditionally performed at the field level using mechanistic crop growth simulation models, and directly up-scaled to the regional and global levels without considering a range of factors intersecting at farm and farming system levels. As an example, these may include farmers' objectives and resource constraints, farm(er) characteristics, rotational effects between subsequent crops or decisions on resource allocation and prioritization of crop management. The objective of this thesis is to gain insights into yield gaps from a farm(ing) systems perspective in order to identify opportunities for sustainable intensification at local level.

Three contrasting case studies representing a gradient of intensification and capturing a diversity of agricultural systems were selected for this purpose, namely mixed crop-livestock systems in Southern Ethiopia, rice based-farming systems in Central Luzon (Philippines) and arable farming systems in the Netherlands. A theoretical framework combining concepts of production ecology and methods of frontier analysis was developed to decompose yield gaps into efficiency, resource and technology yield gaps. This framework was applied and tested for the major crops in each case study using crop-specific input-output data for a large number of individual farms. In addition, different statistical methods and data analyses techniques were used in each case study to understand the contribution of farmers' objectives, farm(er) characteristics, cropping frequency and resource constraints to yield gaps and management practices at crop level.

Yield gaps were largest for maize and wheat in Southern Ethiopia (ca. 80% of the water-limited yield), intermediate for rice in Central Luzon (ca. 50% of the climatic potential yield) and smallest for the major arable crops in the Netherlands (ca. 30% of the climatic potential yield). The underlying causes of these yield gaps also differed per case study. The technology yield gap explained most of the yield gap observed in Southern Ethiopia, which points to a lack of adoption of technologies able to reach the water-limited yield. The efficiency yield gap was most important for different arable crops in the Netherlands, which suggests a sub-optimal timing, space and form of the inputs applied. The three intermediate yield gaps contributed similarly to the rice yield gap in Central Luzon meaning that sub-optimal quantities of inputs used are as important in this case study as the causes mentioned for the other case studies.

Narrowing the yield gap of the major crops does not seem to entail trade-offs with gross margin per unit land in each case study. However, the opposite seems to be true for N use efficiency and labour productivity particularly in Southern Ethiopia and Central Luzon, and to a less extent in the Netherlands. This means that (sustainable) intensification of smallholder agriculture in the tropics needs to go hand-in-hand with agronomic interventions that increase land productivity while ensuring high resource use efficiency and with labour-saving technologies that can reduce the drudgery of farming without compromising crop yields.

Other insights at farm(ing) system level were clearer in Southern Ethiopia than in Central Luzon or in the Netherlands. For example, alleviating capital constraints was positively associated with intensification of maize-based farming systems around Hawassa and increases in oxen ownership (an indicator of farm power) was associated with extensification of wheat-based farming systems around Asella. In Central Luzon, farm and regional factors did not lead to different levels of intensification within the variation of rice farms investigated and the most striking effect was that direct-seeding (and thus slightly lower rice yields) was mostly adopted in larger farms, and used lower amounts of hired labour, compared to transplanting. In the Netherlands, the analysis of rotational effects on crop yields provided inconclusive results but confounding effects with e.g. rented land do not allow to conclude that these are not at stake in this farming system.

This thesis broadens the discussion on yield gaps by moving from the technical aspects underlying their estimation towards the broader farm level opportunities and constraints undermining their closure. Overall, insights from contrasting case studies support conventional wisdom that intensification of agriculture needs to occur in the 'developing South', where yield gaps are large and resource use efficiency low, while a focus on improving sustainability based on sustainable intensification (or even extensification) is more appropriate in the 'developed North', where yield gaps are small and resource use efficiency high.

Keywords: Agronomy, Production ecology, Stochastic frontier analysis, Yield variability, Farm performance, Integrated assessment, Philippines, Netherlands, Ethiopia

Table of contents

CHAPTER 1

Introduction

1.1 Background information

Agriculture has been a crucial activity to provide nutritious food in sufficient quantities to sustain human societies throughout their history. The 'Malthusian catastrophe' predicted for the twentieth century was postponed first through the expansion of cultivated land and later through considerable increases in land and labour productivity (Pongratz et al., 2008; Timmer, 1988). These gains in productivity sector can be attributed to the continuous breeding of plant species (Khush, 2001), the improvement and intensification of crop management practices (Tilman et al., 2002), expansion of irrigated areas (Siebert et al., 2015) and the substitution of labour by energy, and capital, in the production process (de Wit, 1975). Future increases in agricultural output need to come from further productivity gains because continued expansion of cultivated and irrigated land is limited by environmental concerns and existing water supplies (Cassman, 1999).

Intensification of agriculture relied mostly on fossil fuel inputs (Connor et al., 2011; Koning et al., 2008; Tilman et al., 2002) and other non-renewable resources, such as rock phosphate (Cordell et al., 2009). These resulted in a number of externalities such as disruption of the bio-geochemical cycles at global level (Bouwman et al., 2011; Vitousek et al., 2009) and environmental pollution (e.g. Iversen et al., 1998) or loss of biodiversity at local level (e.g. Donald et al., 2011; Féon et al., 2010). These occurred in tandem with a structural transformation in the national economy of modern societies as a result of economic growth (Timmer, 2009). This encompassed a sharp decline in the contribution of agricultural production as a source of income and employment over the past century. The declining economic importance of agriculture, together with its negative externalities and future prospects of resource scarcity, challenge the long-term viability, and sustainability, of the current food system. In this context, sustainable and ecological intensification were proposed as suitable strategies to reconcile agricultural production and environmental quality (Tilman et al., 2011; Cassman, 1999).

There are wide regional differences in the level of intensification and the magnitude of environmental impacts, and these need to be considered in the sustainable intensification debate (see below). Official statistics indicate that cereal productivity increased by about 13 kg ha^{-1} yr^{-1} in Eastern Africa, 55 kg ha^{-1} yr^{-1} in Southeast Asia and 94 kg ha^{-1} yr^{-1} in Western Europe between 1960 - 2014 (Figure 1.1A). During the same period, there was a marginal increase in N fertiliser consumption in Eastern Africa, which contrasts with the steady linear increase in Southeast Asia and with the decline observed in Western Europe after the 1980s (Figure 1.1B). These trends in agricultural intensification are not independent from the broader macro-economic setting of economic development (Zhang et al., 2015; Lassaletta et al., 2014; Mandemaker et al., 2011; Tilman et al., 2011). The share of agricultural production to GDP between 1970 - 2014 declined from ca. 30% to 25 % in Eastern African and

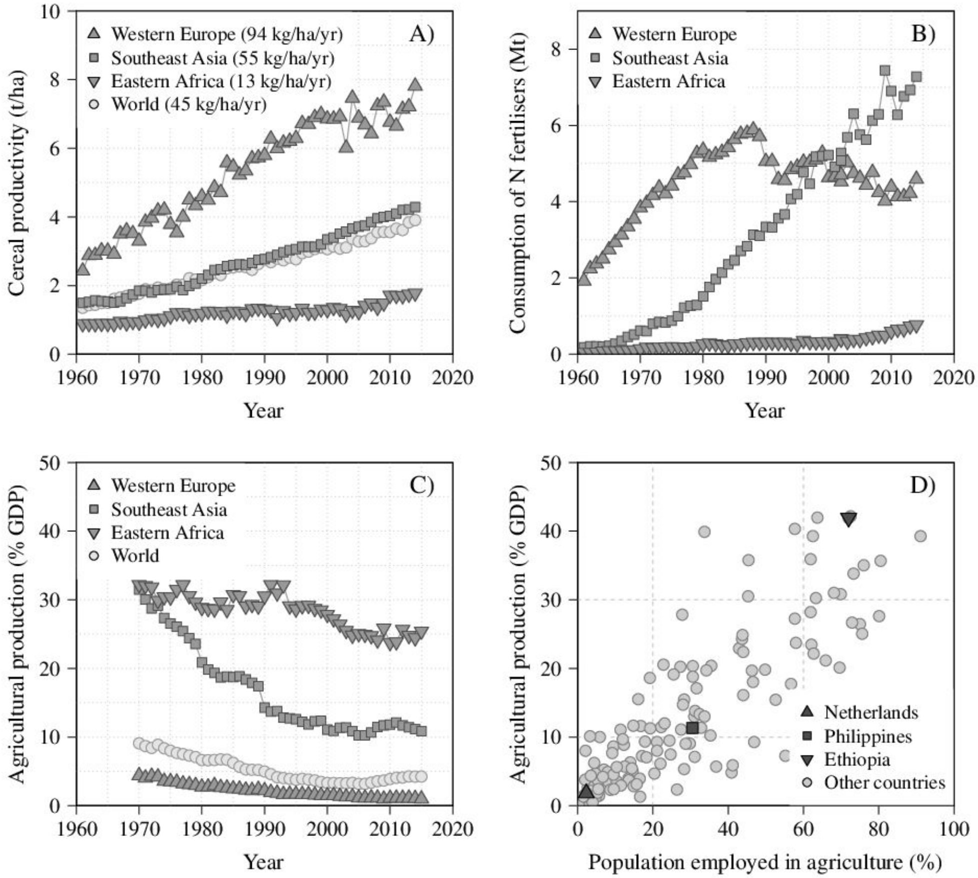

Figure 1.1: Trends in agricultural productivity and development: A) trends in cereal productivity worldwide and for selected regions based on FAOSTAT; B) trends in N fertiliser consumption in selected regions based on FAOSTAT; C) declining share of agricultural production to gross domestic product (GDP) for the entire world and for selected regions based on FAOSTAT; D) contribution of agricultural production to GDP and share of the population employed in agriculture for selected countries based on data from the World Bank and the International Labour Organization (IOL), respectively.

10% in Southeast Asia, while in Western Europe it remained below 5% during the entire period (Figure 1.1C). Large differences between countries can still be observed nowadays (Figure 1.1D).

Yield gaps and resource use efficiencies are important indicators to benchmark agricultural systems (van Ittersum et al., 2013). Traditionally, these have been assessed at the field level and directly up-scaled to the regional level, without considering explanatory factors at farm(ing) systems level. Decisions on resource allocation and prioritization of crop management are generally taken at the farm level (Giller et al., 2011) and need to be considered when explaining yield gaps. Moreover, the state of economic and agricultural development in which a farming system operates also affects its performance. For example, there is still a large potential to increase crop yields in sub-Saharan Africa (Tittonell and Giller, 2013), while governments in Western Europe adopted environmental legislation in the early 1980s to restrict input use (Henkens and van Keulen, 2001). A comparative analyses of farming systems in different stages of intensification is thus highly relevant to contextualize yield gaps and resource use efficiencies.

The aim of this thesis is to assess the scope for sustainable intensification of current agricultural systems at local level. For this purpose, a generic method is required to decompose and explain yield gaps using individual farm data. Some important aspects to be considered at farm level include alternative farmers' objectives, farm(er) characteristics, rotational effects and resource constraints. The thesis builds upon detailed yield gap analyses conducted in three contrasting case studies, namely arable farming in the Netherlands, rice farming in Central Luzon (Philippines) and smallholder farming in Southern Ethiopia.

1.2 The sustainable intensification debate

Ecological and sustainable intensification arouse as strategies to reconcile agricultural production on the one hand and environmental quality and resource scarcity on the other hand (Garnett et al., 2013; Tilman et al., 2011; Cassman, 1999). Ecological and sustainable intensification share a similar definition but differ on the discourse regarding the role of ecological processes for agricultural intensification (Tittonell, 2014; Bommarco et al., 2013). As such, their dominant paradigm lies on increasing resource use efficiency (RUE, i.e. increment output per increment input) through yield gap closure (i.e. increasing output per unit land) while reducing the use, and need, for external inputs (i.e. decreasing input per unit land). In terms of methods, ecological intensification focuses on re-designing current production systems according to the principles of 'agro-ecology' (e.g. Tittonell, 2014) while sustainable intensification looks at improving the 'eco-efficiency' of current production systems through site-specific management interventions (Keating et al., 2010).

Trade-offs are at the core of sustainable intensification, both in its framing and its essence (Pretty and Bharucha, 2014; Struik et al., 2014; Garnett et al., 2013). As an example, the law of diminishing returns states that the highest RUE for a single input is reached at low input rates. Different studies showed that achieving high yields requires a disproportional increase of single inputs, which result in sub-optimal RUE (Tilman et al., 2011; Nijland et al., 2008). This conflicts with the hypothesis that RUE increases with increasing yield levels due to further optimisation of growing conditions (de Wit, 1992). Moreover, it indicates that greater gains in RUE can be obtained through tuning different input levels in less intensive production systems because low levels of one input limit the RUE of other inputs.

The implications of sustainable intensification for economic performance and labour productivity at different scales also remain unclear. However, trade-offs may be expected in case of unfavourable input-output price ratios or additional labour/costs are required to eliminate small crop management imperfections. The paradigm of sustainable intensification is thus contested due to the existence of trade-offs, which require that some choices are made outside the realm of science, and the lack of consensus on the actual definition of 'sustainable' and 'intensification' (Struik et al., 2014). This is further hindered by the lack of strong empirical evidence on how exactly sustainable intensification can be achieved at local level, especially given that its scope is context-specific and dependent on the intensification level of a particular farming system (e.g. Carberry et al., 2013; Koning et al., 2008).

As part of this on-going debate, it is important to provide quantitative evidence of the current magnitude and causes of yield gaps and RUEs in farmers' fields, where sustainable intensification is expected to take place in the years to come. In addition to quantifying yield gaps, it is necessary to identify their drivers at different scales and to make trade-offs between yield gap closure and other objectives explicit. Comparative analyses of farming systems should be pursued to capture diversity in responses at local level.

1.3 Explaining yield gaps at farm level

Yield gap analysis in agronomy is rooted in the concepts of production ecology (van Ittersum and Rabbinge, 1997). Its purpose is to understand the relative contribution of growth-defining, -limiting and -reducing factors to actual yields. Growth-defining factors comprise biophysical conditions (e.g. solar radiation and air temperature) and plant characteristics (related to phenology and physiology), which at optimal supply of all inputs determine the potential yield (Yp). Growth-limiting factors refer to the essential abiotic inputs water and nutrients and these, if at sub-optimal supply, determine water-limited and nutrient-limited yields (Yw and Yn), respectively. Growth-reducing factors include biotic stresses such as pests, diseases and weeds and other abiotic factors such as pollutants. Actual yields (Ya) are a result of water and nutrient limitations and damage from growth-reducing factors.

The yield gap can be defined as the difference between Yp (irrigated conditions) or Yw (rainfed conditions) and Ya (van Ittersum et al., 2013). This is typically done at field level (Lobell et al., 2009) and further up-scaled to the regional level (e.g. van Bussel et al., 2015), while neglecting the broader farm level conditions in which yield gaps are embedded. Farmers make decisions on resource allocation and prioritization of crop management at the farm level and given their personal objectives, resource constraints and farm(er) characteristics (e.g. Giller et al., 2011). These decisions influence both crop and farm performance, due to synergies and trade-offs between different activities taking place in a single farm or between closing the yield gap and optimizing other dimensions of farm performance. Therefore, explaining yield gaps requires in-depth analysis at farm level to understand if, and by how much, crop yields and RUEs can be increased as well as the associated trade-offs of doing so.

Estimation of yield gaps with local-to-global relevance requires a transparent, robust and reproducible protocol following a 'bottom-up' approach (van Ittersum et al., 2013). Crop models have been the preferred tool to quantify yield gaps and to explore the broader 'operational space' within which food production operates (e.g. van Ittersum et al., 2016). They are also useful for more detailed assessments aiming to explain yield gaps in relation to genotype, environment and management interactions (G $\times$ E $\times$ M; Kersebaum et al., 2007; Passioura, 1996) and to quantify yield gaps at cropping systems level (Guilpart et al., 2017). Nonetheless, the fact that crop models can only deal with some of the biophysical aspects of crop production (van Ittersum et al., 2003) indicates that complementary approaches are required to gain further insights into yield gaps from a farm(ing) systems perspective. Those should make use of individual farm level data and capture the role of both biophysical and socio-economic factors. Moreover, they should link yield gaps and RUEs at crop level with decisions on resource allocation and prioritization of crop management at farm level. This requires the development of a suite of methodologies to unpack variation in crop and farm performance across farm systems and offers an entry point to identify options for sustainable intensification.

1.4 The farm level as unit of analysis

The farm is the level at which farmers make decisions on how to organize and manage their resources. Farmers' decisions can be classified as strategic, tactical and operational, depending on their temporal scale and comprehensiveness. Strategic decisions refer to non-routine and long-term decisions such as the choice of a particular production system (crop, animal or mixed) or production technique (e.g. adoption of organic agriculture standards). Tactical decisions stand for medium term decisions such as e.g. the number, type and sequence of activities performed in a farm (e.g. Vereijken, 1997). Finally, operational decisions include the day-to-day activities regarding how inputs/resources are allocated (e.g. Tittonell et al., 2007), and how management operations are prioritized (e.g. Kamanga et al., 2014).

The extent to which growth-defining, -limiting and -reducing factors are optimised for a specific crop depends on factors at different levels (Figure 1.2). Strategic, tactical and operational decisions are reflected in the crop management used by farmers, which in turn explains yield gaps and resource use (in)efficiencies in the biophysical environment of the farm. For instance, growth-defining factors are associated with tactical decisions at cropping system level influencing the length of the growing season. Growth-limiting factors and -reducing factors comprise the quantity of inputs used as well as the time, space and form of application or the method used for biotic control. These further relate to factors at field, farm and regional levels (Figure 1.2). As an example, the timing of application is affected by field conditions and on the availability of family and hired labour, which are associated with household composition, land fragmentation, knowledge, availability of farm power/capital or by labour markets in a given region, among other issues.

Tactical decisions can also affect crop performance by stimulating or inhibiting interactions between different activities at cropping systems level. These may include manipulation of the length of the growing season by adjusting planting and harvesting dates and selecting early or late maturing cultivars (Guilpart et al., 2017; Hochman

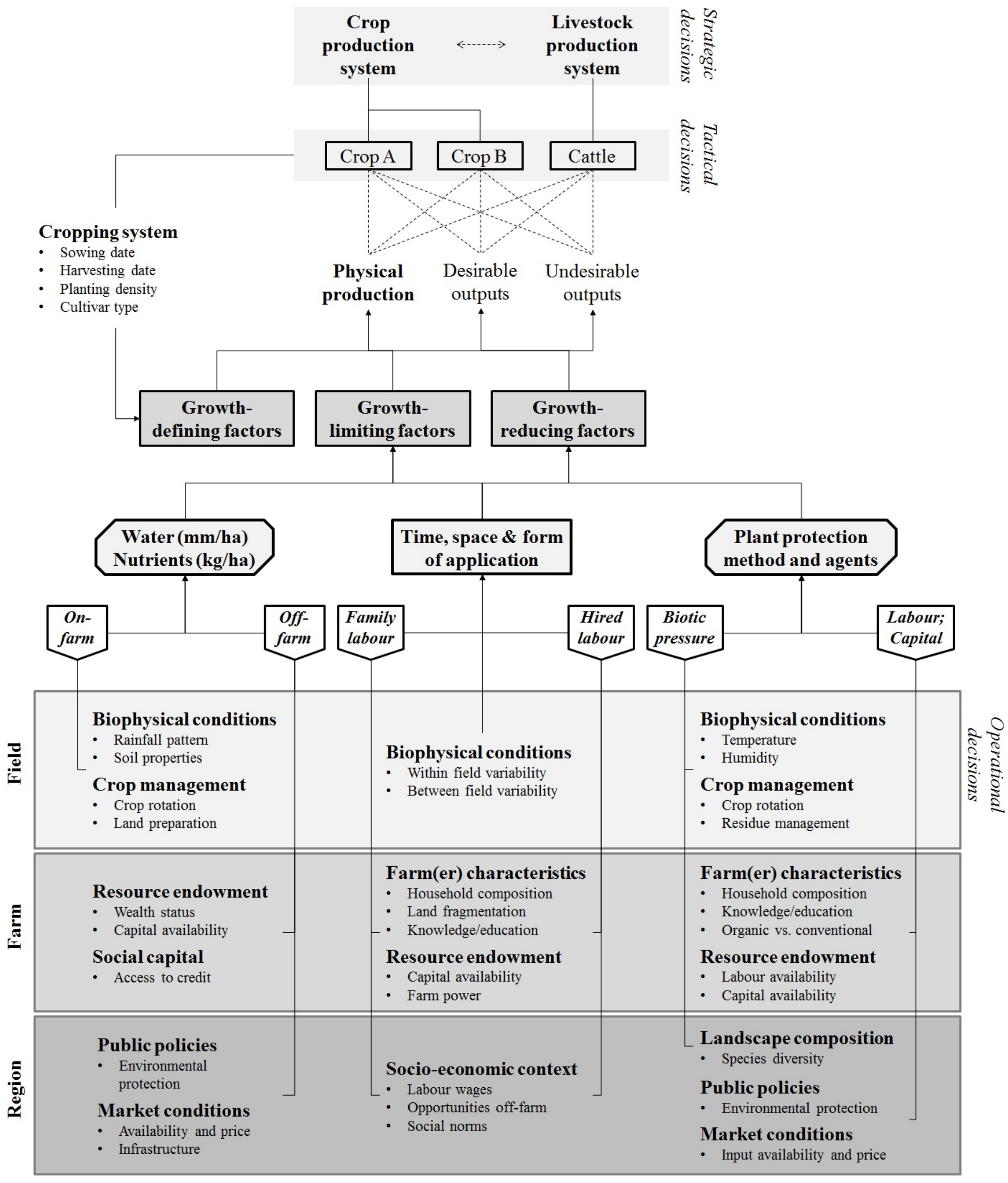

Figure 1.2: Unravelling the contribution of field, farm and regional factors to growth-limiting and -reducing factors and, hence to crop yield gaps in on-farm situations. Cropping system factors associated with the length of the growing season affect growth-defining factors. The factors listed at different levels are not exhaustive and should be taken as examples.

et al., 2014; Fletcher et al., 2011) as well as the diversity, frequency and sequence of crops in a rotation (Andert et al., 2016; Mazzilli et al., 2016; Dogliotti et al., 2003; Struik and Bonciarelli, 1997). For instance, legumes can reduce N fertilizer requirements of subsequent cereal crops (e.g. Plaza-Bonilla et al., 2017) and crop rotations can be used as a strategy to manage reducing factors such as diseases and weeds (Struik and Bonciarelli, 1997). In addition to rotational effects, competition for scarce resources (e.g. labour) in specific periods of the growing season can also reduce crop yields in farms with multiple, and overlapping, activities (Kamanga et al., 2014; Baudron et al., 2012). In those situations, operational decisions will determine the actual impact of alternative resource allocation strategies on crop yields.

The last aspect central to decision-making at farm level is the role of farmers' objectives, resource constraints and farm(er) characteristics. These aspects have deserved a large attention in the bio-economic modelling literature (van Wijk et al., 2014; Janssen and van Ittersum, 2007) but not as much in the yield gap literature. Although inputs and resources should be used as efficiently as possible from an economic and environmental perspective, farmers' objectives are much broader than minimizing yield and RUE gaps. In the words of Struik et al. (2014), and references therein, farmers may strive to achieve other objectives such as "independence of input markets, income from non-agricultural sources, peace of mind, cultural heritage or short-term economic gain". It is unclear whether sustainable intensification through yield gap closure can contribute, and by how much, to achieve such objectives given potential trade-offs at the farm level. This caveat calls for an integrated assessment of yield gaps at the farm level as well.

1.5 Benchmarking farm(ing) systems

The yield gap provides an indication of how (in)efficiently land is used due to (in)efficient use of inputs. This means that yield gaps are closely linked to RUEs (cf. Figures 1.3A and 1.3B; van Noordwijk and Brussaard, 2014; Setiyono et al., 2010; Tittonell et al., 2008b; Sadras and Angus, 2006; Witt et al., 1999), defined as the amount of output produced per unit input used in a given production system. In case of water and nutrients, RUE can be further conceptualized as the product of conversion and capture efficiencies (de Wit, 1992). Conversion efficiency refers to the amount of output produced per unit input uptake and relates mostly to plant architecture (harvest index) and physiology (photosynthetic rate). Capture efficiency stands for the amount of input uptake per unit input applied and it is affected by the amount, time, space and form of input application. Differentiating these two components is useful to identify management practices which can best improve RUE (Giller et al., 2006).

Interactions between different inputs are central in the analysis of yield gaps and RUE across multiple farms (de Wit, 1992). Assuming a situation in which water and N are the only limiting factors to production, it is possible to disentangle the effects of input quantity from the effects of input timing, space and form of application when explaining the yield gap (Figures 1.3A and 1.3B). For Yp and Yw water use efficiency is at its optimum and N use efficiency depends on the amount of water supplied. Differently, the fact that Ya is below the response curve to water and N indicates that both inputs are used inefficiently. Supplying water and N at the right time, form and space results in the projection of Ya on the response curve to N, but not to water. Increasing the amount of N is needed to further narrow the yield gap between Yw and Ya, and to fully project Ya on the response curve to water. Finally, achieving Yp requires that additional water is supplied which leads to gains in N use efficiency by shifting the response curve to N upwards. This example provides a framework to unpack Ya variability across farmers' fields while explaining the yield gap and highlights the importance of considering different inputs simultaneously.

Methods of frontier analysis were developed in agricultural economics to estimate technical (in)efficiency for individual producers based on production functions with multiple inputs and outputs (Thiam et al., 2001; Farrell, 1957). Technical efficiency (TE) refers to the ability to minimize input use in the production of a given output vector (input-oriented TE), or the ability to obtain maximum output from a given input vector (output-oriented TE; Kumbhakar and Lovell, 2000). Input-oriented TE, $\theta \leq 1$, measures the maximum contraction of the input vector that produces of Ya (x^A) in relation to the production frontier (Figure 1.3C) and to isoquant L(Ya) (Figures 1.3D). Output-oriented TE, $\varphi^{-1} \leq 1$, measures the reciprocal of the maximum expansion of Ya that is feasible with the input vector x^A in relation to the production frontier (Figure 1.3C). This implies the expansion of the isoquant L(Ya) to $L(\mathrm{Y_{TEx}})$ since $x^A \in L(\mathrm{Y_{TEx}})$ (Figure 1.3D). If growth-defining, -limiting and -reducing factors determine the production frontier (e.g. Neumann et al., 2010), then the output-oriented measure of technical inefficiency can be interpreted as the yield gap between 'technical efficient yields' ($\mathrm{Y_{TEx}}$, i.e. the maximum yield that can be obtained with current input levels) and actual farmers' yields (Ya), both attained with the input vector x^A.

This method has been widely applied for yield gap analysis at different scales (Henderson et al., 2016; Beddow et al., 2014; Carberry et al., 2013; Hoang, 2013; Neumann et al., 2010). As multiple inputs can be considered in the estimation of the production frontier, it is possible to isolate the contribution of ’input quantity’ and ’input timing, space and form of application’ to crop yields (Figures 1.3A and 1.3B). The quantification of technical efficiency also allows the identification of feasible options for sustainable intensification because technical efficient farms are, by definition, also resource use efficient at a specific input level.

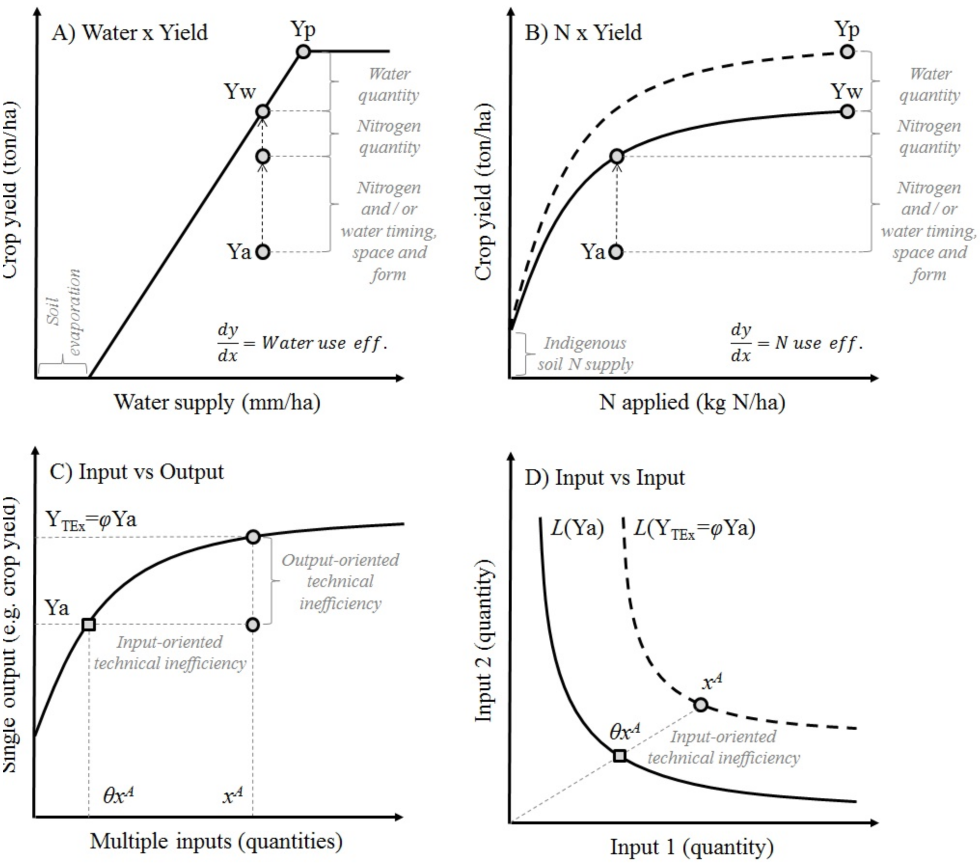

Figure 1.3: Concepts used to decompose and explain crop yield gaps: A) yield response to water supply (French and Schultz, 1984), B) yield response to N applied (Giller et al., 2004), C) representation of technical inefficiency and D) isoquants L(Ya) and $L(\mathrm{Y_{TEx}})$ underlying assumptions on substitutability between inputs to produce respectively Ya and $\mathrm{Y_{TEx}}$ (Kumbhakar and Lovell, 2000). The scalars $\varphi \geq 1$ and $\theta \leq 1$ refer to the maximum expansion of Ya and (radial) contraction of x^A associated with output- and input-oriented technical efficiency, respectively. $\mathrm{Y_{TEx}}$ = technical efficient yields and Ya = actual farmers’ yields. See text for further explanation.

Additional benchmarks are required to gain deeper insights into the causes of yield gaps. As an example, Tittonell and Giller (2013) proposed the concept of 'locally attainable yield', which stands for "*the maximum yield achievable by resource endowed farmers in their most productive fields*". The concept of 'technical efficient yields' (Figure 1.3C) introduced above is another example. The original benchmarks proposed by van Ittersum and Rabbinge (1997), namely Yp and Yw, can be used in combination with frontier analysis to verify if 'technical efficient yields' are within (biophysically) feasible ranges (Stuart et al., 2016; van Ittersum et al., 2013). Moreover, they can be used to further disentangle the yield gap between Yp or Yw and the maximum 'technical efficient yields' and hence, to assess whether or not current technologies used by farmers can achieve the biophysical yield ceilings.

1.6 Objectives and hypotheses

The main objective of this thesis is to explain crop yield gaps at farm(ing) systems level in order to identify options for sustainable intensification at local level. A methodological protocol was developed and applied to individual farm level data to decompose yield gaps into efficiency, resource and technology yield gaps; yield gaps were explained at the farm level based on farmers' objectives, farm(er) characteristics, rotational effects and resource constraints; and options for sustainable intensification were identified through comparative analysis of three case studies with farming systems at different stages of intensification. The specific objectives were to:

1. Review the yield gap explaining factors identified in the literature so far, to assess data availability and suggest improved data collection approaches;
2. Develop and test a method to decompose yield gaps into efficiency, resource and technology yield gaps across contrasting farming systems;
3. Identify constraints and stimuli at field, farm and regional level to narrow rice yield gaps in Central Luzon, Philippines;
4. Explain yield gaps in Dutch arable farms based on cropping frequencies observed in farmers' fields as well as farmers' objectives;
5. Understand if, and to which extent, labour is a limiting factor to maize and wheat yield gaps across smallholder farms in Southern Ethiopia;
6. Delineate the role of yield gap analysis within a development-oriented agronomy, taking into account variability of yield gaps and different livelihood strategies; and
7. Explore options for sustainable intensification through a comparative analysis of farming systems at different stages of intensification.

The hypotheses tested were: 1) the efficiency yield gap is most important in the Netherlands, the technology yield gap is most important in Ethiopia and the three yield gaps are of similar importance in the Philippines; 2) yield gaps are determined by farm level conditions related to farm(er) characteristics and cropping frequencies as well as farmers' objectives and resource constraints and; 3) agriculture needs to be intensified in the 'developing South' while improving sustainability, through sustainable intensification, is more applicable in the 'developed North'.

1.7 Methodological approach

1.7.1 Selection of case studies

Three in-depth case studies were conducted in this thesis to contextualize yield gaps within local biophysical and socio-economic conditions. A case study approach at local level was preferred to a yield gap analysis at global level (Mueller et al., 2012; Licker et al., 2010; Neumann et al., 2010). The agronomic relevance of most global studies published so far is rather limited due to inaccurate representation of biophysical yield ceilings and inadequate representation of crop management practices (van Ittersum et al., 2013). For these reasons, yield gap analysis should be conducted at local level, and up-scaled *a posteriori* using a transparent protocol (van Bussel et al., 2015; Grassini et al., 2015). In this way it is possible to verify estimated yield gaps with on-farm data and experiments and to accommodate the large variation observed between and within farm(ing) systems.

Arable farming systems in Northwest Europe operate close to the potential or water-limited yield (www.yieldgap.org) and RUEs are strongly affected by economic objectives, environmental legislation and production quotas for certain crops. These conditions are very present in the Netherlands, a small country where agricultural intensification is running up to its limits (Bos et al., 2013). For instance, average yields of winter wheat have been above 8.0 t ha^{-1} since the mid 1980s and the average N application rate at national increased sharply up to a maximum of ca. 800 kg N ha^{-1} in the early 1980s, after which it declined stepwise to about half that value from 2010 onwards (Figure 1.4). Moreover, specialized arable farms tend to focus on profit maximisation in the short term while environmental indicators are mostly considered in their strategic planning (Mandryk et al., 2014). Farmers also need to comply with strict environmental regulations regarding for example nutrient use (Schröder and Neeteson, 2008).

Irrigated rice farming in the lowland areas of Southeast Asia have changed dramatically over the past half-century with sharp declines in consumer rice prices, substitution of family by hired labour and adoption of labour-saving technologies (Hossain and Fischer, 1995). Intensification of irrigated rice systems during this period was largely driven by both the expansion of irrigated areas and the adoption of Green Revolution technologies (Cassman and Pingali, 1995). However, rice yield gaps as high as 2.0 - 5.0 t ha^{-1} are still observed in some regions (Laborte et al., 2012). The impact of these structural changes as well as persisting yield gaps are especially present in Central Luzon (Moya et al., 2015), the rice bowl of the Philippines. Annual rice productivity in the Philippines more than doubled in the period 1960 - 2014: from less than 2 t ha^{-1} in the 1960s to ca. 4 t ha^{-1}from 2005 onwards (Figure 1.4A). The average N application to cropland at national level increased over time to a maximum of ca. 50 kg N ha^{-1} after the 2000s (Figure 1.4B).

Contrary to Southeast Asia, smallholder farming in sub-Saharan Africa was largely by-passed by the Green Revolution and yield gaps remain large for all the crops in most regions (van Ittersum et al., 2016; Tittonell and Giller, 2013). The need for sustainable intensification of African agriculture is thus widely acknowledged (Vanlauwe et al., 2014; Pretty et al., 2011). However, it is also acknowledged that sustainable intensification needs to occur in a context where 1) farmers lack investment capacity and expect immediate benefits from farming, 2) farm(ing) systems are very

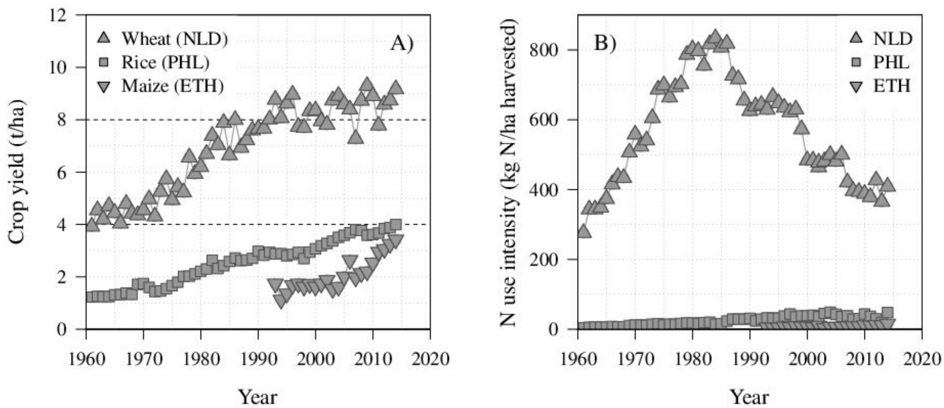

Figure 1.4: Crop yields and N use intensity in the Netherlands, Philippines and Ethiopia for the period 1960 - 2014. Crops yields refer to wheat in the Netherlands, rice in the Philippines and maize in Ethiopia. N use intensity was calculated as the amount of N fertiliser consumed per unit crop area harvested. All data were retrieved from FAOSTAT.

diverse and heterogeneous, 3) farm sizes are often too small to intensify sustainably and 4) labour supply is expected to decline due to migration from rural to urban areas. Ethiopian agriculture takes place in a wide range of agro-ecologies and is no exception to the aforementioned trends: a large share of the rural population lives below the poverty line, average farm sizes are lower than 2 ha and most labour operations are performed either manually or using animal traction (CSA and WB, 2013; Aune et al., 2001). In Ethiopia, maize yields increased from less than 2 t ha^{-1} in the 1990s to ca. 3 t ha^{-1} in 2010 and N application rates at national level remained below 20 kg N ha^{-1} over that period (Figure 1.4).

1.7.2 Crop and farm level analysis

The methodological approach followed in this thesis is summarized in Figure 1.5. A literature review was conducted to synthesize the key yield gap explaining factors identified so far, as well as the methodologies most commonly used for analysis (Chapter 2). Input-output relationships and other socio-economic information (e.g. prices and labour use) were compiled from existing household surveys to quantitatively describe the main farming system in each case study. A methodological protocol combining production ecology (van Ittersum and Rabbinge, 1997) and frontier analysis (Farrell, 1957) was developed to decompose the yield gap of rice in Central Luzon (Chapter 3), winter wheat, spring barley, ware, seed and starch potato, sugar beet and spring onion in the Netherlands (Chapter 5) and, maize and wheat in Southern Ethiopia (Chapter 6). Production frontiers were estimated to assess the contribution of management practices to yield gaps and to estimate and explain the efficiency yield gap, the agronomic equivalent of output-oriented technical inefficiency (Figure 1.3C). Largest, average and lowest yielding farms were compared to assess whether higher yields were obtained due to greater input use and crop models were applied to estimate the biophysical yield ceilings Yp or Yw.

Yield gaps were further explained, first based on farm level conditions regarding farm(er) characteristics across rice farms in Central Luzon (Philippines, Chapter 4), cropping frequencies used in Dutch arable farms (Chapter 5) and resource constraints faced by smallholder farms in Southern Ethiopia (Chapter 6). Second, an integrated assessment of management practices explaining yield gaps was performed based on different farm level indicators. These captured alternative farmers' objectives based on the principles of food production (crop yield and farm production), environment sustainability (N use efficiency), economic performance (gross margin) and labour drudgery (labour productivity). The existence, and magnitude, of synergies and trade-offs between narrowing yield gaps and optimising other indicators was assessed empirically for individual farms.

Opportunities for sustainable intensification were identified within each case study based on the projection of each farm on the production frontier and through comparisons of input levels currently used by farmers to ones required to achieve Yp or Yw. A comparative analysis across case studies was also performed based on cereal yield responses to N and P, N balances at farm level and yield gap closure vis-à-vis resource availability, gross margin per unit land and labour productivity.

1.8 Outline of the thesis

This thesis is structured according to the methodology used (i.e. literature review, yield gap analysis and a critical reflection) and to the diversity of production activities within each case study, i.e. a single crop in the rice farms of Central Luzon (Philip-

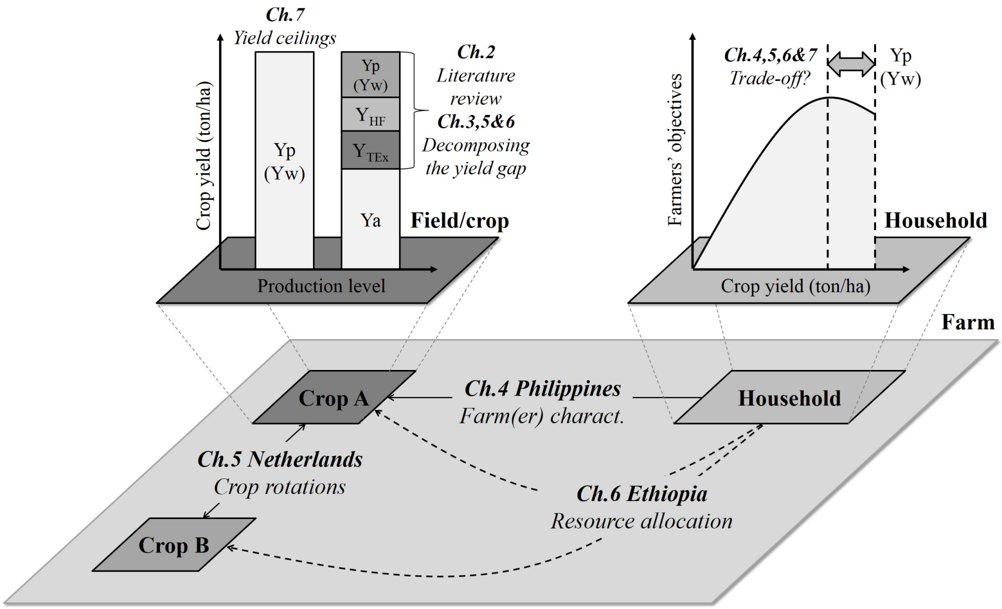

Figure 1.5: Methodological approach and outline of this thesis to decompose and explain yield gaps at crop and farm level. The aspects addressed in each chapter are highlighted in the figure. Yp = climatic potential yield defined according to van Ittersum and Rabbinge (1997); Yw = water-limited potential yield; Y_{HF} = highest farmers' yield; Y_{TEx} = technical efficient yield; Ya = actual farmers' yield.

pines), multiple crops in specialized arable farms in the Netherlands and both crop and livestock activities in smallholder farms of Southern Ethiopia. Chapter 2 summarizes the most important yield gap explaining factors identified and methods used for yield gap analysis to date. A method for decomposing yield gaps into efficiency, resource and technology yield gaps is presented in Chapters 3, 5 and 6 for rice-based farming systems in the Philippines, arable farming systems in the Netherlands and mixed crop-livestock systems in Southern Ethiopia, respectively. Further insights into the effects of farm(er) characteristics, cropping frequencies, labour/capital availability and alternative farmers' objectives to yield gaps in those case studies are provided in Chapters 4, 5 and 6. A reflection paper on the importance to accommodate variability in the yield ceilings and the broader livelihood aspects where farmers operate within yield gap analysis, particularly in developing countries, is provided in Chapter 7. This chapter draws upon empirical on-farm data from Western Kenya and Central Luzon. Finally, Chapter 8 provides the key findings of this thesis and discusses options for sustainable intensification based on a comparative analysis of crop and farm level indicators across the three case studies.

CHAPTER 2

Review of yield gap explaining factors and opportunities for alternative data collection approaches

This chapter is based on:
Beza, E.; **Silva, J.V.**; Kooistra, L.; Reidsma, P. 2017. Review of yield gap explaining factors and opportunities for alternative data collection approaches. *European Journal of Agronomy*. 82, Part B, 206 - 222.

J.V. Silva contributed to the literature review, development and harmonisation of the database, and writing of the manuscript.

Abstract

Yield gap analysis is gaining increased attention, as estimating and explaining yield gaps shows the potential for sustainable intensification of agricultural systems. Explaining yield gaps requires detailed information about the biophysical environment, crop management as well as farm(er) characteristics and socio-economic conditions in which farmers operate. However, these types of data are not always available, possibly because they are costly to collect. The main objective of this research is to assess data availability and data collection approaches for yield gap analysis, and to summarize the yield gap explaining factors identified by previous studies. For this purpose, a review of yield gap studies (50 agronomic-based peer-reviewed articles) was performed to identify the most commonly considered and explaining factors in yield gap analysis. Besides a global comparison, differences between regions, crops and methods were analysed as well. The results show that management and edaphic factors are considered more often to explain the yield gap compared to farm(er) characteristics and socio-economic factors. However, when considered, both farm(er) characteristics and socio-economic factors often explain the yield gap. Fertilization and soil fertility factors are the most considered ones in the management and edaphic categories. In the fertilization group, factors related to quantity (e.g. N fertilizer quantity) are considered more often compared to factors related to timing (e.g. N fertilizer timing). However, factors related to timing explained the yield gap more often when considered. Explaining factors vary among regions and crops. For example, while soil fertility is considered relatively much both in Africa and Asia, it is often explaining in Africa, but not in Asia. Agronomic methods like crop growth simulation models are often used for yield gap analysis, but are limited in the type and number of factors that can be evaluated. By contrast, qualitative methods based on expert knowledge can include the largest range of factors. Although the data included in yield gap analysis also depends on the objective, knowledge of explaining factors, and methods applied, data availability is a major limiting factor. Bottom-up data collection approaches (e.g. crowdsourcing) involving agricultural communities can provide alternatives to overcome this limitation and improve yield gap analysis.

2.1 Introduction

Sustainable intensification of agricultural systems, including the closure of existing yield gaps on currently available agricultural land, has been pointed as a possible pathway to meet the future food demand (Tilman et al., 2011). The concept of 'yield gap' is based on production ecological principles and can be estimated as the difference between a benchmark (e.g. climatic potential or water-limited yield) and the actual yield (van Ittersum and Rabbinge, 1997). This concept is particularly important because it indicates the biophysical potential available to improve agricultural production in a specific location (van Ittersum et al., 2013).

Yield gap analysis provides the foundation for identifying the most important crop, soil and management factors limiting current farm yields (van Ittersum et al., 2013; Lobell et al., 2009; Tittonell et al., 2008a; Lobell et al., 2005). Information on the magnitude of the yield gap, and associated explaining factors, is important for efficiently targeting efforts to increase crop production in a particular farming system (Affholder et al., 2012). For example, a yield gap analysis for cassava in Cambodia revealed that soil nutrients, short crop duration and weed infestation explained much of the yield gap in the area and these factors had to be improved to increase cassava yield (Sopheap et al., 2012). A number of yield gap analyses have been conducted for different crops in different agro-ecological conditions (van Ittersum and Cassman, 2013) and the results of these studies showed that the magnitude and factors that cause the yield gap vary among locations (e.g. Affholder et al., 2012).

Many studies have examined yield gaps at the scale of the region or agro-climatic zone, using aggregated data on crop yields and explaining factors (e.g. Mueller et al., 2012; Neumann et al., 2010). These type of studies are useful to compare different regions in relative terms using harmonized data (van Ittersum et al., 2013). However, yield gap analysis at local level are also necessary to gain deeper insights into the causes of yield gaps at farm level an to integrate biophysical and socio-economical conditions.

The interactions between different activities at farm level, together with resource constraints faced by individual farmers, likely explain why inputs are not optimally allocated across the farm and hence why yield gaps persist (e.g. Tittonell et al., 2008a). Therefore, yield gap analysis at farm and farming system level can contribute to better understand whether or not yield gaps can be narrowed and if so, under which production, economic and environmental conditions (Giller et al., 2006). A major drawback of this type of analysis is the high data standards required, which typically refer to a) large sample size, b) fine resolution and c) great level of detail. Clearly, obtaining information about biophysical characteristics and crop and farm management for individual agricultural activities within a farm, as well as farm and farmer's characteristics and socio-economic conditions for a large number of farms is costly and time-consuming. Nowadays, the proliferation of computing devices like different types of mobile phones equipped with sensors (e.g. GPS), and other similar technologies makes it possible to implement effective and low-cost 'bottom-up' data collection approaches such as crowdsourcing (Ferster and Coops, 2013). These facilitate the collection of relatively large amounts of information directly from local communities (Herrick et al., 2013; Pratihast et al., 2013).

The main objective of this research is to review the yield gap explaining factors identified by previous studies, in order to assess data availability and suggest improved data collection approaches for yield gap analysis. To address this objective, the following steps were undertaken: (1) to provide an overview of factors considered and explaining yield gaps; (2) to identify most commonly considered and often explaining factors of the yield gap at the global, regional and crop levels; (3) to investigate if there are regional similarities or differences in the factors which are commonly considered and explaining yield gaps; (4) to identify the most common data sources for the different factors considered for yield gap analysis; (5) to evaluate to which extent innovative data acquisition methods (e.g. crowdsourcing) are relevant for improving data availability.

2.2 Methodology

2.2.1 Literature search and study selection

A detailed literature search was carried out as starting point for this review. The selection of papers was made through specific searches for peer-reviewed articles on yield gap analysis in agronomic journals with key words 'yield gap', 'potential yield', 'yield variability', 'water-limited yield' and 'yield gap variability'. The initial focus was on a special issue released by Field Crops Research on yield gap analysis (van Ittersum and Cassman, 2013). In addition, whenever peer-reviewed articles related to yield gap analysis were found in the reference list of an already reviewed article, they were analysed and included for our study. However, priority was given to articles which explained yield gaps and/or yield variability rather than only estimating the yield gap. The review was not completely systematic, as using a keywords-based approach resulted in a large amount of papers that were not directly relevant for this review, as they did not explain yield gaps. Although some relevant papers may be missing due to this, the selected papers provide a good basis to reach our objectives.

2.2.2 Review of studies and construction of database

A database was created using MS-Excel 2010 in order to store the information from the selected articles. The database consists of five different tables, namely: 'yield gap', 'determining factors', 'considered factors', 'explaining factors' and 'validation table'. Each of the tables was organised in such a way that information about the five main categories (i.e. climate, edaphic, management, farm characteristics and socio-economic) was stored separately. All of the tables were linked with unique identifiers (IDs) to facilitate information retrieval.

Information about the study sites including the continent, country, administrative region and site names and their respective coordinates were compiled in the 'yield gap' table. When the coordinates of the study sites were not provided, the names of the study sites were used as a geographic reference and Google Maps was used to retrieve the approximate coordinates of the study sites. In addition, information about the scale at which the yield gap was estimated and explained (e.g. farm, field, regional or global), resolution of data collection and the types of crops cultivated were also compiled. In this table, we also included the years in which the yield gap analysis was performed, the data sources used to estimate actual and benchmarking yields as well as the methods used to estimate the benchmarking yield (e.g. name of crop model) and the term(s) used to indicate the benchmarking yield (e.g. potential yield, attainable yield, water-limited yield or economic yield). For studies that explained the yield gap, the explanatory methods used for that purpose (e.g. boundary-line, linear regression) were included in the database as well. Finally, the purpose of the different methods (for e.g. to explain yield gap or yield variability) used within each paper were recorded.

For each of the methods used in a specific paper, the dependent variable (y) and the independent variables (x) were identified and included in the database. The independent variables were included in the 'considered factors' table of the database. Out of these 'considered factors', the ones which explained part of the yield gap and/or yield variability according to the criteria set by the specific paper were included in the 'explaining factors' table.

In order to determine the number of records (entries) per study, the following criteria were used: number of crops considered, number of locations, years in which the yield gap analysis was performed, and methods used to estimate the benchmarking yield and to explain the yield gap. One record is a unique combination of location × crop × year × benchmark yield estimation method × yield gap explanatory method. A total of 270 records with unique identifiers (IDs) were included in the database. For studies which explicitly provided the actual yield (Ya) and the benchmarking yield (Yp or Yw), the magnitude of the yield gap (%) was calculated as the difference between the benchmark yield and the actual yield divided by benchmark yield times 100%. For studies which didn't provide the values explicitly, we didn't calculate the percentage of the yield gap and it was left blank in the database. A 'validation table' was included in the database to harmonize the data and methods used and to ensure consistency across different studies. This table provides an overview of all different variables considered, including their units and was used as the basis for grouping different factors and methods.

2.2.3 Identification of factors

Determining factors

Determining factors are all the input factors that were used to estimate the benchmark yield. As the focus of this analysis is on investigating factors explaining the yield gap or yield variability, and because required determining factors mainly depend on the method used, these factors were not analysed further in this paper. In general, the least data are needed when using the highest farmers' yields as the benchmarking yield because only yield data are required in this case. Conversely, frontier analysis methods require data on yield and input use for a large number of farmers while crop models require data on climatic conditions and cultivars to simulate potential or water-limited yields. Most data are needed when using experimental data/fields, as detailed information is required about the optimal management (treatment) used to obtain the highest possible yields.

Considered and explaining factors

Considered factors are all the input factors that were used as explanatory factors in the analysis of the yield gap or yield variability. Explaining factors are the factors which, out of the considered factors, explained part of the yield gap or yield variability. For example, if correlation or regression analysis was used to explain the yield gap then all the input factors included as independent variables were classified as considered factors. Out of the considered factors those which had a statistical significant relationship with the yield gap were classified as explaining factors. If a method that was used to compare differences between groups (e.g. ANOVA) was used, and if there were statistical significant differences between the treatments and the control plot, then all the factors in the treatments were included into the database under explaining factors. In this case, all factors behind the treatments and control plot were defined as considered factors. If interviews were used to explain the yield gap, all the factors asked for were included as ’considered factors’, and all the factors that were judged to be explaining by the authors were included as ’explaining factors’. The selection of the explaining factors for our database was based on the primary results from the reviewed papers, i.e. additional information cited from other papers was not included in the database. The authors of the papers reviewed were asked to cross-check the ’explaining factors’ identified, and around 50% responded to this.

2.2.4 Grouping factors

Factors in the five main categories were classified into different groups prior to the review. The climatic factors were classified into six groups, namely: radiation, temperature, precipitation, evapotranspiration, wind speed and others. The edaphic factors were classified into four groups, namely: soil type, soil fertility, soil water and slope. The crop and farm management factors were classified into eight groups, namely: land preparation, planting, fertilization, irrigation, weeding, crop protection, crop characteristics and others. The farm characteristics were classified into five groups, namely: income, labour, training, size and intensity, and the socio-economic factors were classified into three groups, namely: institutional, technical and population. An overview with all the individual factors belonging to each of the aforementioned categories and groups is provided online as Supplementary Material.

2.2.5 Analysis of factors explaining yield gaps

The ’considered factors’ and ’explaining factors’ tables were used to quantify the percentage of considered and explaining factors within particular subsets of the database. The percentage of considered factors within a group (e.g. fertilization) was computed as the fraction between the total number of records in the ’considered factors’ table which contained at least one factor from the group of interest and the total number of unique records in the database ($n = 270$). The percentage of explaining factors within a specific group was calculated as the quotient between the total number of records in the ’explaining factors’ table which contained at least one factor from the group of interest and the total number of records in the ’considered factors’ table which contained at least one factor from the same group.

The considered percentage of the individual factors within the specific groups (e.g. considered % of N fertilization quantity) was calculated as quotient between the total number of records which included the individual factor in the 'considered factors' table and the total number of records in the database (i.e. 270 records), multiplied by 100%. The explaining percentage of the individual factors within the specific groups (e.g. explaining % of N fertilization quantity) was calculated as the quotient between the total number of records which included the individual factor in the 'explaining factors' table and the total number of records which included the individual factor in the 'considered factors' table, multiplied by 100%. Besides the global analysis, considered and explaining yield gap factors were also analysed for each continent and for each crop × continent. The same calculations were applied to subsets of the database filtered by continent and/or crop.

To calculate the percentage of factors from a specific data source (e.g. farm survey), first we counted the total number of factors in a specific category (e.g. edaphic), and second we counted total number of factors in a specific category which were provided by specific data source. We then divided the total number of factors provided by specific data source by the total number of factors in a specific category and multiplied by 100%. Considered and explaining percentages for climatic factors were not analysed as these were used mostly as determining factors.

2.3 Results

2.3.1 Crops, continents and yield gaps

A total of 14 different crops were identified and the majority of the studies focused on rice (34%), maize (28%) and wheat (26%). The majority of the studies (64%) included in the database were from Africa and Asia (Figure 2.1 and Table A1, Supplementary Material). Studies conducted in Africa were concentrated in the east, west and southern regions and studies conducted in Asia were concentrated in the southern and south-eastern regions. Studies conducted in other continents were also included but were fewer (38%) compared to the number of studies conducted in Africa and Asia.

Yield gaps for different crops in different continents ranged between ca. 5 and 90% (Table A1, Supplementary Material) and it was large even for one single crop in one country. For example, yield gaps ranged between 14 and 80% for rice in an irrigation scheme in Mauritania (Haefele et al., 2001). Although yield gaps were on

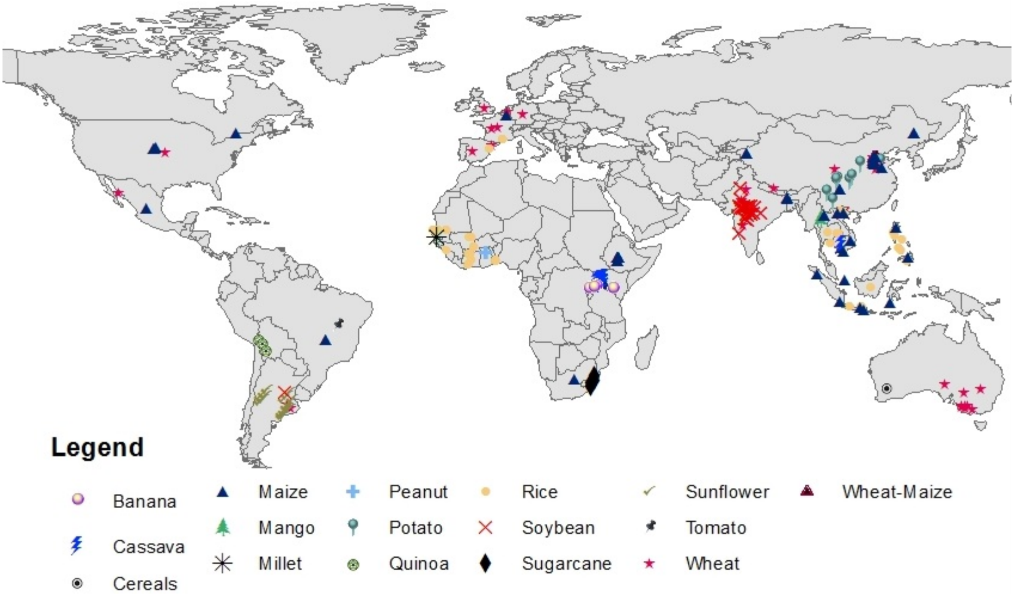

Figure 2.1: Location of the studies and the type of crops included in the review. Although not presented in the map, the global studies of Neumann et al. (2010) and Licker et al. (2010) were also reviewed and included in the database.

average smaller in Europe compared to Africa, this was not necessarily the case for rice, as the yield gap estimated for La Camargue, France (37 - 57%; Delmotte et al., 2011) was similar to the yield gap in an irrigation scheme in West Africa (27 - 49%; Wopereis et al., 1999). Thus, although it is relevant to compare yield gaps obtained by different studies, these may differ not only depending on the crop and location, but also depending on the scale of yield gap estimation, the resolution of data collection or the sources of data and methods used to assess benchmark and actual yields.

2.3.2 Between group comparison of considered and explaining factors

Management factors often explained the yield gap across all crops and locations (Figure 2.2A). Fertilization was the group most often considered, about 45% of the records. In 94% of the records where a factor related to fertilization was considered, it also explained the yield gap. Irrigation had a similar explaining power to fertilization but it was only considered in 15% of the records. Factors related to land preparation and crop characteristics explained the yield gap in 88% and 86% of the records, respectively, but only 10% of the records considered land preparation factors and 16% of the records considered crop characteristics factors. Planting, crop protection and weeding also explained the yield gap in more than 60% of the cases. It is worth noting that only one study included at least one factor from each of the management groups considered (Tanaka et al., 2013).

Edaphic factors were also important to explain the yield gap (Figure 2.2B). In 69% of the records where a factor related to slope was considered, it also explained the yield gap. Factors related to soil fertility were considered by a relatively larger number of records (25%) compared to the other groups of factors and explained the yield gap

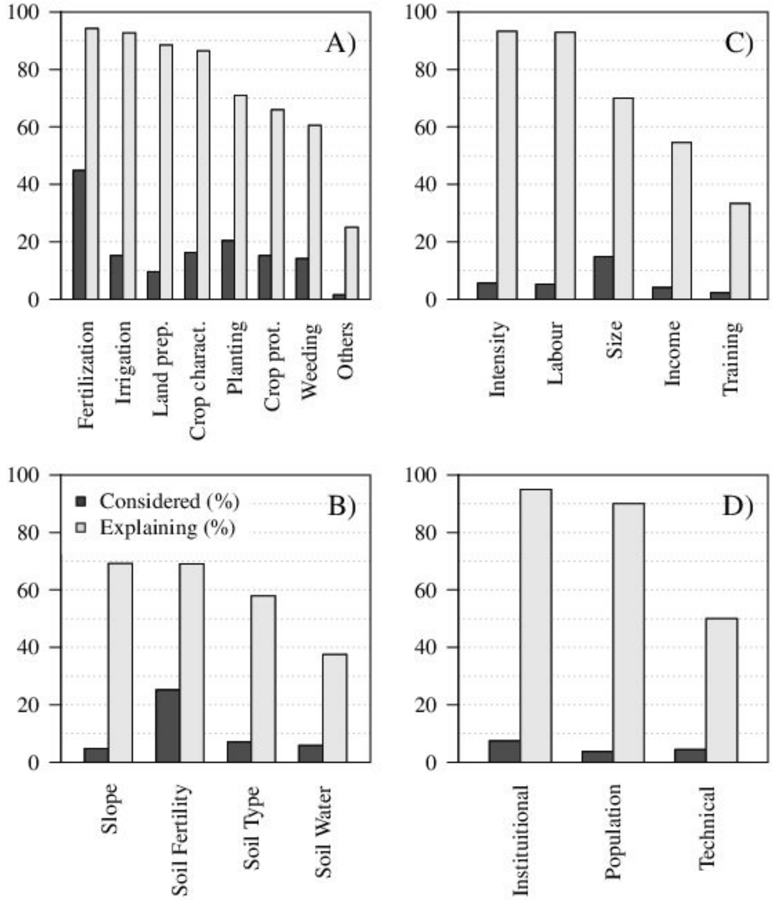

Figure 2.2: Percentage of considered and explaining factors for the different groups within the A) management, B) edaphic, C) farm characteristics and D) socio-economic categories. The different groups are sorted per category in decreasing order of explaining percentage.

in 69% of these records. Soil type explained the yield gap in 58% of the cases while soil water explained the yield gap in 38% of the cases. Compared to management factors, edaphic factors were in general less often considered to explain the yield gap and when considered, they explained yield gap less frequently than management factors (Figures 2.2A and 2.2B). For instance, both soil fertility and fertilization were considered by a relatively large number of records, 25% and 45% respectively but factors in the fertilization group explained the yield gap more often (94%) than factors in the soil fertility group (69%).

Few studies considered farm characteristics when explaining the yield gap, compared to management and edaphic factors (Figure 2.2). However, in 93% of the records where a factor related to intensity (e.g. resource use intensity) or labour (e.g. cost of labour) was considered, it explained the yield gap (Figure 2.2C). Size was the group most often considered in the farm characteristics category and, when considered, factors in this group (e.g. farm area) explained the yield gap in 70% of the records. A maximum of five different farm characteristics were considered by a specific study (Fermont et al., 2009; Haefele et al., 2001) and there were only 16 out of 50 studies which considered farm characteristics to explain the yield gap.

Similarly to farm characteristics, socio-economic factors were not often considered to explain the yield gap either (Figure 2.2D). Compared to other types of factors, the maximum number of socio-economic factors considered by a specific study was also smaller. For example, at most only 3 different socio-economic factors were considered by a single study (Neumann et al., 2010) compared to a maximum of 29 management factors (Delmotte et al., 2011). However, when socio-economic factors were considered, they were often explaining. This is true for socio-economic factors related to population (e.g. rural population density), institutions (e.g. access to fertilizers and credits) and technical assistance (e.g. extension services) as these explained the yield gap in more than 50% of the cases when they were considered (Figure 2.2D).

2.3.3 Within group comparison of considered and explaining factors

As data availability matter as to whether include or exclude specific factors, we analysed 179 unique management factors included in the studies reviewed. An average of three different management factors was considered per record with some studies including none (Kassie et al., 2014; Licker et al., 2010; Neumann et al., 2010; Bhatia et al., 2008) and others including up to 29 (Delmotte et al., 2011). For the most considered and explaining groups, we analysed the factors in detail (Figure 2.3).

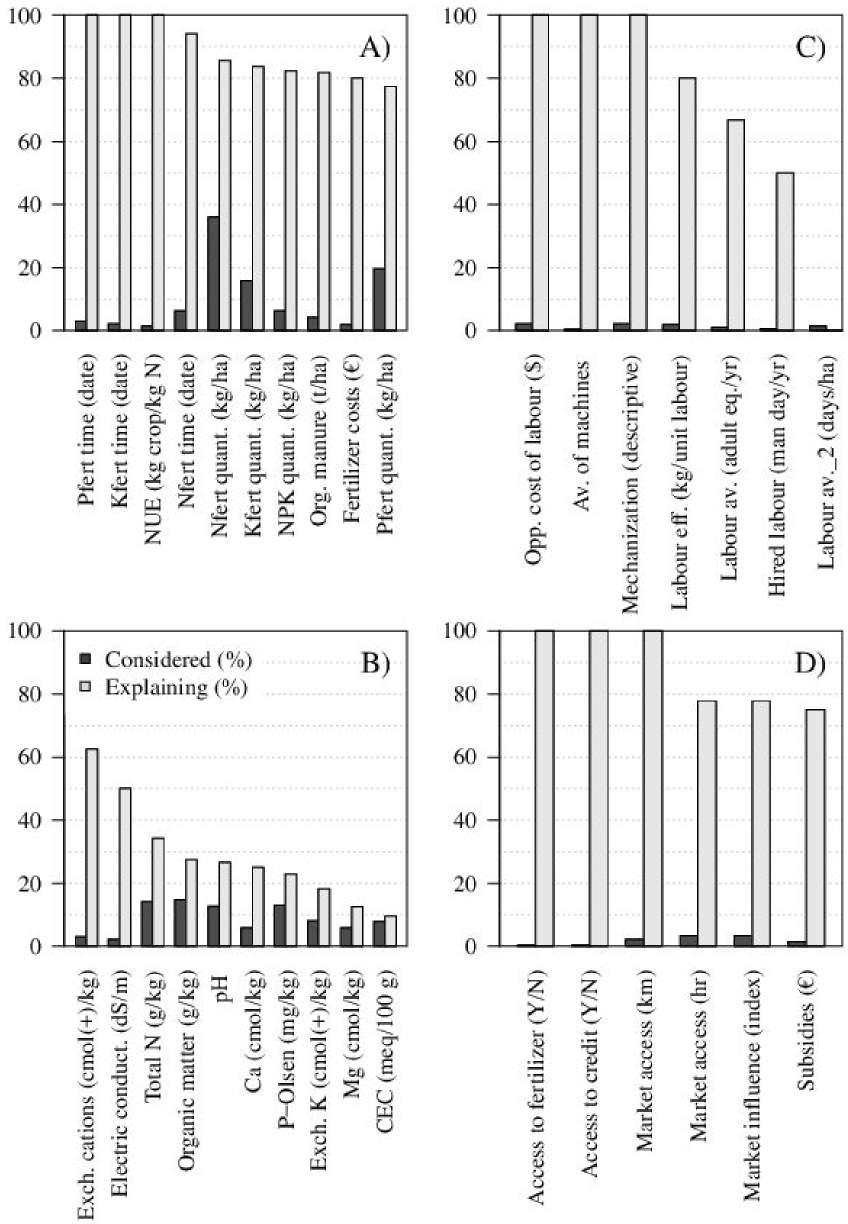

Figure 2.3: Top ten most considered factors in the A) fertilization, B) soil fertility, C) labour and D) institution groups. The different factors are sorted per group in decreasing order of explaining percentage.

Although the quantity of fertilizer used explained the yield gap less frequently than the timing of fertilizer application, data related to fertiliser quantity were more often considered than data related to fertiliser timing (Figure 2.3A). Fertilizer costs were also not often considered, but explained the yield gap in 80% of the cases when considered. The timing of operation was less often considered, but more often explaining, than amounts used in other management groups, such as planting, irrigation and weeding (data provided as online Supplementary Material).

Exchangeable cations and electric conductivity factors of the soil fertility group were not often considered, but when considered, often explained the yield gap (Figure 2.3B). Other factors such as total N, organic matter, pH and P-Olsen were considered by relatively more records, but explained the yield gap less frequently than exchangeable cations and electric conductivity.

Detailed analysis on the labour group of the farm characteristics category showed that opportunity cost of labour, availability of machines and mechanization were the three factors explaining the yield gap most frequently, despite being hardly considered (Figure 2.3C). Besides, labour efficiency and labour availability explained the yield gap in 80% and 67% of the cases, respectively. Looking into the institutional group of the socio-economic category, access to fertilizer, to credit and to markets (km) were the three institutional factors which explained the yield gap the most. Moreover, market access (hrs), market influence and subsides explained the yield gap in more than 75% of the cases when considered (Figure 2.3D).

2.3.4 Spatial differences in considered and explaining factors

Africa

It is evident that factors in the management category were considered more often in Africa than in Asia (Figures 2.4A and 2.4E). In Africa, factors in the fertilization group were often considered (57%) and also often explaining (96%). Moreover, factors related to weeding, planting, crop characteristics and crop protection were often considered but less often explaining, compared to factors in the land preparation category. Factors in the land preparation and irrigation groups were less often considered but explained the yield gap in 89% and 67% of the cases in which they were considered, respectively. With regard to the edaphic category, soil fertility factors often explained the yield gap in Africa compared to factors in the slope, soil water and soil type groups (Figure 2.4B).

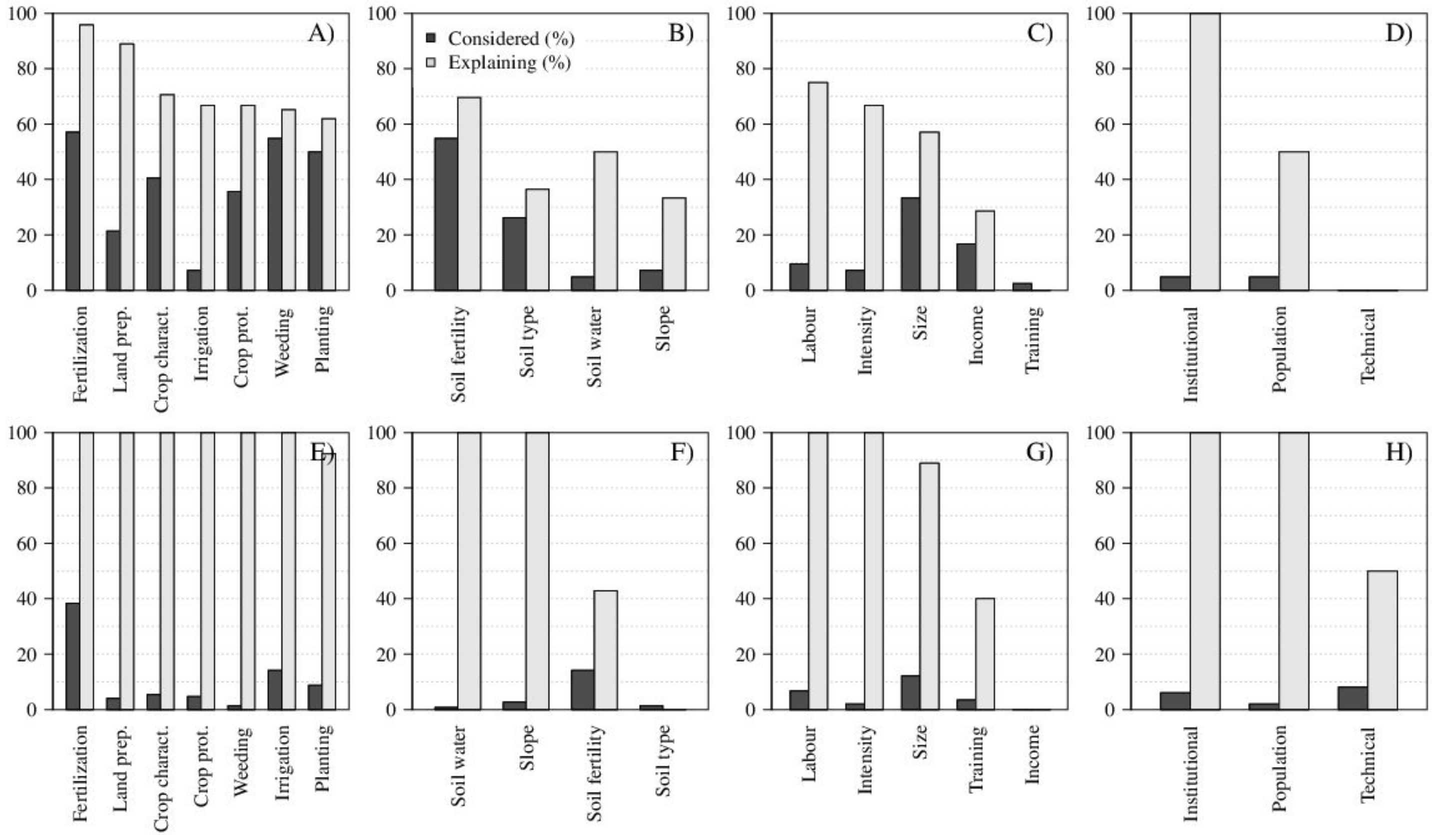

Figure 2.4: Percentage of considered and explaining factors in Africa (upper row, A - D) and Asia (lower row, E - H). The first column shows groups of factors of the management category (A and E), the second column of the edaphic category (B and F), the third column of the farm characteristics category (C and G) and the last column of the socio-economic category (D and H). The different groups are sorted per category in decreasing order of explaining percentage.

The percentage of factors in the management category considered was greater in Africa (Figure 2.4A) than in the global analysis (Figure 2.2A). Fertilization factors were often considered and explaining in Africa and in the global analysis. Slope was the most often explaining edaphic factor in the global analysis (Figure 2.2B), but not in Africa (Figure 2.4B), while soil fertility factors were often considered, and explaining, in Africa and in the global analysis. Size was the most often considered farm characteristic in Africa and in the global analysis, and less often considered factors related to labour and intensity often explained the yield gap in either case (Figures 2.4C and Figure 2.2C). Institutional and population factors also often explained the yield gap in the global analysis and in Africa (Figures 2.2D and 2.4D).

Asia

Similarly to studies in Africa, studies in Asia focused mainly on fertilization and soil fertility factors to explain yield gaps (Figures 2.4E and 2.4F). Soil fertility factors explained the yield gap in 70% of the cases in which they were considered in Africa and only in 43% of the cases in which they were considered in Asia. Fertilization factors had a high explaining power (100%), but this was also true for land preparation, crop characteristics, crop protection, weeding and irrigation factors, which were considered less frequently than fertilization factors. Even though less management factors were considered in Asia than in Africa, they explained the yield gap more often when considered. Soil water and slope were the most important edaphic groups, although they were only considered in less than 5% of the cases. Factors in the labour, intensity and size groups explained the yield gap more often in Asia than in Africa while farm characteristics and socio-economic factors were less considered to explain yield gap in Asia compared to Africa. Overall, a lower number of factors were considered to explain yield gaps in Asia than in Africa.

2.3.5 Crop specific considered and explaining factors

Rice

Rice yield gaps were analysed and/or explained in a total of seven studies in Africa, seven studies in Asia, and two studies in Europe (Table A1, Supplementary Material). These were analysed further without considering two studies on rice yield gaps conducted at global level also included in the database (Figure 2.5).

Factors related to fertilization, land preparation and crop protection often explained the rice yield gap in Africa while factors related to crop characteristics and planting explained the rice yield gap in Europe more frequently (Figures 2.5A and 2.5C).

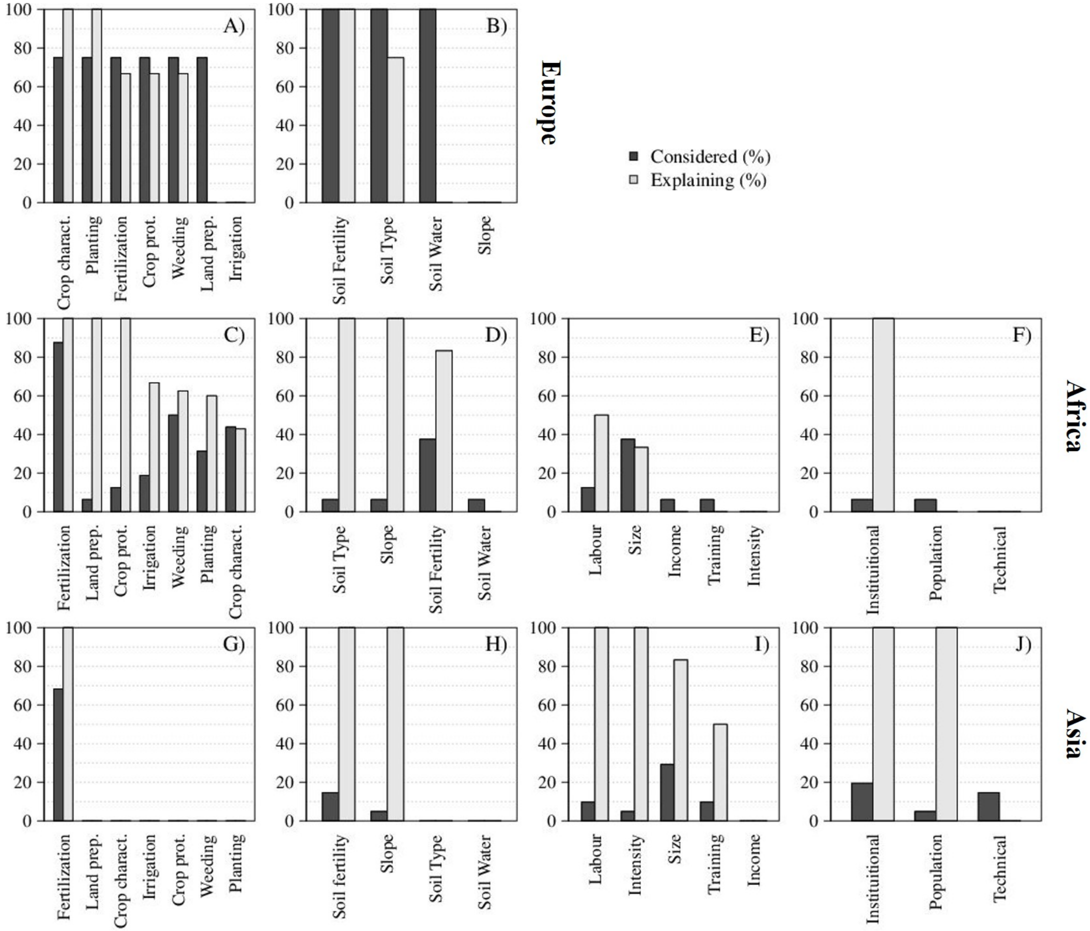

Figure 2.5: Percentage of considered and explaining factors for rice yield gaps in Europe (A - B), Africa (C - F) and Asia (G - J). The first column shows groups of the management category (A, C and G), second column groups of the edaphic category (B, D and H), third column groups of the farm characteristics category (E and I) and fourth column groups of the socio-economic category (F and J). The different groups are sorted per category in decreasing order of explaining percentage.

Land preparation factors were less often considered in Africa than in Europe but often explained the rice yield gap in Africa and never in Europe. Except for land preparation, the management factors considered in Europe and in Africa explained the rice yield gap more often in Europe than in Africa, whereas only fertilisation factors were considered in Asia (Figure 2.5G).

Factors in the crop protection, soil type, slope and soil fertility groups were more important to explain the rice yield gap in Africa (Figures 2.5C and 2.5D) than they were to explain yield gaps in Africa independently of the crop (Figures 2.4A and 2.4B). Conversely, factors related to crop characteristics, soil water and farm characteristics were less important to explain the rice yield gap in Africa (Figures 2.5C - 2.5E) compared to their large contribution to explain yield gaps of other crops in Africa (Figures 2.4A - 2.4C).

Soil fertility was of little importance to explain the yield gap in Asia (Figure 2.4F) but it was important to explain the rice yield gap in this continent (Figure 2.5H). Soil fertility was also always considered, and always explained the rice yield gap, in Europe (Figure 2.5B) while soil type and slope explained the rice yield gap in Africa more often (Figure 2.5D). Soil water often explained yield gaps in Asia but not rice yield gaps in Asia (Figures 2.4F and 2.5H). Conversely, technical factors from the socio-economic category explained the rice yield gap in Asia less often compared to other crops in Asia (Figures 2.5J and 2.4H). Farm characteristics and socio-economic factors were not considered to explain rice yield gaps in Europe.

Maize

A total of four studies conducted yield gap analysis for maize in Africa, six studies in Asia, three studies in North America and one study in Europe and in South America (Table A1, Supplementary Material). Two studies analysed maize yield gaps at global level as well. The analysis of the explaining factors for maize yield gaps were analysed for Africa, Asia and North America only (Figure 2.6) due to the larger number of observations.

Factors in the fertilization, crop characteristics and planting groups explained the maize yield gap in Africa every time they were considered (Figure 2.6C). Crop protection factors were less important to explain maize yield gaps compared to rice yield gaps (Figures 2.6C and 2.5C). Land preparation factors explained the yield gap in 89% of the studies performed in Africa (Figure 2.4A), but were not considered in maize yield gap analysis (Figure 2.6C). Soil fertility was the most considered

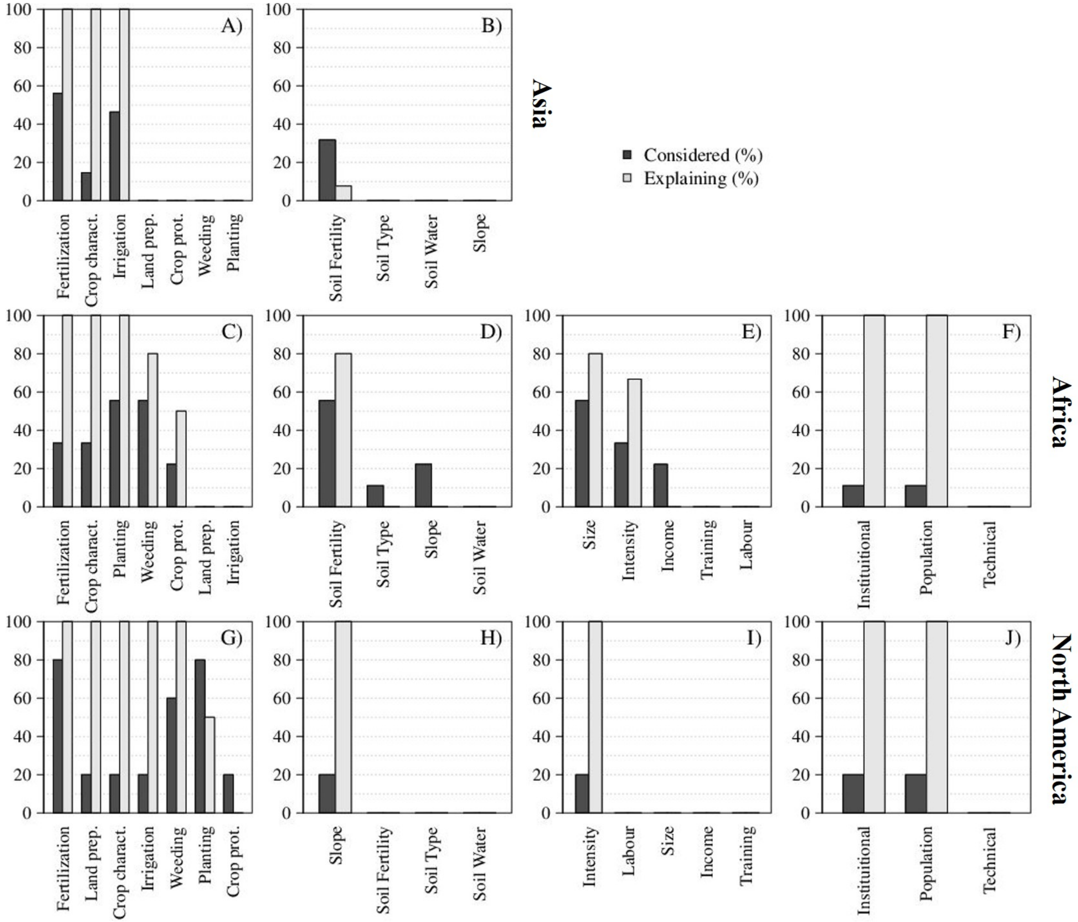

Figure 2.6: Percentage of considered and explaining factors for maize yield gaps in Asia (A - B), Africa (C - F) and North America (G - J). The first column shows groups of the management category (A, C and G), second column groups of the edaphic category (B, D and H), third column groups of the farm characteristics category (E and I) and fourth column groups of the socio-economic category (F and J). The different groups are sorted per category in decreasing order of explaining percentage.

edaphic group and explained the maize yield gap in Africa in 80% of the cases it was considered (Figure 2.6D), whereas soil type and slope were the most important edaphic groups for rice yield gaps (Figure 2.5D). Size and intensity explained the maize yield gap in more than 60% of the cases and factors in the labour group were not considered for maize yield gap analysis in Africa (Figure 2.6E). Although socio-economic factors were considered less often, institutional and population groups always explained the maize yield gap in Africa when considered (Figure 2.6F).

For maize yield gap analysis in Asia, only factors from the management and edaphic categories were considered (Figures 2.6A and 2.6B). Even from the management category, only fertilization, crop characteristics and irrigation groups were considered. Factors in the fertilization and irrigation groups were most considered and always explained maize yield gaps in Asia. Even though crop characteristics were considered less often than factors in fertilization and irrigation groups, they also explained maize yield gaps in all cases they were considered. Factors in the soil fertility group were the only edaphic factors considered and explained maize yield gap analysis in Asia in less than 10% of the cases they were considered (Figure 2.6B).

Factors in the fertilization, weeding and planting groups were most often considered to explain the maize yield gap in North America (Figure 2.6G). Apart from the planting and crop protection groups, all groups of factors in the management category always explained the maize yield gap when considered. Factors in the slope group were considered in 20% of the cases and always explained maize yield gaps in North America (Figure 2.6H). Farm characteristics from the intensity group and socio-economic factors from the institutional and population groups also always explained the maize yield gap when they were considered by studies conducted in North America (Figure 2.6J).

Wheat

A total of four studies conducted yield gap analysis for wheat in Asia, three studies in Europe and two studies in Australia and North America and one study in South America (Table A1, Supplementary Material). There was also one study conducted in South America and two studies conducted at global level but these were not included in this analysis (Figure 2.7).

Factors in the planting group were considered more frequently for wheat yield gap analysis in Asia compared to factors of other groups (Figure 2.7H). Although factors in the fertilization, land preparation, irrigation and crop protection groups were considered less often than factors in the planting group, all these groups always explained the wheat yield gap when considered. Factors related to crop characteristics

and weeding were not considered for explaining the wheat yield gap in Asia. In general, management factors were considered more frequently for wheat yield gap analysis in Asia than for the global analysis (Figure 2.4E) or for maize yield gap analysis in Asia (Figure 2.6A). Slope was the only edaphic group considered and always explained the wheat yield gap in Asia (Figure 2.7I). Although factors in the labour and size groups were considered more often than factors in the intensity group, they all always explained wheat yield gaps in Asia when considered (Figure 2.7J). Factors related to labour were considered more frequently by yield gap studies for wheat in Asia than for other crops (Figures 2.7J and 2.4G). From the socio-economic category, factors in the technical group were considered and explained the wheat yield gap in Asia more often than for yield gaps in Asia (Figures 2.7K and 2.4H). Factors from institutional and population groups were considered less often than factors in the technical group but all always explained the wheat yield gap in Asia.

Studies in Europe only considered factors of the management category for wheat yield gap analysis (Figure 2.7A). Factors of the fertilization and crop characteristics groups were most often considered while factors from land preparation, weeding and irrigation groups were not considered for wheat yield gap analysis in Europe. Factors in the fertilization, crop protection and planting groups, but not crop characteristics, always explained wheat yield gaps in Europe when they were considered.

Factors in the land preparation group were the least considered to explain wheat yield gaps in North America but they always explained the wheat yield gap in this region when considered (Figure 2.7D). Factors in the irrigation group were the most considered for wheat yield gap analysis in North America and explained the yield gap in 86% of the cases. Factors related to fertilization, planting and crop protection were slightly less important to explain wheat yield gaps in North America compared to land preparation and irrigation. Factors in the fertilization, planting and crop protection groups were equally considered (71% of the cases) but fertilization and planting factors explained the yield gap in 60% of the cases, while crop protection factors in only 20% of the cases. In addition, factors from the intensity group or from the institutional and population groups were not often considered but when considered, they always explained the wheat yield gap in North America (Figures 2.7F and 2.7G). For South America, factors in the fertilization and soil fertility groups explained the wheat yield gap more often compared to factors from other groups (data not shown).

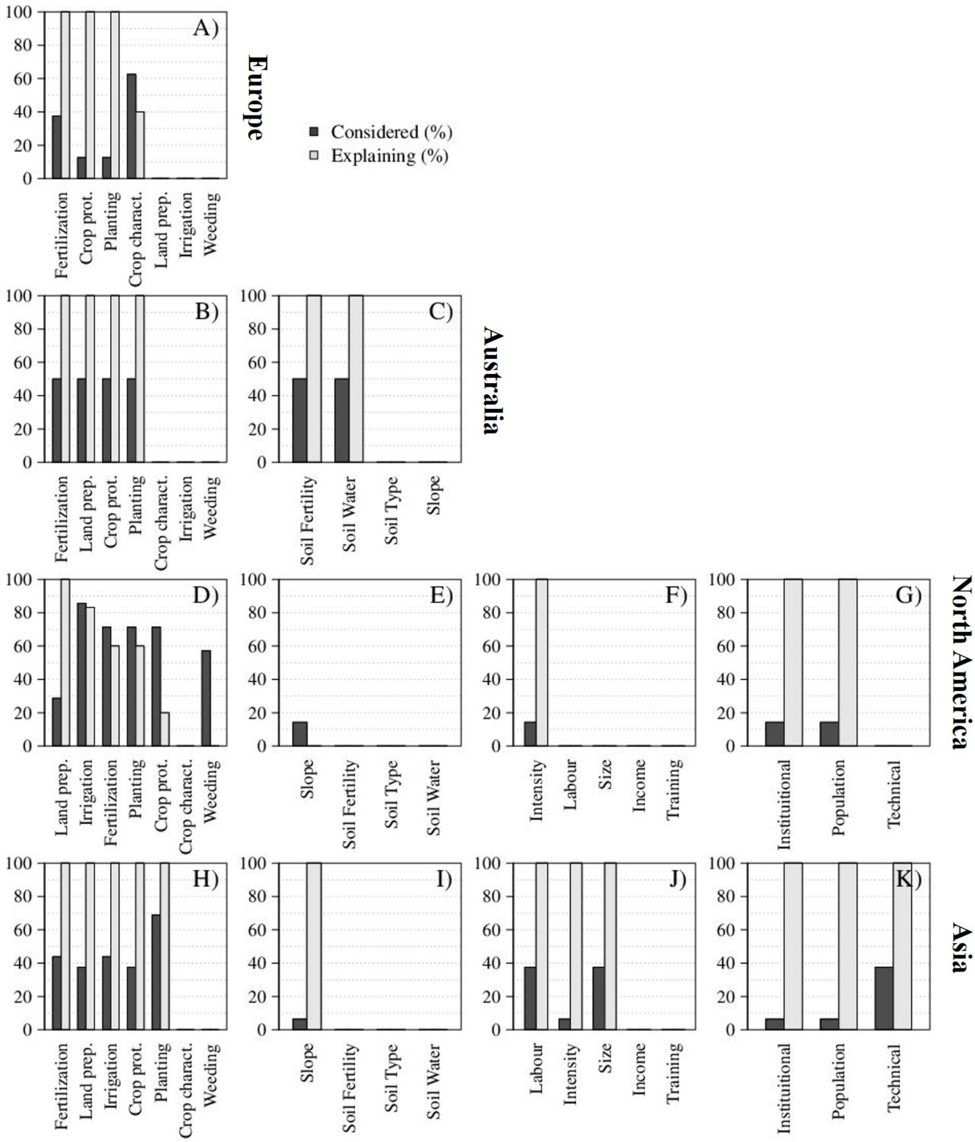

Figure 2.7: Percentage of considered and explaining factors for wheat yield gaps in Europe (A), Australia (B and C), North America (D - G) and Asia (H - K). The first column shows groups of the management category (A, B, D and H), second column groups of the edaphic category (C, E and I), third column groups of the farm characteristics category (F and J) and fourth column groups of the socio-economic category (G and K). The different groups are sorted per category in decreasing order of explaining percentage.

For Australia, factors in the fertilization, land preparation, crop protection and planting groups were considered in 50% of the cases and always explained the wheat yield gap when they were considered (Figure 2.7B). Irrigation, crop characteristics and weeding factors were not considered for wheat yield gap analysis in Australia. From the edaphic category, factors related to soil fertility and soil water also always explained the wheat yield gap when considered while factors related to soil type and slope were never considered. Farm characteristics and socio-economic factors were also never considered for wheat yield gap analysis in Australia.

2.4 Discussion and opportunities for alternative data collection approaches

2.4.1 Methodological approach

The main objective of this paper was to review the yield gap explaining factors identified by previous studies, in order to assess data availability and suggest improved data collection approaches for yield gap analysis. A review was performed based on peer-reviewed journal articles, using the keywords ’yield gap’, ’potential yield’, ’yield variability’, ’water-limited yield’ and ’yield gap variability’. We acknowledge that this terminology is used by a specific group of researchers, and that many studies may be available on similar work which were not included in this analysis. For example, economists use production functions and frontier analysis to explain yields (e.g. Monchuk et al., 2010; Helfand and Levine, 2004; Giannakas et al., 2001; Ali, 1995) and few of such studies were captured by this analysis. Such type of studies focus more on farm(er) characteristics and socio-economic factors, and therefore including those would have increased the percentage of considered factors on these categories.

Our analyses relied on data from a variety of yield gap studies conducted at different scales and this affects the relative importance of the different factors. This is also true for benchmark yields, which differed across different types of studies. As an example, factors related to soil type are important to explain the yield gap between actual and potential yields but not between actual and water-limited yields because they have to be considered in the estimation of the latter. Thus, the relevance of factors related to soil type is smaller when explaining the yield gap between actual and water-limited yields. During the review we also learnt that there is no consistent procedure to report the importance of explaining factors and for that reason we could only compile factors which explained the yield gap based on the criteria used by the individual papers (e.g. significance). A more consistent procedure is required to report the relative importance of explaining factors in future yield gap analyses. Despite these limitations, this review provides a thorough overview of the factors considered and explaining yield gaps for different crops and for different regions across the globe, and therefore it provides a good impression of gaps in required data, and a basis for improved data collection.

2.4.2 Important factors for yield gap analysis

The global analysis showed that factors from the management category were most often considered to explain the yield gap or yield variability compared to the factors in the edaphic, farm characteristics and socio-economic categories. Among the management category, the fertilization group was considered most frequently and explained the yield gap in a large number of cases. However, factors from the irrigation, land preparation and crop characteristics groups were considered less frequently than factors related to fertilization but explained the yield gap in more than 80% of the cases when considered. The importance of fertilization and irrigation factors to explain existing yield gaps at the global level was also indicated by Mueller et al. (2012). Planting, crop protection and weeding were the other groups which explained the yield gap in more than 60% of the cases when considered. Overall, this highlights the importance of crop management in its broadest sense for existing yield gaps.

The global and detailed analysis showed that factors in the edaphic category explained yield gaps less frequently compared to factors in the management, farm characteristics and socio-economic categories (Figures 2.2B and 2.3B). This can be explained by the improved access to technology and resources to solve soil limitations. Farm characteristics and socio-economic factors appeared to be of high relevance to explain yield gaps, but were only considered by few studies. For example, factors related to institutions (e.g. access to fertilizers, credit and market) and labour (e.g. opportunity cost of labour, access to machineries and mechanisation) were relevant to explain the yield gap (Figures 2.3C and 2.3D). The majority of the studies included in this review are agronomic studies and this certainly explains the lower number of factors considered from these two categories.

Our detailed analysis showed that timing of fertilization, irrigation and weeding is less often considered compared to the quantity of inputs used. However, factors related to the timing of crop management operations explained the yield gap more often than factors associated with the quantity of inputs used. The study of Lobell et al. (2005) noted that the timing of irrigation was more important than the number of irrigations, suggesting that the efficient use of water is more important than the total amount of water applied for wheat yields in the Yaqui Valley, Mexico. The issue on timing is important because it stresses that narrowing yield gaps requires the right amount of inputs in the right moment of time. These two aspects of crop management are related to e.g. labour and cash availability at farm level (Chadwick et al., 2015; Gianessi, 2013). The availability of family labour further depends on e.g. household composition and opportunity costs of labour off-farm while hired labour is related to cash availability and market conditions.

The importance of specific factors also depends on the location and on the crop studied, meaning that generalizations should be interpreted with caution. For example, soil fertility was relevant to explain yield gaps in Africa whereas soil water was more relevant to explain yield gaps in Asia (Figures 2.4B and 2.4F). The importance of soil fertility factors to explain yield gaps in Africa was explicitly indicated by several studies (e.g. Affholder et al., 2012; Okumu et al., 2011; Wairegi et al., 2010; Fermont et al., 2009; Tittonell et al., 2008b). Factors related to fertilization, land preparation and crop protection explained rice yield gaps in Africa most often (Figure 2.5C) while crop characteristics and planting were important to explain rice yield gaps in Europe (Figure 2.5A). The latter may be explained by the availability of agricultural inputs (e.g. fertilizers) and machinery in Europe and lack of these resources in many African countries. In addition, explaining factors can be considered in many different ways (e.g. kg N ha^{-1} or fertilization (Y/N), see online Supplementary Material). This means that the way data is collected can also affect the results of the yield gap analysis, as focusing on few factors may bias the results.

2.4.3 Selection of factors for explaining yield gap

We showed that many factors can explain the yield gap, and ideally these should all be considered in a study. However, the number of factors considered by a single study is limited in practice and it is likely to be affected by 1) the objectives of the study, 2) *a priori* knowledge on possible explaining factors, 3) the method used to explain yield gaps and 4) data availability. Below, we will discuss first the first three reasons, and then go in further depth on data availability, and possibilities to improve this aspect in the future.

Objective, previous knowledge and methods used

First, the objective of a study can limit the number of factors considered in the analysis. For example, if the objective of a study is to assess the effect of soil quality problems or iron toxicity on yield, then authors tend to focus on those types of factors in their analyses (e.g. Audebert and Fofana, 2009; van Asten et al., 2003). Moreover, and as previously mentioned, most of the studies analysing yield gaps are of agronomic nature, and therefore focused more on management and edaphic factors compared to farm characteristics and socio-economic factors.

Second, knowledge on possible explaining factors may guide the selection of considered factors for the analysis. Soil fertility explained the yield gap in Africa more often than in Asia (Figures 2.4B and 2.4F) possibly because previous studies have highlighted the importance of soil fertility in Africa, which could influence subsequent studies in this continent to consider more soil fertility factors than other factors (e.g. labour). This is also true in other regions. For example, a study by Anderson et al. (2005) indicated that in Western Australia yields are more constrained by edaphic factors than by management factors. Following these findings, the focus of the study of Oliver and Robertson (2013) was to relate yield gaps to spatial variation in soil properties that are known to limit yield in a water-limited environment.

Third, the method used to explain the yield gap also has an influence on the number and type of factors considered. Table A2 (Supplementary Material) provide an overview of the different methods used to explain yield gaps and of factors used by each method. It is notable that methods like Classification And Regression Tree (CART) analysis and qualitative methods make use of many and different types of factors to explain the yield gap (e.g. Tittonell et al., 2008a), while other methods like crop models make use of a limited number of factors to explain the yield gap (e.g. Abeledo et al., 2008).

Data availability

Data availability is also a limiting factor regarding the number and types of factors used for yield gap analysis in addition to the objective of the study, previous knowledge of explaining factors, and method used. To mention few examples, the study of Neumann et al. (2010) indicated that fertilizer application, one of the most important management strategies to increase actual yields, could not be included in their global analysis due to lack of appropriate data and hence, the yield gap attributed to fertilizer application could not be identified in their study. Lu and Fan (2013) also mentioned that due to lack of data, the EPIC model, which was used in their study to assess yield gaps in the North China Plain, was validated using data from only two experimental stations. The authors indicated model validation with data from two to four other sites would be worthwhile to provide stronger evidence of the utility of the model for yield gap analysis in this region.

Figure 2.8 shows the main data sources for the factors considered for yield gap analysis by the different studies reviewed. It is evident that around 47% of the management factors were from field trials, followed by farm surveys which contributed with 19% of the management factors considered for yield gap analysis. Compiled databases and a combination of measurements and surveys contributed to 12% and 11% of the man-

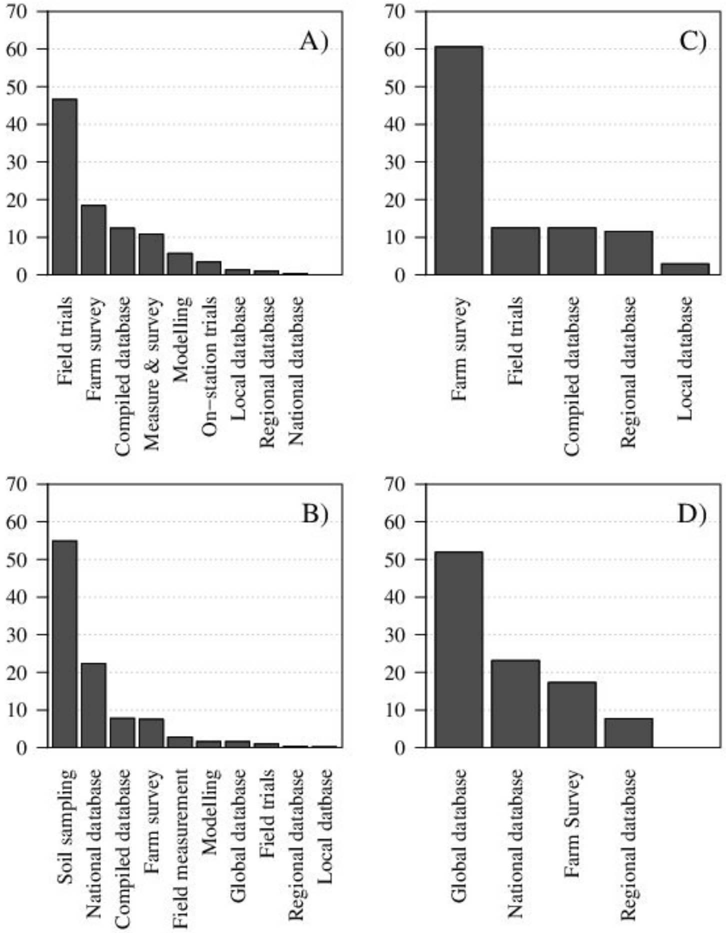

Figure 2.8: Overview of the data sources for factors considered for yield gap analysis per category: A) crop management, B) edaphic conditions, C) farm characteristics and D) socio-economic conditions.

agement factors considered, respectively. Soil sampling was the main data source for edaphic factors, with 55% of the cases, followed by national databases, which contributed with 22% of the edaphic factors considered for yield gap analysis. Farm surveys and global databases were the main data sources for farm characteristics and socio-economic factors, respectively. The fact that the majority of the management data were from field trials, not from real farms, deserves special attention. Data from on-station field trials are often not representative of real farms in terms of soil properties and crop management, particularly in sub-Saharan Africa (Vanlauwe et al., 2016; Tittonell and Giller, 2013). Moreover, it is important to understand what are the actual farm level constraints and technical problems farmers face and how these explain existing yield gaps. One of the reasons not to consider data from real farms for yield gap analysis might be because of the lack of datasets containing detailed field-specific information for a large number of individual farms. Thus, how can we collect more of these relevant factors for yield gap analysis? The next section provides an overview of alternative data collection approaches that can potentially be used to complement traditional methods (e.g. farm survey) in the compilation of some of the relevant field and farm level factors for yield gap analysis.

2.4.4 Opportunities for alternative data collection approaches

Most factors related to management, farm characteristics and socio-economic conditions (e.g. timing of fertilization, education level/age of the farmer, access to fertilizers) can only be obtained by asking farmers directly using farm surveys (e.g. Tittonell et al., 2008a) or through farmers' self-reporting (e.g. Yield Prophet: www.yieldprophet.com.au; Hochman et al., 2012). Other data collection methods, e.g. remote sensing can provide data on e.g. soil organic matter content (Gomez et al., 2008), field size calculated from high spatial resolution satellite images (Schulthess et al., 2012), and crop density (Bai et al., 2011; Thorp et al., 2008). Sensor networks can also provide information on crop canopy (e.g. using upward-pointing digital cameras; Ryu et al., 2012) and soil moisture content (e.g. using DACOM sensor). However, remote sensing measurements are still highly uncertain for heterogeneous and small fields that are common in smallholder agriculture in the tropics (Lobell, 2013) and their potential to provide data for yield gap analysis is also constrained by the availability of satellite images at the required spatial and temporal resolutions.

Most developed countries have a well-developed and organised infrastructure to collect and harmonise agricultural data that are used for different purposes (Paustian, 2013), including yield gap analysis. For example, the Farm Accountancy Data Network (FADN) of the European Commission collects farm characteristics (e.g. crop areas, labour force) and socio-economic data (e.g. subsidies) at farm level, which were used to e.g. assess regional maize yield gaps in Europe (Reidsma et al., 2009a). However, even these well-established and organised databases lack some important factors (Paustian, 2013) for yield gap analysis (e.g. soil and crop-specific management data) and additionally, these are available only as aggregated averages at farm level lack detail for in-depth agronomic analyses (e.g. amount of fertilizer applied). In most developing countries detailed agricultural data are much less harmonised and scarce due to lack of resources to conduct extensive and detailed surveys on agricultural practices (Paustian, 2013). Data compiled in databases at national level (e.g. FAOSTAT) provide a useful first estimate of agricultural activities (Paustian, 2013) that can be used in national and global yield gap analysis, but are inadequate for detailed yield gap analysis (Grassini et al., 2015; van Ittersum et al., 2013). In this context, one possible way to overcome this critical data gap is by 'letting the farmers tell us themselves' (Paustian, 2013).

The proliferation and widespread use of mobile devices, which are equipped with sensors (e.g. GPS) and other related technologies, makes it possible to implement bottom-up data collection approaches like crowdsourcing with which relatively large amounts of information can be directly obtained from local communities (Herrick et al., 2013). Currently, applications are mostly focused on delivering market information so that farmers can make informed decisions on when and where to sell their products (Muto and Yamano, 2009). Provision of information, such as management recommendations and weather forecasts is another area of development where the use of mobile phones played a major role (Aker, 2011). As also suggested by Paustian (2013), the experiences and lessons learnt from aforementioned initiatives can be used to collect more of the relevant factors for yield gap analysis.

Table 2.1 provides an overview of the opportunities and limitations of three types of alternative data collection methods to acquire factors relevant for yield gap analysis: 1) crowdsourcing (CS); 2) remote sensing (RS); and 3) sensor networks (SN). While describing the different data collection methods, special attention is given to the crowdsourcing approach due to its great potential for collecting detailed individual farm data (Table 2.1).

Crowdsourcing is an emerging method for data collection since the advent of widespread access to mobile phones (Belden et al., 2013). It can be used to collect different information related to the timing of an activity (e.g. timing of fertilization, weeding, irrigation) and the quantity of inputs used (e.g. fertilizer applied, number of

Table 2.1: Alternative data collection methods based on crowdsourcing (CS), remote sensing (RS) and sensor networks (SN), and their potential to collect factors relevant for yield gap analysis. Codes: 'It is possible and has already been used' (+), 'Potentially possible to use' (+/−), 'It is not possible/ it has not been used' (−).

Factors	CS	RS	SN	Examples/Reference
Crop yield				
Actual farm yield	+	+	−	**CS**: Farmers self-report the amount of yield they harvested (Hochman et al., 2012; Gittleman et al., 2012). **RS**: (Lobell et al., 2005).
Benchmarking yield	+/−	+	−	**CS**: Asking farmers to provide the maximum yield they have harvested in a specific field for the last few growing seasons. **RS**: to use the maximum yield within the remote sensing estimates as a proxy for the benchmarking yield (Lobell et al., 2002).
Management				
Fertilization	+/−	−	−	**CS**: Timing of fertilization, quantity of fertilizer(s) applied and fertilizer costs.
Planting	+/−	+	−	**CS**: Sowing date, number of plants m^2, seeding method and intercropping (Y/N). **RS**: Crop density (Bai et al., 2011; Thorp et al., 2008), sowing date (Ortiz-Monasterio and Lobell, 2007; Lobell et al., 2003) and intercropping (Jain et al., 2013).
Crop characteristics	+	+	+	**CS**: LAI app (Confalonieri et al., 2013), lodging (Y/N), physiological maturity (using Growing Degree Days app), harvesting time, crop genotype (name of variety planted), and dates of flowering and maturity. **RS**: Canopy cover percentage (Pacheco et al., 2008; DeTar and Penner, 2007), lodging (Zhang et al., 2014). **SN**: LAI (Ryu et al., 2012).
Irrigation	+	+	+	**CS**: Timing of irrigation (Smart ICT- Africa project), irrigation infrastructure, irrigation system, supplementary irrigation (Y/N), water control (good/bad) and number of irrigations. **RS**: Irrigation amount (Droogers et al., 2010). **SN**: DACOM sensor.

Table 2.1: Alternative data collection methods based on crowdsourcing (CS), remote sensing (RS) and sensor networks (SN), and their potential to collect factors relevant for yield gap analysis. Codes: 'It is possible and has already been used' (+), 'Potentially possible to use' (+/−), 'It is not possible/ it has not been used' (−). (*continued*)

Factors	CS	RS	SN	Examples/Reference
Weeding	+/−	+	+	**CS**: Timing of weeding/herbicide application, number of weeding/herbicide applications, weeding method and weed management (score). Moreover, using apps to identify weeds and assess weed pressure (Rahman et al., 2015). **RS**: Weed management (Goel et al., 2003). **SN**: Weed intensity (Sui et al., 2008).
Crop protection	+	+	+/−	**CS**: Digital Early Warning Network, PestNet, Plant clinics of CABI. **RS**: Disease detection (Cao et al., 2013).
Land preparation	+/−	+	−	**CS**: Tillage system, area per crop (%), land levelling (#), fallow reside management and crop residue management. **RS**: Crop residue and tillage practices (Zheng et al., 2014).
Edaphic				
Soil fertility	+	+	+	**CS**: Farmers diagnose soil constraints in the field using soil testing kits (SoilDoc) and transmit the information through SMS and using MySoil app to provide information about pH and organic-matter content of the soil (Shelley et al., 2013). **RS**: Organic matter content of the soil (Gomez et al., 2008). **SN**: SoilCares initiative - Mobile Lab and SoilCares handheld scanner.
Soil water	+	+	+	**CS**: Soil Water app for smartphones (SWApp). **RS**: Soil moisture (Petropoulos et al., 2015). **SN**: Soil moisture sensors (Xiao et al., 2013).
Soil type	+	+	−	**CS**: Using SoilWeb app to provide the soil type of the current location of the phone. **RS**: Soil type identification (Jiji and Nadar, 2016).
Slope	+	+	−	**CS**: Using LandInfoapp to assess slope. **RS**: Using satellite images for Digital Elevation Model extraction (Yanalak et al., 2012).

Table 2.1: Alternative data collection methods based on crowdsourcing (CS), remote sensing (RS) and sensor networks (SN), and their potential to collect factors relevant for yield gap analysis. Codes: 'It is possible and has already been used' (+), 'Potentially possible to use' (+/−), 'It is not possible/ it has not been used' (−). (*continued*)

Factors	CS	RS	SN	Examples/Reference
Farm characteristics				
Labour	+/−	−	−	**CS**: Labour availability, cost of labour, availability of machines and mechanisation.
Income	+/−	−	−	**CS**: Crop income, farm income, household income and production costs.
Training	+/−	−	−	**CS**: Years in school, farming experience (years) and farmer age.
Size	+	+	−	**CS**: Field size (Fritz et al., 2015). **RS**: Field size (Yan and Roy, 2014).
Intensity	+/−	+/−	−	**CS**: Resource use intensity (score), irrigated area (%), irrigated area per grain type (ha).
Socio-economic				
Institutional	+/−	−	−	**CS**: Access to fertilizer (Y/N), access to credit (Y/N) and market access (hrs or km).
Technical	+/−	−	−	**CS**: Technical assistance (#) and extension contacts (score).
Population	+/−	−	−	**CS**: Gender (M/F)

weeding and irrigation operations) in a specific field. Cropping calendar (e.g. sowing date, dates of flowering, maturity and harvest) are other potential management factors that can be collected using this approach. With the ubiquitous availability of smartphones which are equipped with sensors (e.g. GPS), the geo-location (boundary) of a field can also be collected and used as input to calculate accurate field sizes. The camera feature included in many phones can be used to visually record some incidents in a field (e.g. incidence of pest, disease or weed) and these can be used in a later stage by experts to assess the extent of the damage and to identify the type of pest, disease and/or weed that caused the damage. Providing training for selected community members (e.g. focal farmers) to be able to identify pest and/or disease as done by Plant clinics of CABI (www.plantwise.org/plant-clinics/) can help the farmers to get assistance and provide the right information.

Collecting soil fertility data using the crowdsourcing approach might not be as straight forward as collecting management data. However, asking farmers to assess the fertility level of their soils using their own local indicators can potentially be used to complement soil chemical analysis (Desbiez et al., 2004). In addition, an on-farm soil testing kit which allows farmers to diagnose soil constraints in the field and transmit the information quickly through SMS (e.g. SoilDoc) could potentially be used for acquiring soil fertility data. For smartphone/tablet users, applications like 'MySoil' can be used to provide information about pH and organic-matter content of the soil (Shelley et al., 2013).

The crowdsourcing approach also offers opportunities to collect farm characteristics and socio-economic data. Farm(er) characteristics related labour (e.g. labour availability), training (e.g. years in school) and income (e.g. farm income) can be collected using the crowdsourcing approach. Socio-economic data including access to fertilizer (Y/N), access to credit (Y/N), number of technical assistances received and gender of a farmer are some examples of information that can be collected using crowdsourcing. However, to receive accurate and timely information, understanding the motivations of the community (farmers) to participate in crowdsourcing and their incentives to provide the requested information is critical (Roy et al., 2012). Moreover, identifying the right technology (platform) for the farmers to use in the crowdsourcing activity is another important step that needs to be considered while designing a crowdsourcing campaign.

2.5 Conclusion

Many different factors have been considered for yield gap analysis in the studies reviewed. The selection of factors considered by the different studies is influenced by 1) the objectives of the study, 2) knowledge on possible explaining factors, 3) the method used to explain yield gaps and 4) data availability.

Our results show that management factors were considered for yield gap analysis more frequently than edaphic, farm characteristics and socio-economic factors. Although fertilization related factors seem to be considered most often, other management factors like land preparation, irrigation and crop characteristics also often explained the yield gap when considered. The same is true for farm characteristics and socio-economic factors, meaning that future yield gap studies should consider farm characteristics and socio-economic factors as well. Information related to quantity used (e.g. N fertilizer quantity, irrigation amount and number of weeding operations) was more often considered in yield gap analysis than the timing of the different management operations (e.g. N fertilizer timing, irrigation timing and timing of weeding). However, the latter explained the yield gap often. It is important that data about the timing of management operations is also collected and taken into account in future yield gap analysis. The relative importance of different factors to explain yield gaps is location- and crop-specific, meaning that generalizations should be made with caution. Moreover, this requires that approaches for data collection are also location- and/or crop-specific.

Data availability can be increased using bottom-up data collection approaches like crowdsourcing, which might help to collect more of the explanatory factors required for yield gap analysis. This is especially true for management, farm characteristics and socio-economic factors. Crowdsourcing based methods (e.g. farmers send timing information via SMS) is also a promising alternative to acquire real-time information about timing of management operations.

2.6 Acknowledgements

We thank ir. Marcel Lubbers for developing the Excel macros used for data analysis and the corresponding authors of the reviewed studies who replied to our query. This research was carried out as part of the Wageningen UR strategic programme 'Mapping for Sustainable Intensification', 2012 - 2016, funded by Wageningen University and Research.

2.7 Appendix

Selected supplementary tables are provided as Supplementary Material of this thesis. The full supplementary material of the manuscript can be assessed with the online version of the manuscript (https://doi.org/10.1016/j.eja.2016.06.016). The database is available upon request from the first or second author of the manuscript.

CHAPTER 3

Explaining rice yields and yield gaps in Central Luzon, Philippines: An application of stochastic frontier analysis and crop modelling

This chapter is published as:
Silva, J.V.; Reidsma, P.; Laborte, A.G.; van Ittersum, M.K. 2017. Explaining rice yields and yield gaps in Central Luzon, Philippines: An application of stochastic frontier analysis and crop modelling. *European Journal of Agronomy*. 82, Part B, 223 - 241.

Abstract

Explaining yield gaps is crucial to understand the main technical constraints faced by farmers to increase land productivity. The objectives of this study are to decompose the yield gap into efficiency, resource and technology yield gaps for irrigated lowland rice-based farming systems in Central Luzon, Philippines, and to explain those yield gaps using information related to crop management practices, biophysical constraints and available technologies. Stochastic frontier analysis was used to quantify and explain the efficiency and resource yield gaps and a crop growth model (ORYZA v3) was used to compute the technology yield gap. We combined these two methodologies into a theoretical framework to explain rice yield gaps in farmers' fields included in the Central Luzon Loop Survey, an unbalanced panel dataset of about 100 households, collected every four to five years during the period 1966 - 2012. The mean yield gap estimated for the period 1979 - 2012 was 3.2 t ha^{-1} in the wet season (WS) and 4.8 t ha^{-1} in the dry season (DS). An average efficiency yield gap of 1.3 t ha^{-1} was estimated and partly explained by untimely application of mineral fertilisers and plant protection agents. The mean resource yield gap was small in both seasons but somewhat larger in the DS (1.3 t ha^{-1}) than in the WS (1.0 t ha^{-1}). This can be partly explained by the greater N, P and K use in the highest yielding fields than in lowest yielding fields which was observed in the DS but not in the WS. The technology yield gap was on average less than 1.0 t ha^{-1} during the WS prior to 2003 and ca. 1.6 t ha^{-1} from 2003 - 2012 while in the DS it has been consistently large with a mean of 2.2 t ha^{-1}. Varietal shift and sub-optimal application of inputs (e.g quantity of irrigation water and N) are the most plausible explanations for this yield gap during the WS and DS, respectively. We conclude that the technology yield gap explains nearly half of the difference between potential and actual yields while the efficiency and resource yield gaps explain each a quarter of that difference in the DS. As for the WS, particular attention should be given to the efficiency yield gap which, although decreasing with time, still accounted for nearly 40% of the overall yield gap.

3.1 Introduction

Agronomists and agricultural economists have developed different concepts and quantitative methods to estimate and explain yield gaps, i.e. the difference between climatic potential and actual farmers' yields. Agronomic studies traditionally rely on field experiments (e.g. Affholder et al., 2012) and/or crop growth models (e.g. Angulo et al., 2012) to assess the contribution of different management practices to crop yield following concepts of production ecology (van Ittersum and Rabbinge, 1997). The main limitation of these types of studies is that these do not explicitly take into account farmers objectives and constraints (and other socio-economic conditions) because they are usually performed at field and regional levels. On the other hand, production economics deals with the estimation and interpretation of technical and allocative efficiencies using farm level data. Technical efficiency can be defined as the maximum output that can be achieved for a specific level of inputs while allocative efficiency refers to the success of a farm in choosing the optimal proportion of inputs given a pre-defined objective and set of constraints (Farrell, 1957). Although this methodology is highly flexible and versatile (Thiam et al., 2001; Bravo-Ureta and Pinheiro, 1993), its outcomes are heavily dependent on the inputs used and refer to current technologies and practices used by farmers, which are often far below the agronomic optimum.

Different attempts have been made to reconcile agronomic and economic theories for integrated analysis at farm level (Hoang, 2013; de Koeijer et al., 1999; Herdt and Mandac, 1981). Most of these studies use concepts and definitions from both agronomy and agricultural economics, propose methodological modifications of existing methods to gain further insights on existing yield gaps and provide an empirical application of the concepts using farm survey data. To perform meaningful comparative analysis, it is important to estimate and explain yield gaps using a consistent protocol with local to global relevance (van Ittersum et al., 2013) and to acknowledge that yield gaps exist due to suboptimal crop management and/or resource allocation strategies used by farmers given their personal circumstances.

In a recent study, Laborte et al. (2012) estimated rice yield gaps in Central Luzon (Philippines), Suphan Buri (Thailand), Can Tho (Vietnam) and West Java (Indonesia). Rice yield gaps were highest in Central Luzon with a magnitude of about 5.0 - 5.5 t ha^{-1} in both wet and dry seasons. An initial analysis further revealed that actual farmers' yields were positively associated with N fertiliser application and labour use. The authors acknowledged the need for a more thorough yield gap analysis and concluded that "a more in-depth farm survey could shed more light on the explanations of the yield gaps and the differences in performance between average and best-yielding farmers".

In this paper, we propose to combine production ecology with methods of frontier analysis in a theoretical framework and apply this to a longitudinal survey of rice farming households in Central Luzon, Philippines. The objectives of this study are to decompose the rice yield gap into efficiency, resource and technology yield gaps and to explain those yield gaps using information related to crop management, farmers' objectives and constraints, and production technology employed. Specific research questions include 1) what is the magnitude of the partial yield gaps (i.e. efficiency, resource and technology) in rice-based farming systems of Central Luzon; 2) how have those partial yield gaps changed over time and 3) what are the overriding factors referring to crop management explaining the aforementioned yield gaps?

3.2 Theoretical framework

Yield gap analysis can be used to investigate the relative contribution of different growth factors to actual yields (van Ittersum and Rabbinge, 1997). As schematically represented in Figure 3.1, we propose to decompose the rice yield gap into an efficiency, resource and technology yield gap. The efficiency and resource yield gaps can be estimated using frontier analysis, as commonly done in production economics (cf. Herdt and Mandac, 1981), whereas the technology yield gap can be estimated using crop growth models. The relevant concepts and rationale of this framework for its application to lowland irrigated rice-based farming systems of Southeast Asia are described below.

3.2.1 Definition of yield levels

Five yield levels are differentiated to explain the rice yield gap (Figure 3.1). The potential yield (Yp) refers to the maximum theoretical yield achieved by a specific crop genotype in a well-defined biophysical environment. It is traditionally quantified with crop growth models assuming that water and nutrients are optimally supplied to the crop, and pests, diseases and weeds are fully controlled (van Ittersum and Rabbinge, 1997). Yp provides the biophysical benchmark for yield gap analysis in irrigated rice-based farming systems.

The actual yield (Ya) refers to the yield observed in farmers' fields. This can be acquired from surveying individual households at the end of a particular season. Highest farmers' yields (Y_{HF}) is an empirical concept intended to define the maximum

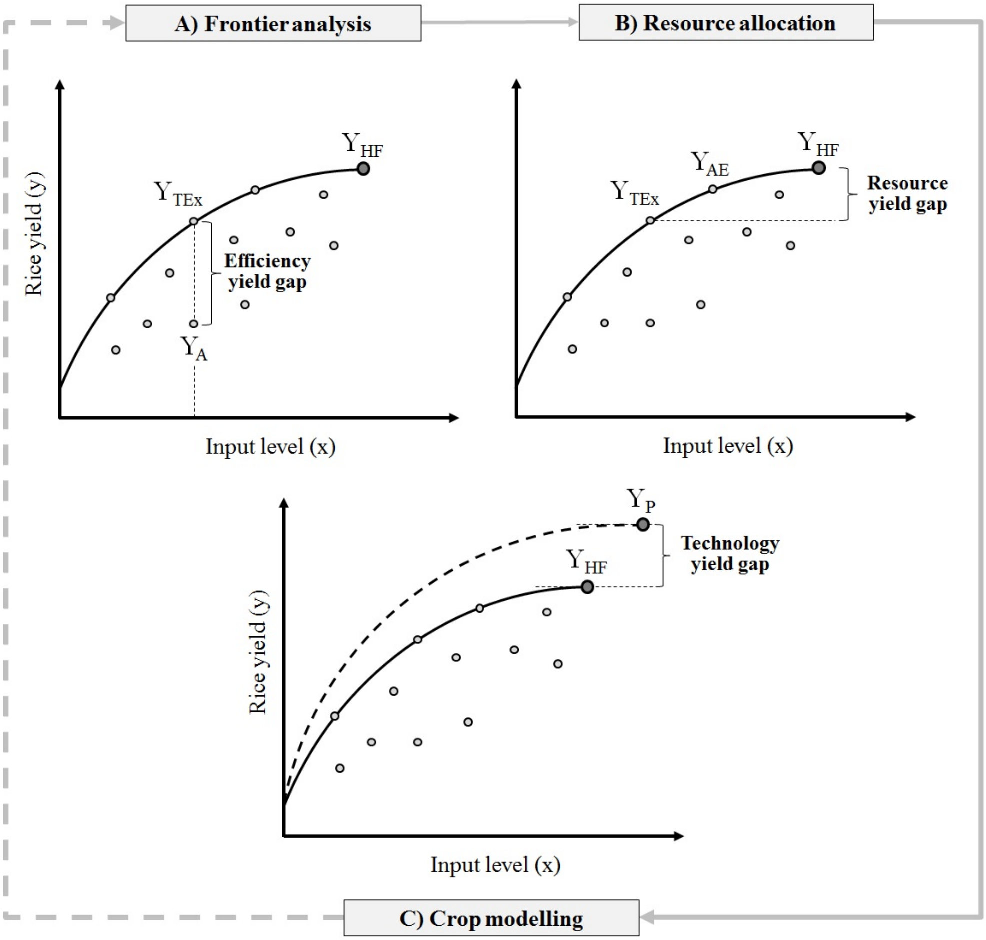

Figure 3.1: General framework for explaining rice yield gaps in lowland irrigated rice-based farming systems. Rice yield (y) is expressed in t ha^{-1}. Input level (x) refers to a vector of input variables defined based on growth-defining, -limiting and -reducing factors which are expressed either as continuous (kg input ha^{-1}) or dummy variables. Single input-output relationships are shown for illustration purposes only. Yp is the potential yield as defined by van Ittersum and Rabbinge (1997). Y_{HF}, Y_{TEx} and Ya are abbreviations for highest farmer's yield, technical efficient yield at a specific input level and actual yield of each individual farm, respectively. Y_{AE} is the allocative efficient yield which can be obtained given farmers' objectives and constraints: it is equal to Y_{HF} from a production perspective. Each dot represents an individual field in a well-defined biophysical environment.

Ya achieved. Y_{HF} can be estimated by calculating the mean of Ya above the 90th percentile or methods of frontier analysis (e.g. van Ittersum et al., 2013). Technical efficient yields (Y_{TEx}) refer to the highest possible yields obtained given observed levels of inputs in a well-defined biophysical environment. Y_{TEx} can be quantified with methods of frontier analysis following Farrell (1957). Within Y_{TEx}, an allocative efficient yield (Y_{AE}) can be identified based on different resource allocation strategies according to farmers' preferences, objectives and constraints.

3.2.2 Efficiency yield gap

We refer to the difference between Y_{TEx} and Ya as the efficiency yield gap (Figure 3.1A). Similar to the economic concept of 'technical inefficiency', it is a one-dimensional measurement of performance which indicates how far each producer is from the (multi-dimensional) production frontier and hence, how much additional output can be produced given observed levels of inputs. Many studies have dealt with its estimation and explanation using different econometric and statistical techniques (e.g. Thiam et al., 2001; Bravo-Ureta and Pinheiro, 1993).

Estimates of Y_{TEx} for each individual producer depend on the input and output variables specified in the analysis. In production economics, production frontiers are traditionally estimated based on inputs regarding land, labour and capital, under the assumptions of perfect input substitutability and that input use is restricted and not easily modified in the short term (Sadoulet and de Janvry, 1995). Here, we propose to specify input and output variables according to concepts of production ecology in order to perform a consistent and agronomically relevant yield gap analysis (van Ittersum et al., 2013). This implies that growth-defining, -limiting and -reducing factors need to be explicitly taken into account, i.e. defined as 'input level' in Figure 3.1, in the estimation of the production frontier. An efficiency yield gap equal to zero indicates that a producer is on the production frontier and an efficiency yield gap greater than zero implies that a producer is below the production frontier (for all the inputs considered).

Following this approach, the fact that two producers use the same level of inputs, in a well-defined biophysical environment, but obtain different levels of outputs (Figure 3.1A) suggests differences in crop management, referring to differences in crop establishment dates, and time, space and form of the inputs applied. The relation between the efficiency yield gap and the crop management employed can be assessed using e.g. a second stage multiple regression coupled with the production frontier (Battese and Coelli, 1995).

3.2.3 Resource yield gap

The difference between Y_{HF} and Y_{TEx} is defined as the resource yield gap from a production perspective (Figure 3.1B). This yield gap captures the trade-off between maximum actual yield (i.e. Y_{HF}) and different resource allocation strategies pursued by farmers to sustain their livelihoods and cope with external shocks (e.g. weather extremes and price variability). A resource yield gap greater than zero indicates that producers may have an objective different from yield maximisation.

The underlying assumption in this approach is that each resource allocation strategy has an optimal (technical and) allocative efficient yield (Y_{AE}), which is likely to differ from Y_{HF} in case economic and environmental objectives are taken into account. As an example, high (low) input prices and low (high) output prices can help explain why Y_{AE} can be lower than (equal to) Y_{HF} for 'profit maximizing' producers. This set up also allows incorporation of the concept of resource use efficiency (RUE) and exploration of trade-offs between a strategy which aims for maximum RUE and a strategy which aims for Y_{HF}. It is acknowledged that understanding resource allocation is far more complex than reducing farmers' objectives to mere hypothetical 'profiles' and this requires a deeper understanding of farmers' behaviour and social relations. However, this simplification may prove useful to understand the contribution of socio-economic, environmental and policy circumstances to the resource yield gap.

3.2.4 Technology yield gap

The technology yield gap can be computed as the difference between Yp and Y_{HF} and can be visualized as an (non-equidistant) upwards shift of the production frontier estimated using Ya (Figure 3.1C). As suggested in Figure 3.1, any new frontier generated will have its own efficiency and resource yield gaps.

Two types of technology yield gaps can be identified depending on whether they result from a total or partial shift of the production frontier. A complete shift of the production frontier can be achieved through the adoption of precision agriculture practices or new varieties and manipulation of sowing practices. Alternatively, a partial shift of the production frontier is input-specific and requires the closure of the resource yield gap of another input, which can be identified, following von Liebig's law of the minimum, as the limiting factor to production. As an example, the provision of irrigation water (i.e. closing the resource yield gap for water) will shift the production frontier of N-input as compared to a rainfed system. A similar shift would occur in case biotic stresses referring to pest, disease and weed pressure would be removed or a technology such as site-specific nutrient management (Dobermann et al., 2002) would be adopted. In addition, the distinction between the technology and the resource yield gaps depends on the sample/survey used in the analysis, i.e. if there are no farmers applying the quantity of N necessary to achieve Yp then N limitation becomes part of the technology yield gap. Crop growth simulation models can be used to generate agronomically optimum production frontiers that achieve Yp, and hence to estimate the technology yield gap.

3.3 Material and methods

3.3.1 Central Luzon Loop Survey

The theoretical framework was tested using the Central Luzon Loop Survey. This farm household survey is an unbalanced panel dataset collected by the International Rice Research Institute (IRRI) every 4 to 5 years during the period 1966 - 2012 with the objective of monitoring changes in crop management and household characteristics over time in the rice-based farming systems of Central Luzon (Moya et al., 2015). The rice fields surveyed are located along the main highway and were systematically selected at a specific distance. Our analysis focused on the surveys covering the period 1979 - 2012, unless otherwise indicated, because irrigation was only widely available in the region after 1975.

Descriptive statistics of the variables used in this analysis are presented in Tables B1 and B2 for the WS and DS, respectively (Supplementary Material). On average, 103 and 59 households cultivating rice were interviewed in the WS and DS, respectively. The sample size is lower in the DS because fewer households cultivated rice during this season, possibly due to water-related constraints or cultivation of other crops. There are more fields than households surveyed because 60% of the households cultivated more than one field in a particular season (out of which 82% cultivated two fields). The average farm size was about 1.2 ha over the period analysed.

3.3.2 Description of study site

We refer to Central Luzon as the region covering the administrative provinces of Bulacan, La Union, Nueva Ecija, Pampanga, Pangasinan and Tarlac. The total harvested rice area in this region increased from 0.7 to 0.9 million ha during the period 1987 - 2014 (PSA-BAS, 2015). Such area expansion can be attributed to a steady increase in irrigated rice area from 70% of total rice area in 1987 to 83% in 2014. According to national statistics, actual rice yields in the irrigated areas increased from about 3.3 t ha^{-1} in 1987 to 5.2 t ha^{-1} in 2014, and remained on average 19% higher than the national average.

Central Luzon can be described as a flat alluvial floodplain surrounded by the Zambales mountains on the West and the Sierra Madre mountains on the East. Most of the soils are suitable for irrigated rice production: they were primarily formed from the deposition of sediments from the adjacent mountains, which were often rejuvenated with volcanic sediments, exhibit medium to heavy textures (clay loam to clay) and have an aquic soil moisture regime (Gines et al., 2004). Results from a soil survey conducted in November 2014 in 96 fields included in the household survey indicated the existence of four predominant textural classes (clay, clay loam, loam and silty clay loam) and an average organic C of 1.23 ± 0.36 g C kg^{-1}, P-Olsen of 14.06 ± 12.68 mg P kg^{-1} and exchangeable K of 0.26 ± 0.24 $cmol_c$ K kg^{-1} across all classes (Silva et al., 2014).

The region is characterized by a moist tropical monsoon climate (Figure 3.2). Precipitation averages 1500 mm yr^{-1}, 89% of which occurs from May to October. Minimum temperatures are rather constant throughout the year with a mean value of 23.1°C. Maximum temperatures have a mean value of 32.2°C throughout the year and are slightly higher between March and May. Solar radiation is higher in the first half of the year with a mean value of 19.5 MJ m^{-2} and a maximum of 24.3 MJ m^{-2} occurring in April. Due to cloudiness, mean solar radiation is about 15.8 MJ m^{-2} during the second half of the year. Double rice cropping systems are the most common farming system in the region with a wet season (WS) crop between June/July and September/October and a dry season (DS) crop between December/January and March/April. Sometimes farmers replace the DS rice crop by vegetables or maize, particularly in areas where irrigation water is not readily available.

Actual rice yields have been rather constant through the WS with a mean value of 3.5 and 3.8 t ha^{-1} for modern varieties type 3 (Mv3) and 4 (Mv4), respectively (Table B1; Supplementary Material). During the DS there was a slight increase in actual rice yields over time and mean values were 4.1 and 4.9 t ha^{-1} for Mv3 and Mv4 type of varieties (Table B2; Supplementary Material). Applications of N, P and K were always higher in the DS than in the WS and increased with time: in 1979-1980 an average farmer applied 75 kg N ha^{-1}, 15 kg P ha^{-1} and 8 kg K ha^{-1} compared with 106 kg N ha^{-1}, 29 kg P ha^{-1} and 22 kg K ha^{-1} in 2011-2012. For pesticides, it is important to highlight the decrease in insecticide use and the slight increase in herbicide use particularly in the DS. There was a sharp increase in seed use during the 1980s and 1990s, particularly in the WS due to a shift from transplanting to direct seeding. However, a decrease in seed use was observed after the late 1990s due to re-adoption of manual transplanting in some farms.

3.3.3 Stochastic frontier analysis

Stochastic frontier analysis (Kumbhakar and Lovell, 2000; Meeusen and van den Broeck, 1977; Aigner et al., 1977) was used to estimate and explain the efficiency yield gap for rice farming in Central Luzon. We opted for this parametric approach of frontier analysis because it explicitly separates the effects of technical inefficiency and statistical noise, which is justified in our analysis. Further information about model selection and evaluation is available online as Supplementary Material.

The stochastic frontier model with inefficiency effects was applied to two subsets of the household survey. First, panel data for the period 1979 - 2012 were used to characterize the production frontier and quantify the efficiency yield gap. Secondly, cross-sectional data for the period 2011 - 2012 were used to assess the contribution of field-specific soil properties to the efficiency yield gap. The parameters and error terms of the stochastic frontier models and the second stage regression for technical inefficiency effects were estimated simultaneously with the method of maximum likelihood of the R package *frontier* (Coelli and Henningsen, 2013).

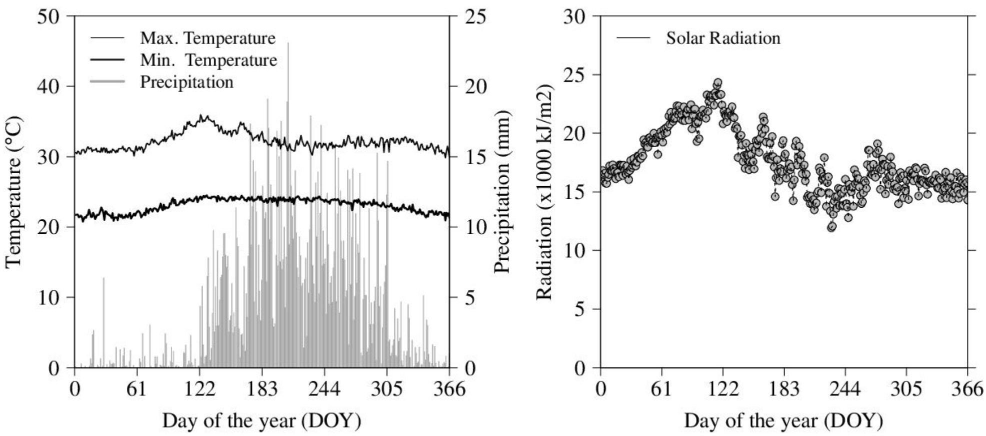

Figure 3.2: Mean daily maximum and minimum temperature (°C), rainfall (mm) and solar radiation (x1000 kJ m^{-2}) for the period 1983 - 2003 in Muñoz, Nueva Ecija. *Source*: GSOD and NASA Power databases.

Panel data: 1979 - 2012

The stochastic production frontier and efficiency yield gaps (Equation 1) were estimated based on quantitative input-output data for a set of unique combinations of household $\times$ field $\times$ year $\times$ season. Actual rice yields were corrected to 14% moisture content (y, kg ha^{-1}) and used as dependent variable.

The vector of independent variables, x_k, was designed to analyse the contribution of growth-defining, -limiting and -reducing factors to crop yield. A dummy variable was considered to control for differences between rice varieties which were classified into modern varieties type 1 - 3 (Mv3) and and type 4 (Mv4), following the classification proposed by Launio et al. (2008). As opposed to Neumann et al. (2010), growth-defining factors referring to weather conditions (e.g. solar radiation and temperature) were not explicitly included as input variables because Central Luzon is assumed to be a climatic homogeneous region (van Wart et al., 2013; Moya, 2001). Instead, a season-dummy and a year-variable were included to capture differences between WS and DS and the different years, respectively. Finally, we included a set of variables in linear and quadratic form, and interaction terms between them, to capture the contribution of growth-limiting (i.e. nitrogen, phosphorus, potassium and seed use, kg ha^{-1}) and -reducing factors (i.e. insecticide and herbicide use, kg active ingredient ha^{-1}) to actual yields as well as a dummy variable with transplanting and direct seeding establishment methods. The y and x_k variables included in the production frontier were mean-scaled and ln-transformed prior to the analysis.

The efficiency yield gap was explained with a second stage multiple regression (Equation 2) coupled with the stochastic production frontier following the specification of Battese and Coelli (1995). The number of fertilizer and pesticide applications (#) were considered as explanatory variables, z_{it}, in the estimation of the production function which best fitted our data since this information was available for all observations (Model 1, n = 1406). The impact of the timing of the first and second fertilizer and pesticide application (expressed in days after transplanting, DAT) on the efficiency yield gap was further assessed by including them as z_{it} variables into an alternative stochastic frontier model fitted to the data containing information on timing (Model 2, n = 848). No data transformations were performed to the z_{it} variables.

Likelihood ratio tests of alternative stochastic frontier models were used to select the specification which best fitted the data (details are available online as Supplementary Material). As a result, we used a stochastic frontier model with a translog functional form, observation-specific technical efficiency scores and, non-constant and neutral technical change to estimate the efficiency yield gap and Y_{TEx} of farm i in year t:

$$\begin{aligned} ln\, y_{it} \; &= \; \alpha_0 \; + \; \sum_k^K \beta_k \; ln\, x_{kit} \; + \; \frac{1}{2} \sum_k^K \sum_j^K \theta_{kj} \; ln\, x_{kit} \times ln\, x_{jit} \; + \\ & \quad + \; \delta \, T \; + \; \lambda \, T^2 \; + \; v_{it} \; - \; u_{it} \qquad (3.1) \\ u_{it} \; &= \; b_0 \; + \; \beta' \, z_{it} \; + \; \beta'_T \, T \; + \; \epsilon_{it} \qquad (3.2) \\ \text{EffYg}_{it} \; &= \; 1 \; - \; exp(-u_{it}) \qquad (3.3) \\ \text{Y}_{\text{TEx}_{it}} \; &= \; \text{Ya}_{it} \times exp(-u_{it})^{-1} \qquad (3.4) \end{aligned}$$

where y_{it} is the dependent variable and x_{kit} and z_{it} are the independent and explanatory variables defined as before; v_{it}, u_{it} and ϵ_{it} are random errors to be estimated under specific distributional assumptions (see Battese and Coelli, 1995); and α_0, β_k, δ, b_0, β' and β_T' are the parameters to be estimated. The input-specific parameter β_k can be interpreted as an output elasticity at the sample mean, i.e. the responsiveness in percent of rice yield to a one percent increase of a particular input (in case θ_{kj} = 0). The translog functional form is captured through θ_{kj} which specifies both interaction and quadratic terms of x_k. The parameter δ captures the constant annual rate of technological progress or regress. Last but not least, λ indicates whether the process of technological change is non-constant (i.e. not linear) over time.

Cross-sectional data: 2011 - 2012

The soil properties included as input variables in the frontier model were clay and sand content (%), pH in H_2O, P-Olsen (mg P kg^{-1}), exchangeable-K ($cmol_c$ K kg^{-1}) and soil organic C (%). These data were merged with crop management data collected for the same fields in the last round of the household survey, resulting in a sample size of 54 and 53 fields during the 2011 WS and 2012 DS, respectively.

The impact of field-specific soil properties on the efficiency yield gap was assessed using a simple Cobb-Douglas stochastic frontier model, i.e. a functional form representing the technological relationship between the amount of inputs and the amount of output produced by those inputs (expressed in ln and neglecting interactions and quadratic terms). The aforementioned soil properties were used as input variables in addition to growth-defining, -limiting and -reducing factors previously specified. The choice for a simpler model was justified by our interest in assessing the impact of soil variables on the efficiency yield gap rather than in characterizing exhaustively the production frontier for this subset of the data. All variables were constructed as described in Section 3.2.1 and no z_{it}-variables were included. Differences in the efficiency yield gap were compared for a model which did not take into account soil properties (Model 3) and one which included soil properties as independent variables (Model 4).

3.3.4 Resource allocation

The resource yield gap was analysed from a production perspective by comparing Y_{TEx} and Y_{HF} (equivalent to Y_{AE} from a production perspective) for each unique combination of household × field × year × season. This yield gap indicates by how much yields could increase in case all technical efficient fields would use as much inputs as the highest-yielding fields observed in the dataset. Y_{TEx} was quantified using the stochastic frontier Model 1 specified above and Y_{HF} as the mean across fields of Ya values above the 90^{th} percentile. In the case of N-specific resource yield gaps, fields with N application levels greater than the average input to achieve Y_{HF} were considered to have a resource yield gap equal to zero.

To explain the resource yield gap from a production perspective, rice fields were grouped according to their Ya values. Besides Y_{HF}, we identified the average of lowest yielding fields (Y_{LF}) as the mean across fields of Ya below the 10^{th} percentile. Average yielding fields (Y_{AF}) are the fields with Ya between the 10^{th} and 90^{th} percentiles. Statistical differences in the mean input use observed in fields with Y_{HF} and Y_{LF} were assessed using the Mann Whitney *U* test (p-value $\leq$ 5%) under the alternative hypothesis that rice yield and input use are lower for Y_{LF} than for Y_{HF}. This was performed per season for seed, fertiliser and pesticide use but results are presented only for N, P and K because these were the most important ones. This analysis was the starting point to assess to which extent greater rice yields can be obtained with more inputs and hence, to understand if input use is sub-optimal in the study region.

3.3.5 Crop modelling: ORYZA v3

The ORYZA v3 model was able to simulate rice yield relatively well for the varieties IR72 and NSIC Rc222 (RMSE between 10 - 20%) but tended to overestimate experimental yields by about 1.0 t ha^{-1} (modelling efficiency between 0.65 and 0.97). Standard crop parameters were used for crop development rates, leaf and stem growth, photosynthesis rate, partitioning coefficients and N concentration in different organs across all simulations performed. These have been previously calibrated by Bouman and van Laar (2006) and Qin et al. (2014) for IR72 and NSIC Rc222, respectively. IR72 and NSIC Rc222 are representative of the Mv3 and Mv4 varieties, respectively.

Details of model calibration and evaluation are available online as Supplementary Material.

ORYZA v3 was then applied to simulate a production frontier of N-limited yields (Yn) by means of a sensitivity analysis on the rate and timing of N application, assuming no water limitations (i.e. water balance turned-off). For this purpose, a set of 2500 random N management strategies were created for which N-limited yields were subsequently simulated. Each strategy comprised five different N applications within the growing season. The exact date of each N split was randomly generated within a pre-defined range: 1) basal application 20 - 30 days after sowing (DAS), 2) top-dressing 30 - 40 DAS, 3) top-dressing 40 - 50 DAS, 4) top-dressing around panicle initiation 50 - 70 DAS, and 5) top-dressing around flowering 70 - 90 DAS. For each split, a N rate varying from 0 to 80 kg N ha^{-1} was randomly computed. The total N applied ranged between 0 and 400 kg N ha^{-1} and was normally distributed.

Model simulations were performed for the varieties IR72 and NSIC Rc222 and for 36 different sowing dates within a year for the period 1984 - 2012. The input data were solar radiation, minimum and maximum temperature from a local weather station located in Muñoz (Nueva Ecija, Figure 3.2) and detailed soil profiles of four different soil types (clay, loam, clay loam and silty clay loam; Silva et al., 2014). This set up resulted in a total of 20.16 million genotype $\times$ environment $\times$ management combinations which were simulated with ORYZA v3. Yp, Yn, the quantity of N needed to achieve Yp and the magnitude of the technology yield gap were analysed for the mean sowing date observed in the household survey in a particular year. Results for the temporal variation of the technology yield gap are presented for IR72 for the periods 1986 - 1987 and 1994 - 1995 and for NSIC Rc222 for the periods 2003 - 2004 and 2011 - 2012. In addition, production frontiers of Mv3 and Mv4 varieties are presented for the period 2003 - 2004 to illustrate technological change due to varietal shift.

3.4 Results

3.4.1 Efficiency yield gap

Properties of the production frontier

The sign, magnitude and significance level of parameter estimates obtained for Model 1 and 2 are rather similar (Table 3.1). The dummy 'variety' was not statistically significant meaning that the model did not capture yield differences between Mv3

Table 3.1: Parameter estimates of the stochastic frontier models estimated for lowland irrigated rice-based farming systems of Central Luzon, Philippines. Model 1 and 2 were applied to panel data (1979 - 2012) and differ in the timing variables included as inefficiency effects. Model 3 and 4 were applied to cross-sectional data (2011 - 2012) and differ in the soil properties included in the production frontier. Significance codes: '***' 0.1% '**' 1% '*' 5%.

Variable	Model 1	Model 2	Model 3	Model 4
Production frontier				
Intercept	**0.118** ***	**0.098** *	-0.088	-0.228
Nitrogen	0.079	0.090	0.064	0.088
$Nitrogen^2$	-0.052	**0.161** *		
Phosphorus	-0.051	-0.088	-0.030	-0.054
$Phosphorus^2$	0.064	0.052		
Potassium	**0.191** ***	**0.166** *	0.122	**0.155** *
$Potassium^2$	0.040	0.057		
Seed	**0.116** *	0.129	**-0.078** *	**-0.089** *
$Seed^2$	**0.113** *	0.027		
Insecticide	0.018	0.008	0.029	0.030
$Insecticide^2$	0.014	**0.025** *		
Herbicide	0.029	-0.009	0.008	0.015
$Herbicide^2$	0.000	-0.017		
Method_Transplanting	**0.085** **	**0.071** *		
Variety_Mv4	-0.046	-0.048	0.089	0.072
Season_DS	**0.111** ***	**0.135** ***	**0.386** ***	**0.372** ***
Year	**-0.016** ***	**-0.016** ***		
$Year^2$	**0.004** ***	**0.004** ***		
Clay content				0.004
Sand content				0.001
Soil organic C				-0.071
Available P-Olsen				0.003
Exchangeable K				0.013
Nitrogen × Phosphorus	0.034	0.039		
Nitrogen × Potassium	-0.032	-0.070		
Nitrogen × Seed	-0.017	-0.054		
Nitrogen × Insecticide	-0.018	-0.020		
Nitrogen × Herbicide	**-0.044** *	**-0.053** *		
Phosphorus × Potassium	-0.027	-0.050		
Phosphorus × Seed	0.062	0.061		
Phosphorus × Insecticide	-0.016	-0.015		
Phosphorus × Herbicide	0.024	0.032		

Table 3.1: Parameter estimates of the stochastic frontier models estimated for lowland irrigated rice-based farming systems of Central Luzon, Philippines. Model 1 and 2 were applied to panel data (1979 - 2012) and differ in the timing variables included as inefficiency effects. Model 3 and 4 were applied to cross-sectional data (2011 - 2012) and differ in the soil properties included in the production frontier. Significance codes: '***' 0.1% '**' 1% '*' 5%. *(continued)*

Variable	Model 1	Model 2	Model 3	Model 4
Potassium × Seed	**-0.117** **	-0.074		
Potassium × Insecticide	0.024	0.029		
Potassium × Herbicide	-0.001	0.003		
Seed × Insecticide	-0.009	-0.010		
Seed × Herbicide	0.004	0.019		
Insecticide × Herbicide	0.007	0.010		
Method_Transp × Nitrogen	-0.072	-0.103		
Method_Transp × Phosphorus	0.091	0.103		
Method_Transp × Potassium	**-0.115** *	-0.062		
Method_Transp × Seed	0.068	-0.039		
Method_Transp × Insecticide	0.012	0.018		
Method_Transp × Herbicide	-0.020	0.016		
Variety_Mv4 × Nitrogen	-0.044	-0.024		
Variety_Mv4 × Phosphorus	0.020	0.007		
Variety_Mv4 × Potassium	**-0.110** *	**-0.098** *		
Variety_Mv4 × Seed	-0.066	-0.052		
Variety_Mv4 × Insecticide	-0.011	-0.007		
Variety_Mv4 × Herbicide	-0.027	-0.032		
Season_DS × Nitrogen	**0.181** ***	**0.139** **		
Season_DS × Phosphorus	0.081	**0.109** *		
Season_DS × Potassium	-0.005	0.016		
Season_DS × Seed	-0.032	-0.044		
Season_DS × Insecticide	0.007	0.022		
Season_DS × Herbicide	0.021	0.034		
Inefficiency effects				
Intercept	-0.453	-1.271		
Year	**-0.070** ***	**-0.057** **		
No. fertiliser applications	**-0.141** *	-0.088		
Timing 1^{st} fertilisation		**-0.016** *		
Timing 2^{nd} fertilisation		**0.010** *		
No. pesticide applications	-0.023	0.030		
Timing 1^{st} spray		**0.016** **		
Timing 2^{nd} spray		**0.013** **		

Table 3.1: Parameter estimates of the stochastic frontier models estimated for lowland irrigated rice-based farming systems of Central Luzon, Philippines. Model 1 and 2 were applied to panel data (1979 - 2012) and differ in the timing variables included as inefficiency effects. Model 3 and 4 were applied to cross-sectional data (2011 - 2012) and differ in the soil properties included in the production frontier. Significance codes: '***' 0.1% '**' 1% '*' 5%. *(continued)*

Variable	Model 1	Model 2	Model 3	Model 4
Model evaluation				
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	**0.482** ***	**0.432** ***	**0.175** ***	**0.158** ***
$\gamma \ = \sigma_u^2 \ / \ \sigma^2$	**0.939** ***	**0.940** ***	**0.920** ***	**0.886** ***
TE scores (%)	74.0	75.4	74.8	76.2
Sample size (n)	1397	844	107	107

'TE scores' refer to the average of the field-specific Technical Efficiency scores (i.e. 100 minus the efficiency yield gap expressed in %) obtained for each stochastic frontier model.

and Mv4 varieties. The dummy 'season' and variable 'year' were both statistically significant at 0.1%, which indicates that there were important intra- and inter-annual variations of climatic conditions affecting Ya. Linear and quadratic 'year' variables were statistically significant and their signs indicate that the rice farms surveyed experienced both technological regress and progress, i.e. a period in which rice productivity declined was followed by a period in which rice productivity increased.

As for crop management, statistically significant positive effects were found between Ya and potassium (K), seed, seed^2 and crop establishment method. K use was an important production input not only due to its linear and positive effect on Ya but also due to the statistically significant negative interactions with seed, crop establishment method and variety. The effect of seed use on Ya was positive and increased with higher levels of seed use. As for the crop establishment method, fields in which rice was manually transplanted exhibit greater Ya than fields in which the crop was direct seeded. There was no significant linear and quadratic effect of nitrogen (N) on Ya. However, the significant effects of N $\times$ season and N $\times$ herbicide interactions indicate that the effect of N was season-dependent (i.e. Ya increased with increasing levels of N in the DS) and less positive when herbicide use was high (i.e. greater yield responses to N occur when lower amounts of herbicides are applied), respectively. In addition, the effect of N^2 became significant when timing variables are included in Model 2.

Explaining the efficiency yield gap

The mean efficiency yield gap and its frequency distribution were similar in the WS (26.5%, Figure 3.3A) and DS (25.4%, Figure 3.3B). In both seasons, there are few fields which have an efficiency yield gap $< 5\%$ while most of the fields exhibit an efficiency yield gap of 5 - 30%. However, there are also fields in which the efficiency yield gap is greater than 50%, which suggests that management is performed inefficiently.

The mean Y_{TEx} is 5.1 and 5.7 t ha^{-1} in the WS and DS, respectively (Figures 3.3C and 3.3D). As expected, Y_{TEx} was on average greater in the DS than in the WS most likely because of the more favourable growing conditions during this time of the year in terms of solar radiation and temperature. From these estimates, a mean efficiency yield gap for rice-based farming systems in Central Luzon of 1.3 t ha^{-1} was calculated for both growing seasons. As observed in Figures 3.3E and 3.3F, fields with greater Ya exhibited a lower efficiency yield gap.

The total number of fertiliser applications had a statistical significant negative effect on the efficiency yield gap whereas the effects of the total number of pesticide applications were not statistically significant (Table 3.1). The number of splits affects the agronomic use efficiency of the inputs used but a high number of applications does not necessarily imply smaller efficiency yield gaps, since the timing of the operations is crucial as well. The timing of the first and second fertiliser applications were statistically significant at 5%. The efficiency yield gap was negatively related to the timing of the first fertiliser application. This indicates that fields in which fertiliser was applied at earlier dates exhibited relatively greater efficiency yield gaps than fields which received fertiliser at later dates. The opposite was true for the timing of the second fertiliser application. As for pesticide applications, the coefficients for both timing variables were positive and statistical significant at 1%, which indicates that fields which received an earlier application of pesticides exhibited a lower efficiency yield gap as compared to fields in which biotic control was done at a later timing.

The effect of crop establishment date was not included in the stochastic frontier analysis due to incomplete data, but its variability across farmers' fields may also partly explain the efficiency yield gap. As can be observed in Figure B1 (Supplementary Material), crop establishment dates varied widely and most did not coincide with the sowing date for which Y_{HF} was achieved.

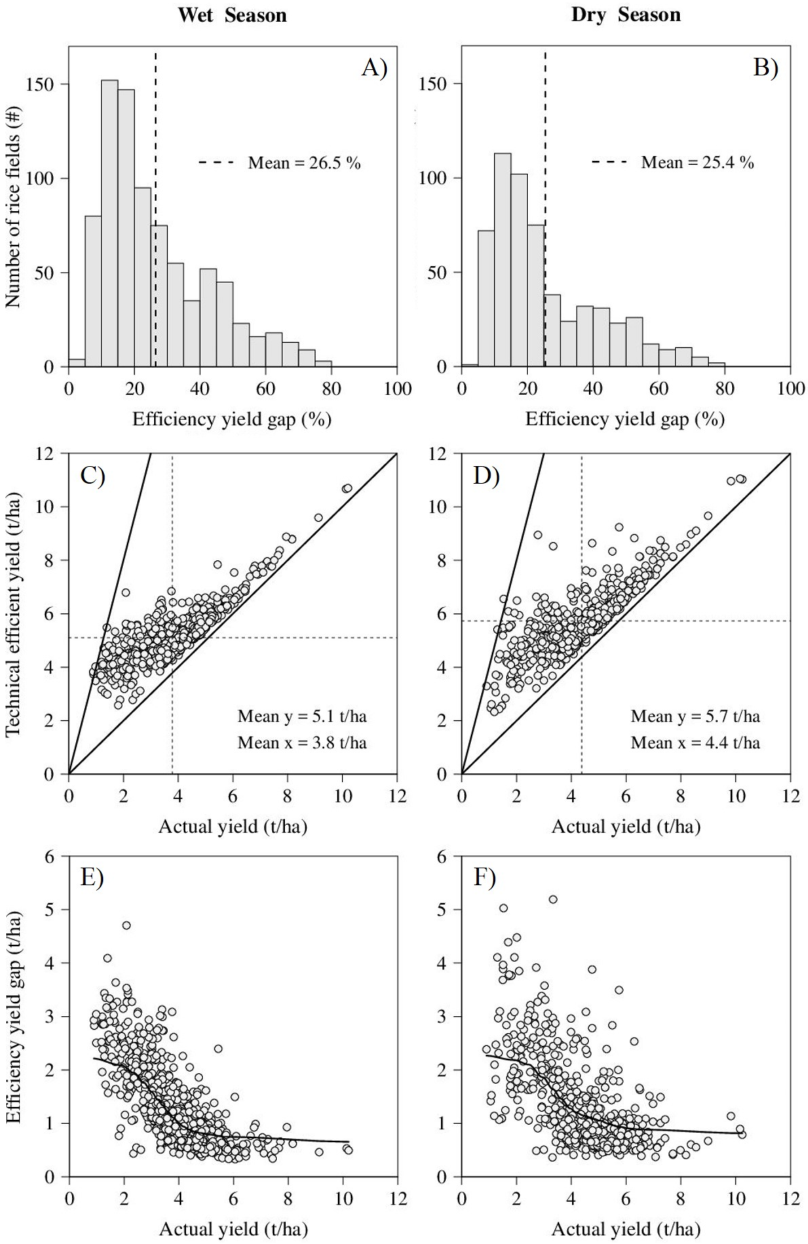

Figure 3.3: Efficiency yield gap for lowland irrigated rice-based farming systems in Central Luzon, Philippines: A) and B) show the distribution of the efficiency yield gap; C) and D) illustrate the relationship between Y_{TEx} and Ya (solid lines are the 1:1 and 1:4 lines) and E) and F) provide its magnitude for each rice field analysed in the WS (n = 827) and DS (n = 579), respectively. Dashed lines in A) and B) are the mean efficiency yield gap (%) and in C) and D) the mean Ya and Y_{TEx} (t ha^{-1}); solid lines in E) and F) are fitted polynomial regressions.

Effect of soil properties

The only difference between the production frontier of Model 3 and 4 is the inclusion of soil variables in Model 4, which were all non-significant. In Model 3 only the parameter estimate for seed and season were statistically significant, while in Model 4 the parameter estimated for potassium was also statistically significant. Differences in parameter estimates between Model 1 and Model 3 were expected *a priori* because Model 1 was estimated using data from a much wider number of years.

The Y_{TEx} values of Model 3 and 4 were close to the 1:1 line, with most values lower for Model 4 (Figure 3.4A). Model 3 slightly overestimated the efficiency yield gap as compared to Model 4 (Figure 3.4B) and the average difference in efficiency yield gap between the models was ca. 1.5% (Table 3.1). These indicate that for a proper estimation of the efficiency yield gap for rice farming systems in Central Luzon variation in soil types and properties was of minor importance. In addition, the discrepancy in absolute terms between efficiency yield gaps estimated by both models increased as the magnitude of the efficiency yield gap increased.

3.4.2 Resource yield gap

Resource yield gaps from a production perspective

A mean resource yield gap of 1.0 and 1.3 t ha^{-1} in the WS and DS, respectively, were estimated for the multiple input stochastic frontier presented in Table 3.1 (Figures

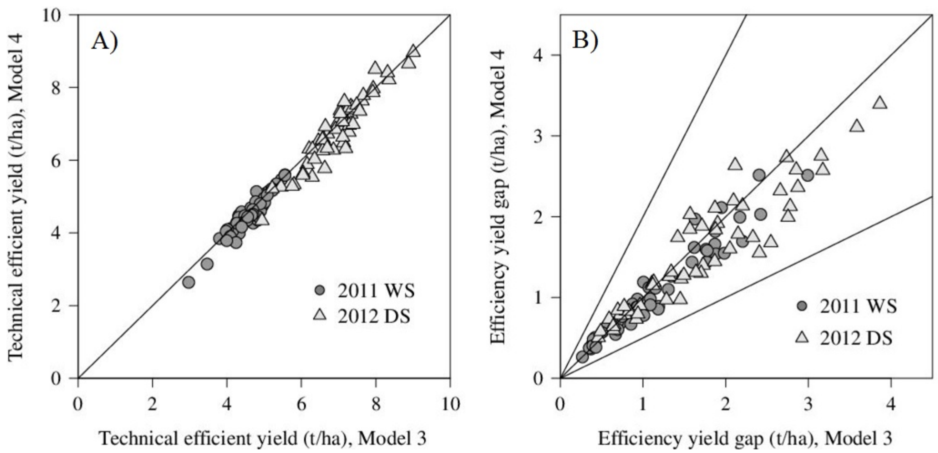

Figure 3.4: Effect of field-specific soil proprieties on a) Y_{TEx} and b) efficiency yield gaps of rice-based farming systems in Central Luzon for the 2011 WS and 2012 DS. Further information about the specification of 'Model 3' and 'Model 4' is provided in Table 3.1.

3.5A and 3.5B). This suggests that sub-optimal quantities of one or more input, i.e. the ones which are statistically significant in Table 3.1, were applied in average and lowest yielding fields. N-specific resource yield gaps are discussed below.

Yield responses to N applied were clear in the 2012 DS but not evident for the 2011 WS (Figures 3.5C and 3.5D). The yield range observed for Y_{TEx} at specific N application levels indicates that there were other limiting factors to production apart from N and it justifies the use of a multiple input frontier approach. Sub-optimal N

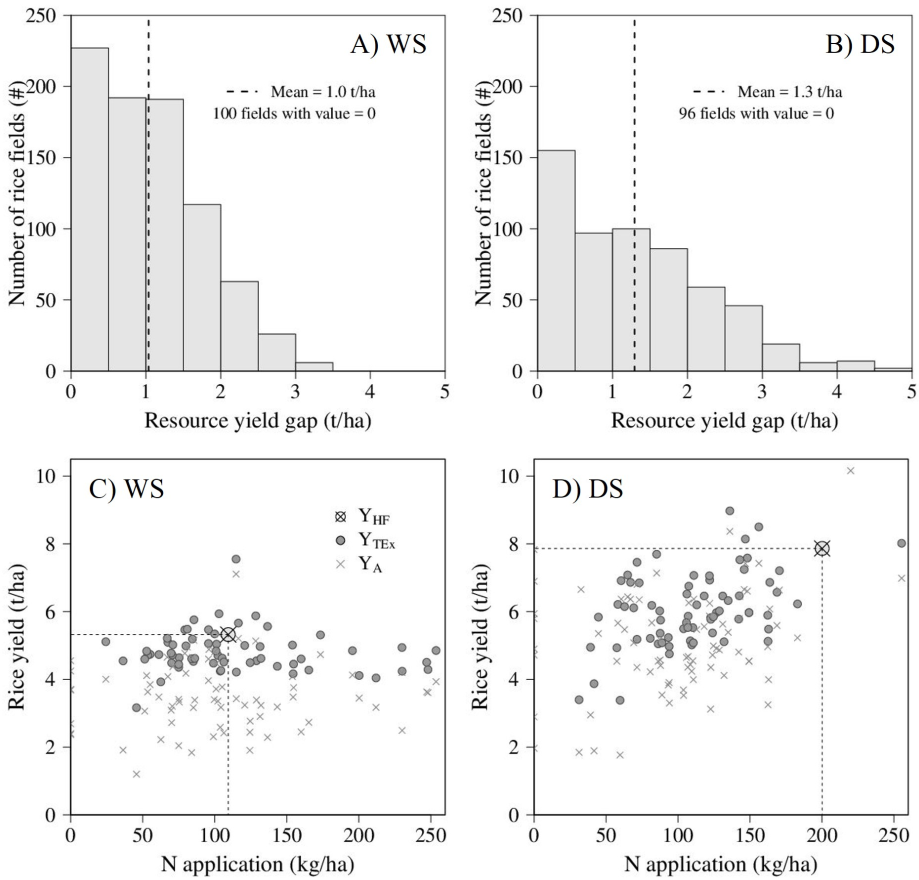

Figure 3.5: Resource yield gap for lowland irrigated rice-based farming systems in Central Luzon, Philippines: A) and B) illustrate the distribution of the resource yield gap across the period 1979 - 2012 (WS, n = 902 and DS, n = 609) ; and C) and D) show the relationship between N applied (kg ha^{-1}) and Y_{TEx} and Ya (t ha^{-1}) for individual rice fields in 2011 WS and 2012 DS, respectively.

application resulted in a mean resource yield gap of 0.8 and 1.2 t ha^{-1} in the WS and DS, respectively (data not shown). In addition, 317 and 140 fields had a resource yield gap due to N equal to zero in the WS (n = 902) and DS (n = 609), respectively, because N application levels in these fields were equal to or greater than the mean N application used in the highest-yielding fields.

Nutrient-specific resource yield gaps

There was a large variation in rice yields and N, P and K application rates among fields. No yield response to applied fertilisers was observed in the WS whereas there was a positive yield response in the DS (Figure 3.6). Rice yields ranged between 2.0 and 5.3 t ha^{-1} during the 2011 WS and 2.6 and 7.9 t ha^{-1} during the 2012 DS. Fertiliser application rates varied from 25 - 250 kg N ha^{-1}, 0 - 95 kg P ha^{-1} and 0 - 55 kg K ha^{-1}. The ranges in rice yields, N, P and K use reported for the period 2011 WS - 2012 DS were similar for the other years included in the household survey.

Fertiliser application rates to achieve Y_{HF} and Y_{LF} during the 2011 WS were not statistically different (Figures 3.6A, 3.6B and 3.6C). N, P and K use in both highest and lowest yielding fields were similar in this period: 103 vs 83 kg N ha^{-1}, 28 vs 22 kg P ha^{-1} and 21 vs 16 kg K ha^{-1}. Differences in N use between Y_{HF} and Y_{LF} were statistically significant only in the first three WS years (1979, 1982 and 1986). Statistically significant differences for P use were found for five WS years (1979, 1986, 1990, 1994 and 1999) and for K use in only two WS years (1986 and 1999). We can conclude that although there seems to be a resource yield gap in the WS, it is rather small and there is no relationship between N, P and K use and rice yield, and these inputs are generally not limiting production during this growing season.

Differences in N and K application in the 2012 DS between highest and lowest yielding fields were statistically significant (Figures 3.6D and 3.6F). This was also the case for the DS years 1980, 1991, 1995 and 2004. The highest yielding fields received a greater P application than the lowest yielding fields but the difference was only significant in 1980 and 2004. On average, approx. 188 kg N ha^{-1}, 67 kg P ha^{-1} and 49 kg K ha^{-1} were applied in highest yielding fields during 2012 DS as compared to approx. 79 kg N ha^{-1}, 21 kg P ha^{-1} and 9 kg K ha^{-1} in the lowest yielding fields. These results indicate that highest yielding fields exhibited high(er) yields in the DS partly because they received a high(er) N, P and K application rate.

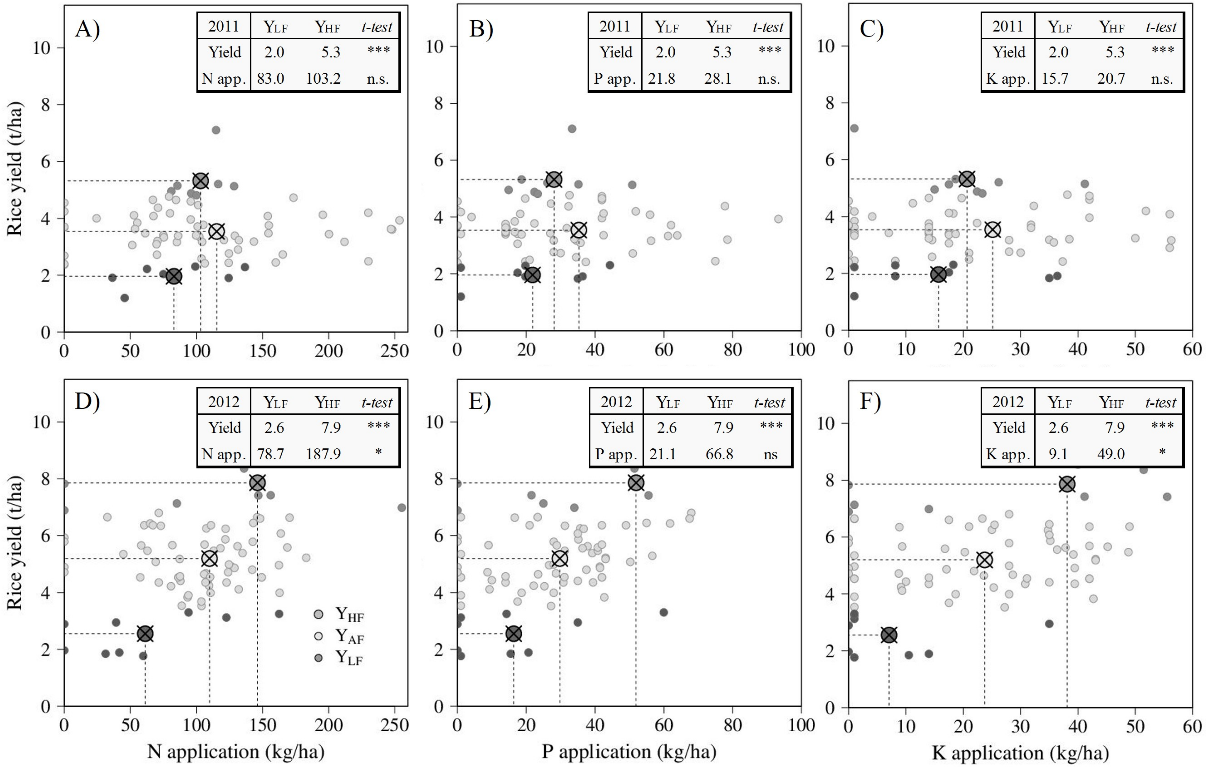

Figure 3.6: Relationship between rice yield (t ha^{-1}) and N (A and D), P (B and E) and K (C and F) application (kg ha^{-1}) during the 2011 - 2012 crop year. A) to C) refer to the 2011 WS and D) to F) refer to the 2012 DS. Individual fields were classified as highest, average and lowest yielding fields based on the Ya values. Comparisons between N, P and K rates in highest and lowest yielding fields were performed with the Mann-Whitney *U* test and results are shown in the inset tables to each figure ('n.s.' = not significant; '*' = significant at 5% level; '***' = significant at 0.1% level).

3.4.3 Technology yield gap

Magnitude and temporal variation

No major differences in Yn were observed across the four soil types used in the simulations hence, results are presented only for the 'clay' soil which is the most common soil type across the fields sampled (Silva et al., 2014). The simulated production frontiers captured well the technological means employed by farmers in terms of N and water management as most Ya values were smaller or equal to the simulated Yn (Figure 3.7).

The magnitude and variability of Yp, Yn, Y_{HF} and Ya is presented for four different periods in Figure 3.7. Yp was always higher in the DS (8.2 - 9.8 t ha^{-1}) than in the WS (6.0 - 7.6 t ha^{-1}). Similarly, Y_{HF} was also always higher in the DS (6.0 - 7.9 t ha^{-1}) as compared the WS (5.1 - 6.8 t ha^{-1}).

The technology yield gap was much larger in the DS (0.3 - 3.2 t ha^{-1}) than in the WS (0.3 - 1.8 t ha^{-1}). A particularly small technology yield gap was estimated for the WS periods prior to 2003, i.e. 0.3 - 0.9 t ha^{-1}, but it increased to about 1.6 - 1.8 t ha^{-1} during the most recent WS periods. This is due to the large Yp simulated for 2003 WS (7.6 t ha^{-1}) and the small Y_{HF} observed in 2011 WS (5.3 t ha^{-1}). For the DS periods, Yp was always far greater than Y_{HF} except for 2012 DS. The small technology yield gap estimated in 2012 DS, i.e. 0.3 t ha^{-1}, is due to a low simulated Yp of 8.2 t ha^{-1} and a high observed Y_{HF} of 7.9 t ha^{-1}.

The response of simulated yields to N applied is in agreement with the law of diminishing returns (Figure 3.7). Under irrigated conditions, N application rates of 120 - 150 and 180 - 200 kg N ha^{-1} seemed necessary to achieve Yp of IR72 in the WS and DS, respectively. This level was even greater for NSIC Rc222: 180 - 200 kg N ha^{-1} in the WS and 250 - 300 kg N ha^{-1} in the DS, respectively. N application rates greater than such thresholds (i.e. resource yield gap due to N equal to 0) are not likely to result in further yield responses. In most of the years, N application rates observed in fields which achieved Y_{HF} were far below the optimal N application rate required for Yp which partly explains the technology yield gap.

Technological change due to varietal shift

Different technologies, e.g. varieties, are characterized by different production frontiers and each of them has its specific efficiency, resource and technology yield gaps. This situation is illustrated in Figure 3.8 for the period 2003 - 2004 in which both

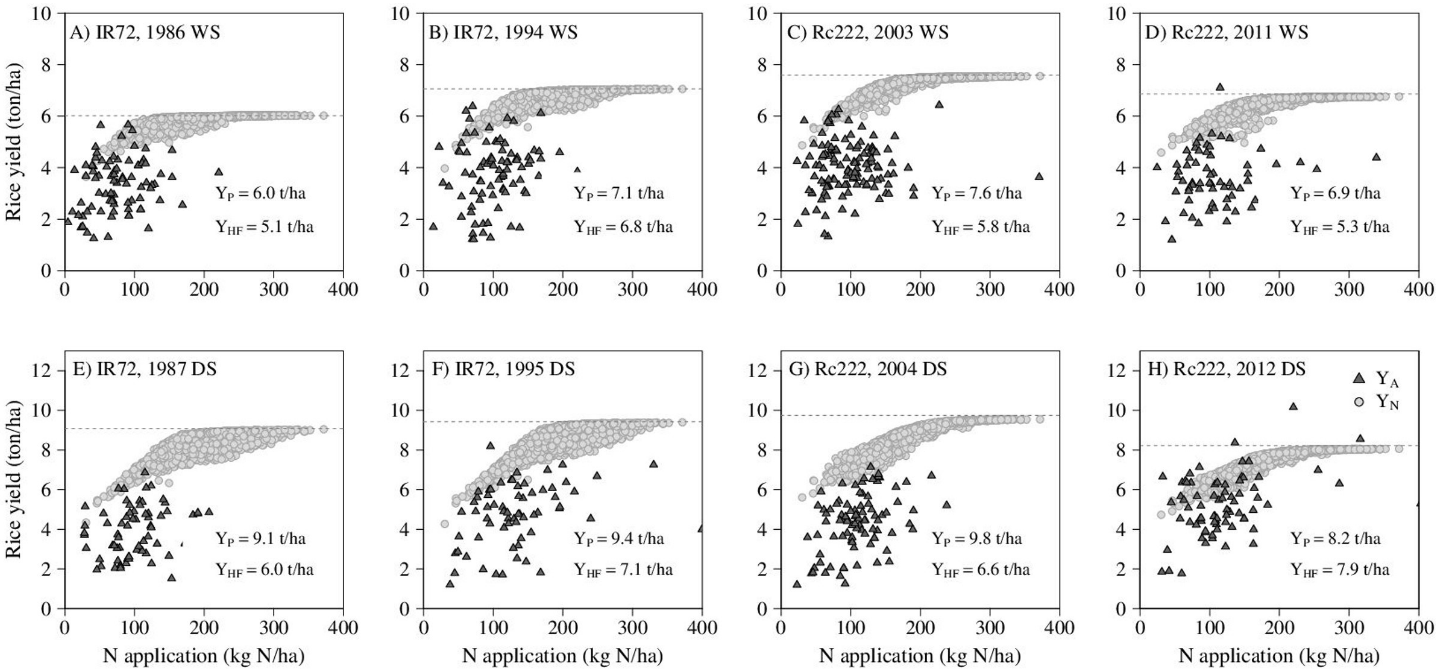

Figure 3.7: Trends in Yp, Yn in a clay soil and Ya for different N application rates in lowland rice farming systems in Central Luzon, Philippines. Figure A), B), C) and D) refer to the WS while E), F), G) and H) refer to the DS. Yp for IR72 and NSIC Rc222 is presented as a dashed line for the two years before and after 2000, respectively. Yp and Yn were computed with ORYZA v3 and Y_{HF} and Ya were obtained from the Central Luzon Loop Survey. Simulated yields are presented for the mean sowing date observed in the Central Luzon Loop Survey in each year × season.

Mv3 (benchmarked by IR72) and Mv4 (benchmarked by NSIC Rc222) varieties were cultivated in similar proportions.

Yp of IR72 was 6.0 and 8.7 t ha^{-1} during the 2003 WS and 2004 DS, respectively. For the same period, Yp was much larger for NSIC Rc222: 7.6 and 9.8 t ha^{-1}, respectively. The difference in genetic potential between the two varieties indicates that it is appropriate to consider variety-specific production frontiers in the estimation of the technology yield gap for the period in which the varietal shift occurred.

The technology yield gap observed for Mv3 varieties was 0.7 and 2.2 t ha^{-1} and for Mv4 varieties 1.8 and 2.8 t ha^{-1} in 2003 WS and 2004 DS, respectively (data not shown). The larger technology yield gap of Mv4 varieties observed in both seasons can be explained by the relatively similar Ya between Mv3 and Mv4 varieties as compared to the large differences in Yp of IR72 and NSIC Rc222. In addition, yield responses to N application were variety specific, i.e. greater N levels are required to achieve Yp of Mv4 than of Mv3 varieties. This differentiated N management was not clear in the farmers' N management practices (mean difference in N application between Mv3 and Mv4 varieties was ca. 10 kg N ha^{-1} in both 2003 WS and 2004 DS), which further explains the larger technology yield gap estimated for Mv4 varieties.

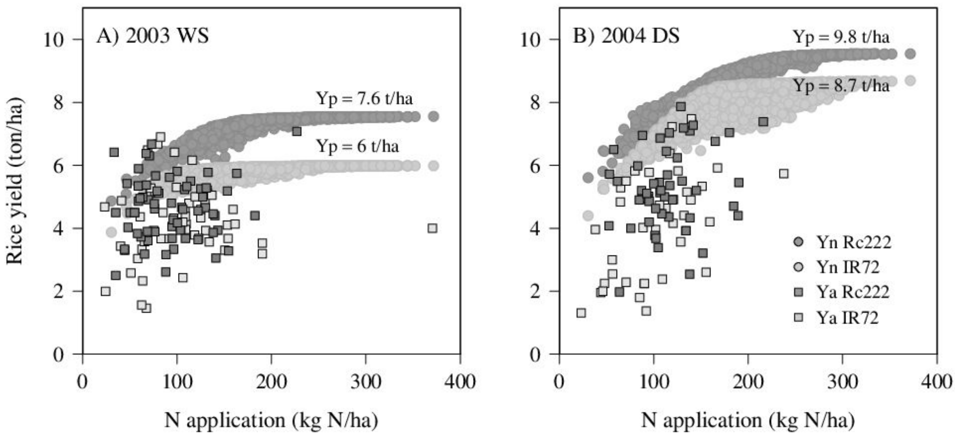

Figure 3.8: Technological change due to varietal shift in A) 2003 WS and B) 2004 DS for rice farming in Central Luzon, Philippines. Production frontiers (Yp and Yn) were computed using ORYZA v3 and Ya are actual farmers' yields in Central Luzon.

3.5 Discussion

3.5.1 Explaining rice yield gaps in Central Luzon

The mean rice yield gap was 3.2 t ha^{-1} in the WS and 4.8 t ha^{-1} in the DS (Figure 3.9 and Table 3.2). These yield gap estimates are slightly lower than the ones reported by Laborte et al. (2012) which can be explained by the longer time span of the current analysis and different input weather data used for the simulation of Yp (more precisely there were some high/unrealistic solar radiation values in the weather station used by Laborte et al. (2012), data not shown).

Our analysis showed that genetic yield potential progress in rice based-farming systems of Central Luzon was more pronounced than yield progress over the period 1979 - 2012. Our estimates of Yp for IR72 during the WS were on average 6.1 t ha^{-1} as compared to 8.7 t ha^{-1} during the DS, while for NSIC Rc222, a Yp of 7.4 and 9.6 t ha^{-1} was simulated for the WS and DS, respectively. These Y values are well within the ranges of 6.0 - 8.0 t ha^{-1} in the WS and 9.0 - 11.0 t ha^{-1} in the DS simulated by Angulo et al. (2012) and Dobermann and Witt (2004). As for Ya, farmers' yields were 3.5 t ha^{-1} in the WS on average with no temporal trend during the period 1979 - 2012 (Figure 3.9A; Table B1; Supplementary Material) while a slight increase from 3.9 t ha^{-1} in 1980 up to 4.8 t ha^{-1} in 2012 was observed during the DS (Figure 3.9B; Table B2; Supplementary Material). Yield stagnation in the WS may be associated with low input use due to climatic risk (i.e. typhoons) which is further exacerbated by the fact that many farmers cultivate rice for home consumption rather than to maximize rice production.

A mean efficiency yield gap of 1.3 t ha^{-1} was estimated for rice farming in Central Luzon (Figures 3.3C and 3.3D) with the significant tendency to decrease during the period 1979 - 2012 (Figure 3.9). Only minor differences in the efficiency yield gap (ca. 1.5%) were observed when soil properties were explicitly included in the stochastic production frontier (Table 3.1). The efficiency yield gap was associated with sub-optimal timing of application of fertilisers and pesticides. The average timing of the first and second fertiliser application were 19 and 41 DAT and roughly 30% of the fields surveyed received a single fertiliser application while the majority of fields (54%) received two. In contrast, four N-fertiliser splits are usually recommended for rice cultivation in Southeast Asia with the first application up to 10 DAT and the second at panicle initiation i.e. about 40 DAT (R. Buresh, personal communication). Our findings partially contradict the technical recommendation since fields which received the first fertilisation later than 19 DAT had smaller efficiency yield gaps than those fields with an earlier first fertilisation (Table 3.1). This may be due to the fact that hardly any farmer is performing the four N applications recommended and in that case fertiliser applications close to panicle initiation may result in greater rice yields than applications right after transplanting. In addition, the diversity and variability in biotic factors make it very hard to determine the optimal timing of operations aiming at minimizing the impact of pests and diseases. Multi-level and species-specific control mechanisms should be implemented (Savary et al., 2012) and insecticides may not be applied at all (Heong, 2014). Although not directly tested in our analysis, we expect that differences in crop establishment dates across fields may also explain part of the efficiency yield gap particularly in the DS (Figure B1; Supplementary Material).

The mean resource yield gap was rather small with a value of 1.0 and 1.3 t ha^{-1} for the WS and DS, respectively (Figure 3.9). N, P and K application rates explained part of the yield differences observed between highest and lowest yielding fields in the DS but not in the WS (Figure 3.6). This illustrates that macro-nutrients are available to farmers in the region but they were not being used with the optimal rates from a

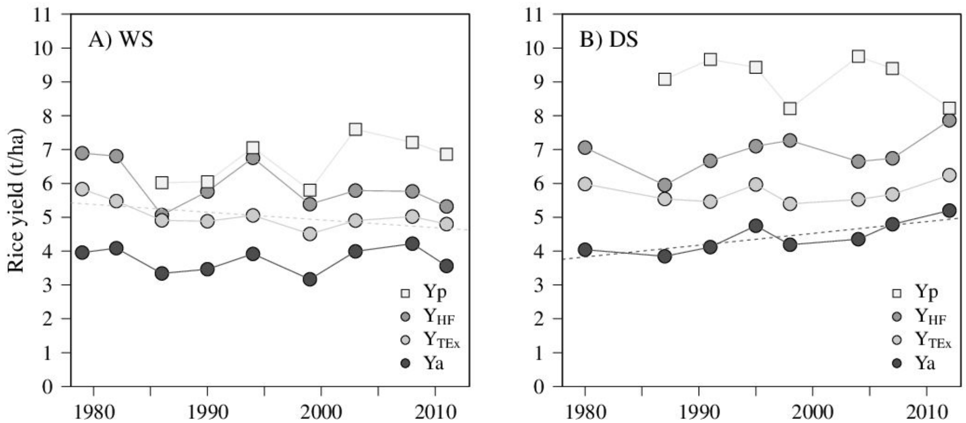

Figure 3.9: Trends in Yp, Y_{HF}, Y_{TEx} and Ya during the A) WS and B) DS for lowland irrigated rice systems in Central Luzon, Philippines. Mean Y_{HF}, Y_{TEx} and Ya were computed for all the fields included in the household survey, independently of the variety cultivated. Yp for the period 1984 - 2002 was simulated using IR72 and for the period 2003 - 2012 using NSIC Rc222; no Yp is presented prior to 1986 due to lack of weather data to perform the simulations. Statistically significant time trends at 5% significance level are presented as dashed lines.

production perspective, most likely due to accessibility issues (e.g. cash constraints) or farmers' preferences (i.e. risk avoidance vs production maximisation). Indeed, farmers may not be maximising production due to high input:output price ratios or climatic risks (e.g. typhoons) to which they are exposed. This situation is further aggravated by the lack of capital (which determine the capacity for investment in the beginning of the next growing season) and needs to be put into the context of off-farm activities and other sources of income. The contribution of these issues for rice yield gaps and resource use efficiencies will be investigated in Chapter 4.

The technology yield gap represents the highest share of the yield gap in all DS periods analysed with a mean of 2.2 t ha^{-1} (Figure 3.9B). This can be explained possibly by sub-optimal (relative to Yp) N application in highest-yielding fields and water and/or other limitations in highest-yielding fields with N application rates greater than the optimum N required to achieved Yp (Figure 3.7). Water stress may be particularly important in the DS to explain the difference between Yp and Y_{HF} because water release and distribution in Central Luzon is mostly controlled by the National Irrigation Administration (NIA) and disputes for irrigation water among municipalities in the region have been common over the past half-century (Barker and Levine, 2012). During the WS, the technology yield gap has been consistently lower than 1.0 t ha^{-1} prior to the year 2003 and it increased to ca. 1.6 t ha^{-1} in the period 2003 - 2012 (Figure 3.9B). This is most likely associated with the adoption of Mv4 varieties which have greater genetic potential and require far greater N application rates to achieve Yp (Figure 3.8). In some years, Ya values greater than simulated Yn and Yp were observed (e.g. Figure 3.7H). Possible explanations are misclassification of varieties (e.g. in our approach hybrids were classified as Mv4 varieties), particularly in 2012 DS when 25% of rice area was planted to hybrids (Laborte et al., 2015), errors occurring during the farm surveys which are filled out based on recall (e.g. unrealistic yield data given by the farmer) and/or site specific differences in biophysical conditions.

3.5.2 Methodological considerations

Explaining yield gaps in a quantitative manner, using the theoretical framework presented in Figure 3.1 is particularly demanding due to the number of interacting and confounding factors embedded in Ya and requires the application of a combination of approaches from different disciplines to individual field level data obtained through farm household surveys. Those include static and empirical models based on econometric techniques (stochastic frontier analysis, Kumbhakar and Lovell, 2000) and dynamic and mechanistic crop growth simulations models (ORYZA v3, Bouman et al., 2001). The need for combining methodologies built upon different theories poses considerable methodological challenges and can only be implemented to large and high quality datasets such as the Central Luzon Loop Survey.

We used stochastic frontier analysis to characterize the production frontier which best fitted the data and to assess the contribution of the timing component of crop management to the efficiency yield gap. Agronomists often argue that this methodology is not appropriate for assessing the limiting factor to production because of its underlying assumptions on the functional form and on perfect substitutability between agronomic inputs (Zhengfei et al., 2006; van Ittersum and Rabbinge, 1997). Our analysis is certainly subject to such limitations. However, we considered a generic functional form including linear and quadratic terms for each variable and interaction terms between variables in order to assess how individual inputs affect Ya (Table 3.1). We deliberately did not include input variables referring to land, labour and capital in Equation 3.1 since we argue that those variables explain rather than determine rice yield gaps. Instead, we included explicitly growth-defining, -limiting and -reducing factors and assess their contribution to Ya. This may not be fully compatible with the assumption of perfect input substitutability underlying (stochastic) frontier analysis since agronomic inputs (e.g. nutrients and water) cannot be directly substituted by each other as they have different biological roles in crop growth. To handle this methodological drawback, we checked whether the nutrient use efficiencies computed from the household survey were within agronomic meaningful ranges (which was the case; data not shown), which ensures that Y_{TEx} can be obtained with the observed nutrient levels. The latter was done with three quadrant plots (van Keulen, 1982) using the survey data and the internal use efficiencies, corresponding to maximum dilution (and accumulation) of N, P and K in the rice crop reported by Witt et al. (1999). In addition, the crop growth simulation model ORYZA v3 could be used to cross-validate the results obtained with frontier analysis (i.e. check the biophysical ranges of the production frontier). However, the fact that crop models only incorporate crop responses to a limited set of inputs plus their high data quality standards and requirements may hinder their application to a large set of individual farmer' fields.

The concept of Y_{AE} was only explored in this paper from a production perspective and hence defined as Y_{HF}. Input use in farmers' fields is often not optimal from a production perspective because it is related to the availability, accessibility and effective use of the inputs required. Given inputs are locally available, producers may still not access them due to capital constraints, intrinsic objectives and cultural preferences or superimposed (environmental) regulations limiting input use. Therefore, it is important to put sustainable intensification, and closure of yield gaps, into the context of the broader socio-economic environment and hence to make explicit existing trade-offs between production and other objectives. Further manipulation of the stochastic production frontier combined with optimization routines based on linear programming could be performed for this purpose.

The use of ORYZA v3 proved useful to determine the technology yield gap. Once calibrated and properly evaluated, the mechanistic nature of crop growth simulation models makes them a useful tool to assess causal relationships between yield and N management. However, this is often translated into high data requirements and standards referring to the biophysical conditions (weather and soil properties). In addition, such models assume that P and K are in optimal supply as well as no incidence of pests, diseases and weeds. These drawbacks could be overcome by following a target-oriented approach to quantify the required level of P and K (e.g. Janssen et al., 1990) in combination with expert knowledge to estimate the amount of pesticides and labour required to achieve a pre-defined target yield. Alternatively, the technology yield gap can be explained using methods of frontier analysis (cf Neumann et al., 2010) and/or on-farm trials (cf. Herdt and Mandac, 1981). Following the approach of Neumann et al. (2010), frontier analysis can be used in case farm survey data are available for farming systems across similar agro-ecological regions given that Yp is achieved in farmers' fields in at least one region. On-farm trials can be used to test alternative technologies by closing the resource yield gap of particular inputs (e.g. water, nitrogen or sowing density) in field conditions. By using this methodology, it becomes possible to compare not only yield responses to N application but also how other inputs limit crop yield.

3.6 Conclusion

A theoretical framework was introduced to explain yield gaps in lowland irrigated rice-based farming systems in Southeast Asia. The approach used recognizes three partial yield gaps, namely efficiency, resource and technology yield gaps and was tested for Central Luzon (Philippines) using the Central Luzon Loop Survey (1979 - 2012). The efficiency yield gap captures the influence of timing, spacing and form of given inputs applied. The resource yield gap is associated with the application rates of different inputs. The technology yield gap is used as benchmark for comparing technologies currently employed by farmers with agronomically optimum technologies which can achieve the climatic potential yield.

The mean yield gap estimated for the period 1979 - 2012 was 3.2 t ha^{-1} in the WS and 4.8 t ha^{-1} in the DS (Table 3.2). An average efficiency yield gap of 1.3 t ha^{-1} was estimated for rice farming in Central Luzon between 1979 - 2012, independently of the season. This was explained by crop management referring to the timing of the first and second application of mineral fertilisers and pesticides. The efficiency yield gap decreased over time; in the WS this could be attributed to a decrease of technical efficient yields over time and stagnation of actual yields while in the DS this was due to an increase in actual yields. The mean resource yield gap was small in both seasons but slightly larger in the DS (1.3 t ha^{-1}) than in the WS (1.0 t ha^{-1}). A comparison of N, P and K use across highest and lowest yielding fields indicated that application of macro-nutrients explains differences between the two groups in the DS but not in WS, when other (climatic) factors may be at stake. The technology yield gap was consistently lower than 1 t ha^{-1} in the WS periods prior to 2003 and ca. 1.6 t ha^{-1} between 2003 - 2012. In all DS periods, it represented the largest share of the rice yield gap with a mean of 2.2 t ha^{-1}. Varietal shift and sub-optimal application of inputs (e.g quantity of water and N) are the most plausible explanations for this yield gap during the WS and DS, respectively.

Further research is necessary to understand the importance of farmers' objectives, farm resources and socio-economic conditions for rice yield gaps and resource use efficiency in Central Luzon, Philippines. This should provide further insights about possible trade-offs between different resource allocation strategies (e.g. maximising production vs minimising risk) as well as the relationship between crop manage-

Table 3.2: Decomposing the rice yield gap into efficiency, resource and technology yield gaps in Central Luzon, Philippines (1979 - 2012). Data is presented in t ha^{-1}. The signs $<$, $>$ and $=$ stand for significant decreasing, increasing or constant yield gap values over time, respectively.

	Wet season	Dry season	Remarks
Efficiency Yg	1.3 ($<$)	1.3 ($<$)	Untimely application of fertilisers and pesticides
Resource Yg[a]	1.0 ($<$)	1.3 ($=$)	Sub-optimal quantity of N, P and K in the dry season
Technology Yg[b]	0.9 ($>$)	2.2 ($=$)	Varietal shift and sub-optimal quantity of inputs
Rice Yg	3.2 ($>$)	4.8 ($<$)	

[a] Data of the resource yield gap refer to the multiple-input stochastic frontier.
[b] Estimations of the technology yield gap refer to the period 1986 - 2012 only.

ment and land, labour and capital. Last but not least, it remains unclear what are the farm-level implications of closing rice yield gaps in terms of rice self-sufficiency and household income. Insight on this is necessary if one is interested in explaining yield gaps from a socio-economic perspective and in understanding the potential contribution of sustainable intensification for agricultural development.

3.7 Acknowledgements

We are grateful to 1) Ms. Piedad Moya (IRRI SSD), Ms. Esther Mariano (IRRI SSD) and Ms. Marie-Rose San Valentin (IRRI SSD) for making the Central Luzon Loop Survey readily available for analysis and discussing preliminary findings, 2) Dr. Grigorios Emvalomatis (University of Dundee) for helping in the estimation of the stochastic frontier models, 3) Dr. Tao Li (IRRI CESD), Mr. Manuel Marcaida III (IRRI CESD) and Dr. Pepijn van Oort (AfricaRice) for helping with the parametrization of ORYZA v3, 4) Mr. Ferdinand Corcuera (IRRI CESD), Dr. James Quilty (IRRI ES) and the farmers included in the loop survey for their help during the soil survey and, 5) the IRRI GIS Team, Dr. Roland Buresh (IRRI CESD), Mr. Teodoro Correa Jr (IRRI CESD), Dr. Glenn Gregorio (IRRI PBGB) and Ms. Alpha Grace Tabanao (IRRI PBGB) for data sharing. This research is part of the Wageningen UR strategic programme 'Mapping for Sustainable Intensification' and was funded by Wageningen UR and the Global Rice Science Partnership (GRiSP).

3.8 Appendix

Detailed results about the statistical tests conducted to select the stochastic frontier model which best fitted the data and about the calibration and evaluation of ORYZA v3 are provided as Supplementary Material to the online version of this manuscript (https://doi.org/10.1016/j.eja.2016.06.017). The Central Luzon Loop Survey, including the soil survey conducted in 2014, can be freely accessed in http://ricestat.irri.org/fhsd/php/panel.php (household survey) and http://ricestat.irri.org/fhsd/php/survey.php (soil survey).

CHAPTER 4

Intensification of rice-based farming systems in Central Luzon, Philippines: Constraints at field, farm and regional levels

This chapter is under review as:
Silva, J.V.; Reidsma, P.; Velasco, M.L.; Laborte, A.G.; van Ittersum, M.K. 2017. Intensification of rice-based farming systems in Central Luzon, Philippines: Constraints at field, farm and regional levels. *Agricultural Systems*.

Abstract

Understanding the opportunities for sustainable intensification requires an integrated assessment at field, farm and regional levels of past developments. Two hypotheses regarding current rice production in Central Luzon (Philippines) were developed for this purpose. First, we hypothesize that there are trade-offs between rice yields, labour productivity, gross margin and N use efficiency and, second, that farm(er) characteristics and socio-economic conditions at farm and regional level affect the management practices used by farmers. These hypotheses were tested using two household surveys characterizing rice-based farming systems in Central Luzon in terms of changes over time (1966 - 2012) and spatial variability. Over the past half-century there was an increase in the proportion of irrigated fields and adoption of improved varieties, which allowed the cultivation of a dry season rice crop in Central Luzon. Moreover, transplanting has been replaced by direct-seeding and herbicides substituted hand-weeding. These resulted in greater rice yields and labour productivity, and contributed to gradual transition from subsistence to commercial farming systems, as observed in the increasing proportion of hired labour and rice sold. Our results indicate the existence of a trade-off between rice yields, labour productivity and N use efficiency as yield levels maximising labour productivity and NUE were ca. 25% and 35% lower than climatic potential yield in the wet and dry season, respectively. At field level, this can be explained by 1) the use of transplanting as crop establishment method, which resulted into higher yields but lower labour productivity as compared to direct-seeding, and 2) the high N application levels, which led to higher yields but lower N use efficiency. In contrast, yield levels which maximised gross margin were ca. 80% of the climatic potential in both wet and dry seasons, so there was little trade-off between rice yields and economic performance. Regarding the second hypothesis results were not always conclusive. As an example, N application per ha was negatively associated with farm size and the timing of the first fertiliser application was positively associated with household size and with the number of parcels. More intensive practices, and better farm performance, were recorded in the province at the heart of the irrigation system. We thus conclude that closing rice yield gaps in the production systems of Central Luzon incurs trade-offs with environmental and social objectives at field and farm levels but less with economic objectives. However, we could not clearly show whether, and to what extent, management practices used by farmers are influenced by farm or regional level constraints.

4.1 Introduction

Sustainable intensification has been proposed as a strategy to raise productivity and resource use efficiency, through its focus on yield potential, soil quality and precision agriculture (Tilman et al., 2011; Cassman, 1999). This is particularly relevant in developing countries where large yield gaps persist (Tittonell and Giller, 2013; Laborte et al., 2012). Understanding the opportunities for sustainable intensification requires analysis of possible constraints at field, farm and regional levels. At field level, crop management packages need to fit farmers' needs not only in terms of high yields and resource use efficiencies, but also labour productivity and profitability. At farm and regional levels, structural changes may be needed to alleviate farmers' resource constraints (in terms of land, labour and capital), to ensure equitable natural resource management and to reduce the biophysical and economic risk of farming.

Yield gap analysis of rice-based farming systems in Central Luzon, Philippines, revealed that crop establishment method (transplanting and direct-seeding), nitrogen (N), potassium (K) and seed use were significantly related with rice yields (Chapter 3). Moreover, the timing of the first and second applications of fertiliser and pesticides explained part of the variation in the efficiency yield gap, i.e. the difference between technical efficient and actual yields. Technology yield gaps, defined as the difference between climatic potential and highest farmers' yields, were mostly attributed to sub-optimal water and nutrient use as compared to input requirements to achieve the climatic potential yield. From an agronomic perspective, overcoming these limiting factors requires the use of 'management packages', which can achieve yield levels close to the climatic potential yield while ensuring high resource use efficiency (Savary et al., 2012; Dobermann et al., 2002; Bouman et al., 2001). Moreover, these management packages need to suit the livelihoods of smallholder farmers.

An integrated assessment of agricultural systems considering drivers at different scales can be used as a framework to identify constraints to sustainable intensification. This requires developing indicators, which capture different dimensions of smallholder farming systems and exploring possible synergies, and trade-offs, among those indicators (Klapwijk et al., 2014). In this way, it is possible to assess whether technically feasible management packages for yield gap closure fit the needs and characteristics of smallholder farmers, the farm level consequences of using a particular management package, and the ideal management package to narrow the yield gap considering multiple objectives and resource constraints. In addition, it is necessary to understand how factors at farm and regional levels interact with crop management practices (e.g. resource availability, farm(er) characteristics and socio-economic conditions), as these define the 'operating space' for agricultural production (for Asian agriculture see e.g. Studwell, 2013; Takahashi and Otsuka, 2009; Estudillo and Otsuka, 1999; Ledesma, 1980).

The objective of this paper is to identify constraints and stimuli to intensification of rice-based farming systems in Central Luzon, Philippines at field, farm and regional levels. The two research questions we aim to answer are: 1) are there trade-offs between rice yields, labour productivity, gross margin and N use efficiency for the management packages used by farmers, and if so, what are their magnitudes, and 2) how do farm(er) characteristics and socio-economic conditions at farm and regional levels affect farmers' management practices and the aforementioned indicators? For these purposes, we analysed a panel household survey of about 100 rice farmers during 1966 - 2012 (Central Luzon Loop Survey, Moya et al., 2015) and an unpublished cross-sectional household survey of 1,800 rice farmers conducted in 2013 - 2014 (Metrics and Indicators for Tracking in GRiSP) in Central Luzon, Philippines.

4.2 Conceptual framework

For irrigated conditions, the yield gap can be defined as the difference between climatic potential yield (Yp) and actual yields (Ya). Farmers can reduce yield gaps through improved crop management but this may not be their only objective. Moreover, yield gaps and associated management packages need to be evaluated in terms of different indicators and explained by farm and regional conditions (Figure 4.1). We formulated two hypotheses for this purpose, which are explained below.

The first hypothesis is that management practices associated with small yield gaps perform sub-optimally in terms of labour productivity, gross margin and N use efficiency (cf. van Ittersum et al., 2013; Lobell et al., 2009; Cassman et al., 2003). Trade-offs between indicators are partially the result of the management packages used (step 1 in Figure 4.1). In case of rice, modern varieties can achieve higher Ya (and Yp) than traditional varieties but these also have more labour requirements for crop management, harvesting and threshing (Estudillo and Otsuka, 1999). Herbicide use can result in greater Ya and labour productivity, as compared to hand-weeding, as it allows for more timely control of weeds due to lower labour requirements. As for crop establishment, significantly higher Ya was observed for transplanted compared to direct-seeded rice in Chapter 3, while direct-seeding requires less labour for crop establishment but more labour for weed control (Pandey and Velasco, 2002). These examples illustrate how interactions between management practices at field level can affect performance at farm level.

We expect that closing yield gaps is associated with higher revenues, costs, labour requirements and fertiliser application levels (step 2 in Figure 4.1). Trade-offs between maximising gross margin and minimising yield gaps are likely to occur when output prices are low and input prices are high. In addition, closing yield gaps requires greater amount of labour for crop management activities such as fertilisation, weeding, pest control and harvesting but adoption of capital intensive, labour-saving, technologies (e.g. tractors, herbicides and mechanical threshers) can mitigate potential trade-offs between high Ya and high labour productivity. Finally, closing yield gaps in lowland irrigated rice systems requires relatively high amounts of N (150 -

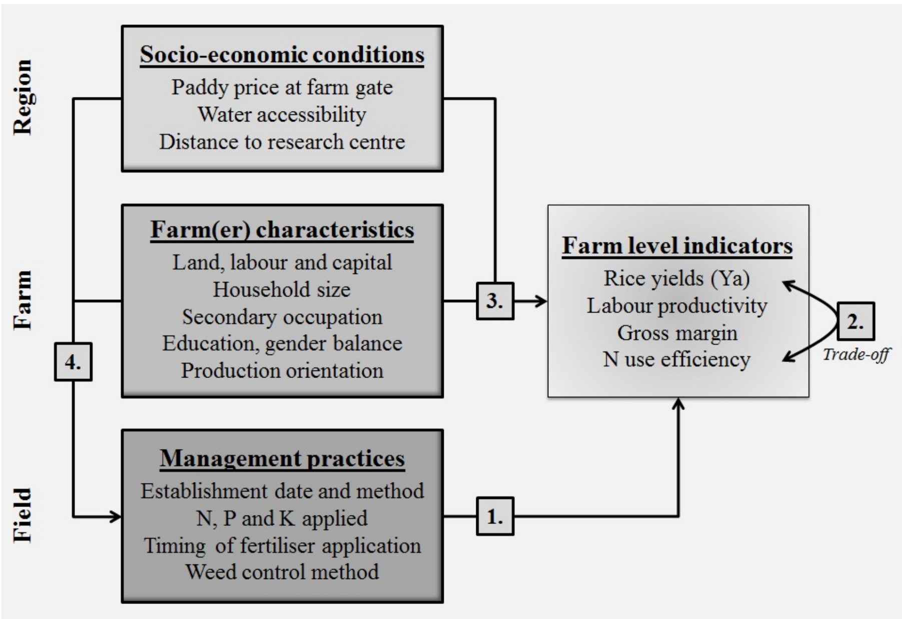

Figure 4.1: Integrated assessment of rice yield gaps and associated management practices in Central Luzon, Philippines. The relationship between management practices and rice yields is detailed in Chapter 3. Temporal dynamics are not explicitly represented in this figure but were taken into account throughout the analysis. In this paper, we assess how different management packages perform in terms of different indicators (step 1), the existence of trade-offs between closing yield gaps and optimising other farmers' objectives (step 2), and how farm level determinants and regional conditions affect rice yield gaps and farm performance (step 3) and management practices (step 4).

200 kg N ha^{-1}, Dobermann et al., 2000), which, following the law of diminishing returns, results in sub-optimal N use efficiency (Cassman et al., 1998; Kropff et al., 1993) and increases the risk of crop failure due to lodging (Lampayan et al., 2010).

The second hypothesis is that management practices used by farmers, and hence the variability observed in farm level indicators, can be explained by the availability of farm resources (land, labour and capital), farm(er) characteristics and regional conditions (Villano et al., 2015; Takahashi and Otsuka, 2009; Estudillo and Otsuka, 1999; Kerkvliet, 1990). This is depicted in steps 3 and 4 in Figure 4.1. For instance, according to Erguiza et al. (1990) supervision of hired labour, cost of transplanting labour and use of credit were among the most important determinants of the crop establishment method used by rice farmers in Nueva Ecija. In Central Luzon, farmers adopted direct-seeding and increased mechanisation (Moya et al., 2004) perhaps due to the increasing importance of capital and non-rice income in the region (Takahashi and Otsuka, 2009). We expect non-rice income to be positively associated with fertiliser use because farmers are more likely to purchase fertilisers if more capital is available. Moreover, we anticipate that smaller farm size and greater labour availability lead to more timely crop establishment and fertiliser application dates. Finally, regional conditions can affect the land preparation and crop establishment dates because they determine the pattern of water release and availability (Tabbal et al., 2002; Loevinsohn et al., 1993). They may also affect the farm level indicators due to differences in water accessibility (Barker and Levine, 2012) and implementation of land reform programs (Ledesma, 1980).

Understanding the constraints to intensification of crop management requires an integrated assessment considering drivers at different scales because farm and regional level conditions are likely to influence the management practices adopted by farmers, and their performance. As constraints are not static, temporal trends should also be analysed (Falconnier et al., 2015; Valbuena et al., 2015; Takahashi and Otsuka, 2009; Iráizoz et al., 2003).

4.3 Material and methods

4.3.1 Household surveys

Data from two different household surveys conducted by the International Rice Research Institute (IRRI) in Central Luzon were used in this study (Figure 4.2). Double rice cropping is common in the region with a wet season (WS) crop cultivated

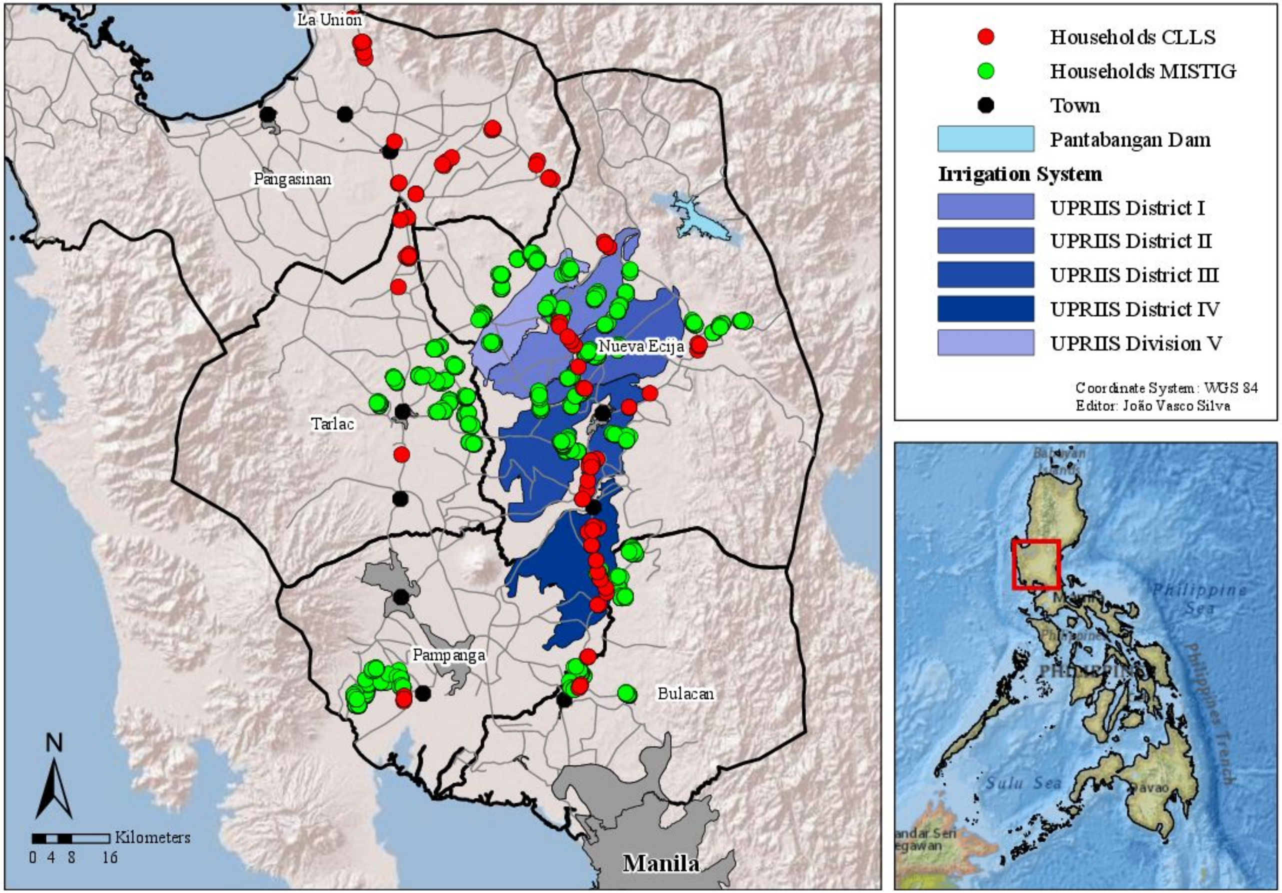

Figure 4.2: Map of the study region with locations of the households surveyed and the Upper Pampanga River Integrated Irrigation System (UPRIIS). Note that La Union and Pangasinan are not part of Administrative Region 3 (Central Luzon) but are geographically situated in the central part of Luzon.

between June - July and September - October and a dry season (DS) crop cultivated between December - January and March - April. The WS crop is cultivated with or without irrigation while the DS crop is only possible with irrigation.

The Central Luzon Loop Survey (CLLS) was collected at farm level every 4 to 5 years during 1966 - 2012 with the objective of monitoring changes over time in crop management and household characteristics in rice-based farming systems (Moya et al., 2015). The survey covers the provinces of Bulacan, La Union, Nueva Ecija, Pampanga, Pangasinan and Tarlac. The rice fields surveyed were selected at a specific distance along the main highway. On average, 103 and 59 households cultivating rice were interviewed, respectively, in the WS and DS. The sample size in the DS is lower than in the WS because fewer households cultivated rice during the former, due to water-related constraints and/or cultivation of other crops. Trends in crop management, farm performance indicators and farm/regional level conditions can be found in Figure 4.3.

A more comprehensive household survey was conducted in 2014 under a project that aimed to provide 'Metrics and Indicators for Tracking in GRiSP' (MISTIG, GRiSP stands for Global Rice Science Partnership). A three-stage sampling procedure was used to identify the households to be surveyed in the top four rice producing provinces of Central Luzon (Bulacan, Nueva Ecija, Tarlac and Pampanga). Fifteen municipalities with more than 2,000 ha of total rice area were randomly selected with the number of municipalities selected per province proportional to the rice area in the province. In each municipality four villages were randomly selected. In each village, twenty farmers were identified using a systematic sampling procedure, which resulted in a sample size of 1,800 rice farming households. A semi-structured questionnaire was administered using tablets to gather crop management information for the largest rice parcel and farm level conditions referring to the 2013 DS and 2014 WS.

4.3.2 Farm level indicators

Four principles were used to describe different dimensions of rice farming in Central Luzon relevant for farmers and regional stakeholders. Those were rice production, farming drudgery, economic viability and environmental sustainability. There are slight differences in the calculation of the indicators depending on the household survey because the CLLS contains information for all plots cultivated by a household (60% of the households cultivated more than one field per season, out of which 82% cultivated two fields) while the MISTIG survey contains information only for the largest parcel cultivated by each household. The importance of each principle and the calculation of the different indicators are explained below.

Rice production. Increasing rice production is important not only to obtain greater economic returns but also to meet rice self-sufficiency requirements at household level. Rice production is represented by rice yield (Ya, t ha^{-1}). In case of the CLLS, in which data for multiple fields in a single farm are available, Ya was calculated as the quotient between rice production at 14% moisture content and the sum of the area of the individual rice fields cultivated by each single farm. For the MISTIG survey, Ya refers to the yield level reported by each farmer for his/her largest parcel and this was standardized at 14% moisture content. Climatic potential yields (Yp) were simulated with ORYZA v3 for the period 1984 - 2012 using local weather and soil data and taking into account farmers' sowing dates and changes in rice varieties (Chapter 3).

Farming drudgery. The drudgery of farming is represented by the indicator labour productivity (kg rice labour-day^{-1}, ld). Increasing labour productivity in Central Luzon has been achieved partly through adoption of labour-saving technologies, which contribute to a reduction of the tedious and exhausting farm work and frees household members to engage in off-farm activities. Labour productivity was calculated for the two household surveys as follows:

$$\text{Total Labour}_i \text{ (ld ha}^{-1}\text{)} = \sum_{j}^{6} \text{Labour}_{i,j} \text{ (ld ha}^{-1}\text{)} \tag{4.1}$$

$$\text{Labour productivity}_i \text{ (kg ld}^{-1}\text{)} = \frac{\text{Rice yield}_i \text{ (kg ha}^{-1}\text{)}}{\text{Total Labour}_i \text{ (ld ha}^{-1}\text{)}} \tag{4.2}$$

where i stands for an individual farm in case of the CLLS, or field in case of the MISTIG survey, and j stands for management operations (i.e. land preparation, crop establishment, fertiliser application, pesticide/herbicide application, hand-weeding and harvesting/threshing). Data for hand-weeding were not available in the MISTIG survey and hence, labour productivity values may be over-estimated from the available data. 'Labour' includes both family and hired labour.

Economic viability. Rice farming is an important activity in Central Luzon and its economic viability, as a source of income for rural households, depends on the returns provided by rice cultivation. For example, assuming a gross margin maximising farmer behaviour, it is likely that a low input-output price ratio will stimulate closure of (resource) yield gaps in farmers' fields while a high input-output price ratio will hinder the use of inputs, resulting in an economic optimum yield below Yp (van Dijk et al., 2017; Koning et al., 2008). Thus, gross margin (PhP ha^{-1}) was used as an economic indicator and estimated as follows:

$$\text{Revenue}_i\ (\text{PhP ha}^{-1}) = \text{Rice Kept}_i\ (\text{kg ha}^{-1}) \times \text{Rice Price}_i\ (\text{PhP kg}^{-1}) \quad (4.3)$$

$$\begin{aligned}\text{Total Cost}_i\ (\text{PhP ha}^{-1}) = & \sum_{m}^{4} \text{Materials}_{i,m}\ (\text{PhP ha}^{-1}) + \\ & + \sum_{j}^{6} \text{Labour}_{i,j}\ (\text{PhP ha}^{-1}) + \\ & + \text{Land Rent}_i\ (\text{PhP ha}^{-1}) \end{aligned} \quad (4.4)$$

$$\text{Gross Margin}_i\ (\text{PhP ha}^{-1}) = \text{Revenue}_i\ (\text{PhP ha}^{-1}) - \text{Total Cost}_i\ (\text{PhP ha}^{-1}) \quad (4.5)$$

where i stands for an individual farm in case of the CLLS, or field in case of the MISTIG survey, m for the materials used in rice production (i.e. seeds, fertilisers, pesticides and water) and j stands for management operations (i.e. land preparation, crop establishment, fertiliser application, pesticide/herbicide application, hand-weeding and harvesting/threshing). Data for hand-weeding were not available in the MISTIG survey but this may hardly affect gross margin estimates because costs of family labour were not considered in the calculation (see below). 'Rice Kept' corresponds to the total rice harvested minus the amount of rice used to pay in kind the harvest and threshing operations per unit area and 'Paddy Price' is the unitary rice price received by each household. Historical prices reported in the CLLS were standardized to 2005 prices using the consumer price index (Central Bank of the Philippines, consulted on 19-10-2014) while prices from the MISTIG survey are reported 2014 prices. Our estimates of gross margin correspond to a 'best case scenario' because 1) the quantity of rice sold by each household is generally lower than the quantity of rice kept after harvest, as households may save rice for seed or may store it for home consumption and 2) family labour was not explicitly considered as a cost in the analysis as we assume farmers have a non-economic incentive to grow rice (e.g. rice self-sufficiency at household level).

Environmental sustainability. Nitrogen use efficiency (NUE) is one relevant indicator to assess the environmental sustainability of rice farming because rice cultivation in the lowland areas of Southeast Asia is an intensive activity, with two to three rice crops grown in a single year (Dobermann et al., 2002). Moreover, its long-term productivity and sustainability depend heavily on the use of mineral fertilisers, particularly N (Pampolino et al., 2008; Dobermann et al., 2000). For a given N input, a low NUE points to high environmental losses while a high NUE reduces this risk. NUE (kg N output kg^{-1} N input) was estimated as follows:

$$\text{N Output}_i \text{ (kg N ha}^{-1}) = \text{Rice Yield}_i \text{ (kg ha}^{-1}) \times \text{N concentration (\% N)} \quad (4.6)$$

$$\text{N Input}_i \text{ (kg N ha}^{-1}) = \text{Fertiliser N}_i \text{ (kg N ha}^{-1}) + \text{Indigenous N (kg N ha}^{-1}) \quad (4.7)$$

$$\text{N use efficiency}_i \text{ (kg N kg}^{-1} \text{ N)} = \frac{\text{N Output}_i \text{ (kg N ha}^{-1})}{\text{N Input}_i \text{ (kg N ha}^{-1})} \quad (4.8)$$

where i stands for an individual farm in case of the CLLS, or field in case of the MISTIG survey. We assumed 'N concentration' to be 1.1% in the rice grains (Witt et al., 1999). 'N input' included N application by farmers as mineral fertiliser and 'Indigenous N' from irrigation sediments, rain dust and biological N fixation (ca. 50 kg N ha^{-1} $crop^{-1}$; Dobermann, 2000). NUE values ranging between 0.6 - 0.7 kg N kg^{-1} N are considered optimal in irrigated lowland rice farming systems (Haefele et al., 2008, 2003; Witt et al., 1999). However, values above this range can be observed at very low N application rates or in very efficiently managed systems (Dobermann, 2005).

4.3.3 Step 1: Performance of management packages

Management packages used by rice farmers were summarized from the CLLS and their performance in terms of yield, labour productivity, gross margin and N use efficiency (step 1 in Figure 4.1) was assessed using linear mixed models.

Identification of management packages

The management packages used by farmers during the past half-century were identified through consecutive subsets of the CLLS. We analysed management packages rather than individual practices to take into account interactions between different management practices (e.g. Lampayan et al., 2010; Lantican et al., 1999; Mandac and Flinn, 1985). First, irrigated fields were separated from rainfed fields to capture differences in water management. Second, fields cultivated with traditional (Tv) or modern rice varieties (Mv) were differentiated. Further distinction within Mv was not done because of little temporal overlap between Mv3 and Mv4 type of varieties (Figure 4.3C; see Launio et al., 2008, for a definition of Mv3 and Mv4). Third, the crop establishment method was used to differentiate between transplanting and direct-seeding. Fourth, weed management was used to separate fields in which hand-weeding was performed or herbicides were applied. Finally, fields with 'High N applied' (i.e. with N application equal or higher than the overall mean N application level) were separated from fields with 'Low N applied' (i.e. lower than the overall mean N application). The timing of management operations was not considered in this categorization of the management packages for the sake of simplicity. Descriptive statistics of input use, farm resources and farm(er) characteristics per management package are provided in Tables B3 and B4 (Supplementary Material).

Linear mixed model

The performance of management packages (with at least 10 observations) in terms of the different performance indicators was assessed using a linear mixed model, considering management package and province as fixed effects and year within household as random effects. Province effects were included to capture differences in access to irrigation, markets (e.g. distance to the capital) and innovations (e.g. distance to research center). Year within household was included to capture both year-specific climatic conditions and trends in technology use for each household. Season effects were taken into account by fitting the same model to the WS and DS data separately to control for biophysical differences (e.g. solar radiation and rainfall). The linear mixed models were fitted using the R package *nlme* (Pinheiro et al., 2016) and predicted means and standard errors for each management package were obtained with the R package *predictmeans* (Luo et al., 2014). Differences were assessed at 5% significance level.

4.3.4 Step 2: Trade-off analysis between indicators

Trade-offs between indicators (step 2 in Figure 4.1) were firstly assessed by comparing best-performing farms with respect to each indicator, and secondly by comparing all farms using boundary line analysis.

Comparison of best-performing farms

The top 10^{th} percentile of Ya, labour productivity, gross margin and NUE were used to identify different 'farm types', which we defined as the farms following a specific objective. For instance, the top performing farms in terms of rice yield (i.e. the farms with Ya above the 90^{th} percentile of this indicator) were classified as 'production maximising farms'. 'Labour productivity maximising farms', 'gross margin maximising farms' and 'NUE maximising farms' were defined following the same approach. One-way analysis of variance (ANOVA) and Tukey's test were used to assess statistically significant mean differences in farm level indicators between the four farmers' objectives considered. This analysis was done with the CLLS for the WS and DS separately, using pooled data from all the years.

Boundary-line analysis

The relationship between Ya and the other indicators was studied using boundary line analysis (Fermont et al., 2009; Shatar and McBratney, 2004; Schnug et al., 1996; Webb, 1972). The underlying reason for using this benchmarking method was that fields/farms with similar Ya may have different labour productivity, gross margin or NUE due to differences in e.g. mechanisation levels, input-output prices and N management, respectively. Therefore, this method was applied to 1) identify the optimal level of an indicator for each given Ya, 2) estimate the Ya with the maximum level of each performance indicator and 3) assess how differently farms with similar Ya performed in terms of the different indicators. In this way, we can obtain quantitative insights into synergies and trade-offs between increasing Ya and optimising other indicators using individual farm data. The boundary lines were estimated for the WS (irrigated and rainfed) and the DS (irrigated) data separately in three consecutive steps, which are described in the Appendix (Section 4.8). We used the MISTIG survey in this analysis to avoid confounding effects of technological change over time and because it is the most recent, representative and comprehensive household survey available for the production system studied. Data manipulation and model fitting were performed using R software (R Core Team, 2013).

4.3.5 Steps 3 & 4: Drivers of indicators and crop management

The drivers of farm level indicators (step 3) and crop management practices (step 4 in Figure 4.1) were analysed using multiple linear regressions applied to the CLLS with pooled data over different years. The intercept and coefficients of the different models were estimated using ordinary least squares (OLS). The dependent variables considered were Ya (t ha^{-1}), labour productivity (kg ld^{-1}), gross margin (PhP ha^{-1}) and NUE (kg N kg^{-1} N), as well as N use (kg N ha^{-1}), timing of first fertiliser application (days after sowing, DAS) and crop establishment date (day of the year, DOY). The three management variables were aggregated at the farm level using the mean across different fields. The aforementioned management practices were selected because they explain the resource and efficiency yield gaps (Chapter 3). The independent variables were chosen to capture a range of farm(er) characteristics and socio-economic conditions, as specified in Table 4.1. Endogenous variables to each indicator were removed prior to model estimation (e.g. paddy price and gross margin). Model estimation was done for the WS and DS separately using the lm function of R software, after checking for correlations between independent variables. All variables were ln-transformed so that parameter estimates can be interpreted as elasticities, i.e. percentage change of the dependent variable in case of a 1% change in the independent variable. Data on off-farm and other income, which together are defined as non-rice income, were not collected for the years 1967 DS, 1971 DS, 1974 WS, 1975 DS, 1980 DS, 1995 DS, 1998 DS, 1999 WS and 2003 WS (Figure 4.3), hence observations from these years were excluded from the regression analysis.

In case of categorical management factors (e.g. crop establishment method), comparisons of means of farm and regional variables across groups of farms with contrasting management practices were performed using t-tests. Farm and regional variables were ln-transformed prior to the analysis to homogenize variances between the different groups. Further group comparisons between farms using herbicide or hand-

Table 4.1: Description of the independent variables used in the multiple regression analysis.

Variable	Unit	Description
Farm size	ha	Size of land cultivated with rice
Parcels	no.	Number of parcels cultivated with rice
Rented	%	Proportion of rented land cultivated with rice
Shared	%	Proportion of sharecropped land cultivated with rice
Hired labour	%	Proportion of hired labour for rice production
Power	factor	Source of farm power used for land preparation; levels = Animal, Animal & Tractor, Tractor
Off-farm income	factor	Income derived from off-farm employment (Y/N)
Other income	factor	Income derived from remittances, pension, rentals (Y/N)
Household size	no.	Total number of household members
Gender	%	Proportion of males in each household
Education	yr	Years of formal education of the household head
Members	no.	Number of household members farming their own land, working as agricultural labourers or working in the non-agricultural sector
Rice sold	%	Proportion of total rice production which is sold
Paddy price	PhP kg^{-1}	Unitary price of unmilled rice received by each household
Province	factor	Province in which household is located (cf. Figure 4.2)
Year	factor	Year in which the survey was conducted (cf. Figure 4.3)

weeding, and all macro-nutrients (N, P and K) or some macro-nutrients (N, P or K) were performed to gain insights into the resource yield gap. No group comparisons were performed between Tv and Mv because there is hardly any temporal overlap between these types of varieties (Figure 4.3C). Only significant results are reported but non-significant results are provided in Tables B5 and B6 (Supplementary Material).

4.4 Results

4.4.1 Temporal changes in rice farming

The most remarkable temporal changes in management practices, performance indicators, farm(er) characteristics and regional conditions over the past half-century are summarized in Figure 4.3. It is worth noting that the time period covered by the CLLS starts in the year of release of the semi-dwarf variety IR8, which spearheaded the Green Revolution in Southeast Asia. This contributed to a series of structural transformations in rice-based farming systems in the region, and particularly in Central Luzon, the rice bowl of the Philippines.

The opening of the Pantabangan dam in 1978 contributed to an increase in the proportion of farms using irrigation in the WS and allowed the cultivation of a second rice crop during the DS (which would not be possible without irrigation and the availability of short cycle modern varieties; Figure 4.3A). Traditionally, farmers used manual transplanting of rice under flooded conditions to establish the crop but that was gradually replaced by direct-seeding in the beginning of the 1980s (Figure 4.3B). This shift occurred particularly in the DS and only marginally in the WS. However, there was a sharp increase again in the proportion of farms transplanting rice between 2007 DS and 2012 DS. The adoption of modern rice varieties occurred continuously and without major setbacks in both seasons (Figure 4.3C): the photo-period sensitive Tv cultivated in the late 1960s were fully replaced by semi-dwarf Mv1 in the early 1970; these were replaced by pest- and disease-resistant Mv2 in the late 1970s, which were further replaced in the late 1980s by Mv3 that have better grain quality and, finally, Mv4 (that have higher Yp) became dominant in the early 2000s. There was a sharp increase in the use of herbicides, i.e. less than 10% of the surveyed farms used herbicides in 1970 while nearly 80% used herbicides after 1990 (Figure 4.3D).

Rice yields increased over time in both seasons: Ya was around 2 t ha^{-1} in 1966 WS and 1967 DS, 3.6 t ha^{-1} in 2011 WS and 5.2 t ha^{-1} in 2012 DS (Figure 4.3E). A statistically significant yield progress was observed in the DS, while in the WS yield progress was observed up to 1990 and afterwards yields stagnated at ca. 3.8 t ha^{-1}.

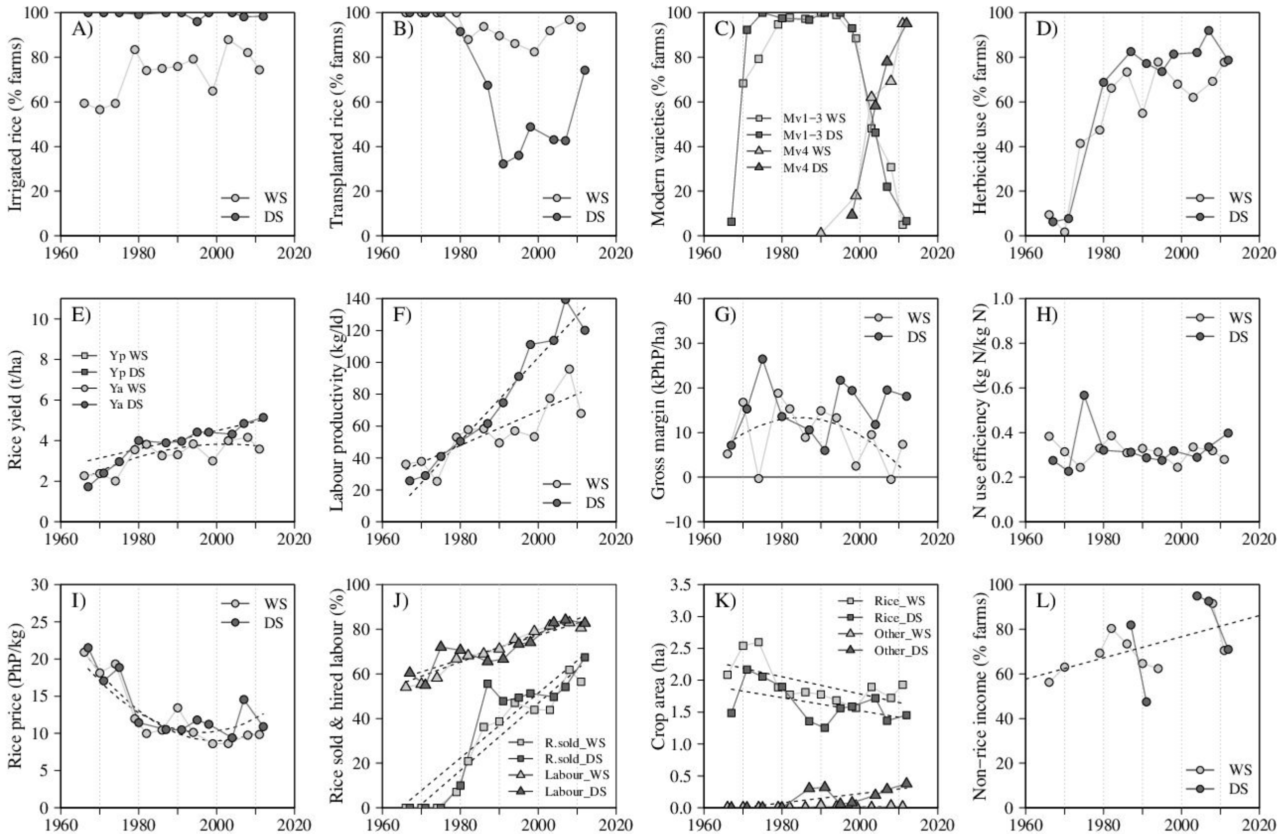

Figure 4.3: Temporal changes in crop management practices (A - D), performance indicators (E - H) and farm(er) characteristics and regional conditions (I - L) over the past half-century in rice-based farming systems in Central Luzon, Philippines. Further details about the simulation of Yp in E) can be found in Chapter 3. Dashed lines show statistically significant linear and quadratic regressions and paddy prices were standardized to 2005 prices using the consumer price index. Temporal changes in household characteristics are provided in Supplementary Material (Figure B3). WS = wet season; DS = dry season. *Data source*: Central Luzon Loop Survey.

Yield gaps were on average 3.0 and 4.6 t ha^{-1} in the WS and DS, respectively. Labour productivity increased at a rate of 1.1 kg ld^{-1} per year in the WS and of 2.6 kg ld^{-1} per year in the DS because of yield increases and adoption of labour-saving technologies (Figure 4.3F). The average labour productivity observed after 2000 was ca. 70 kg ld^{-1} in the WS and 120 kg ld^{-1} in the DS. Mean gross margin from rice farming in the WS increased up to 13.5 kPhP ha^{-1} (ca. 250 € ha^{-1}) in 1985 and sharply declined afterwards, while there was no statistically significant trend for gross margin in the DS (Figure 4.3G). No statistically significant changes over time were observed for NUE either, which remained ca. 0.3 kg N kg $^{-1}$ N in both seasons (Figure 4.3H).

Major changes also occurred in farm(er) characteristics and socio-economic conditions. Paddy prices halved from ca. 20 PhP kg^{-1} in 1966 -1967 to ca. 10 PhP kg^{-1} from 1998 - 1999 onwards and the time trends were fairly similar for both seasons (Figure 4.3I). There was a significant linear increase in the proportion of rice sold over time in both seasons: in 1979 - 1980 farmers sold about 5% of their total rice production, which increased to nearly 60% in 2011 - 2012 (Figure 4.3J). The area cultivated with rice declined by about 0.5 ha between 1966 - 2012 (partly due to land conversion to residential areas) while there was a small increase in the area of other crops (e.g. maize and vegetables) during the DS (Figure 4.3K). Hired labour as a percentage of total labour increased linearly from about 60% to 80% in both seasons indicating an increasing importance of hired labour compared to family labour over time (Figure 4.3J). Hence, rice farmers in Central Luzon shifted their production orientation over the past half-century from subsistence to market-oriented. Finally, the proportion of farms with access to non-rice income tended to increase over time due to job opportunities available elsewhere in the economy as well as remittances (Figure 4.3L).

4.4.2 Step 1: Performance of management packages

An overview of the most commonly used management packages over the last half-century is provided in Figure B2 (Supplementary Material). In short, the package that was mostly used in the WS was 'Irrigation + Modern Variety + Transplanting + Herbicide + N, P and K'. The only difference observed in the DS was the use of direct-seeding instead of transplanting as a crop establishment method starting in the late 1980s. Differences between management packages in terms of Ya (Table 4.2) are mostly due to N applied and the use of transplanting which is similar to results obtained in Chapter 3. The most important driver of labour productivity was crop establishment method, particularly in the DS, as direct-seeding resulted in significantly

Table 4.2: Comparison of farm level indicators across different management packages used by rice farmers in Central Luzon, Philippines, 1966 - 2012. Details per management package are provided in Supplementary Material (Tables B3 and B4). Codes: 'Mean' = predicted means, 'SE' = standard error of the means, 'Signif.' = letter-based representation of pairwise comparisons per season at 5% significance level (no overlap in letters between pairs means a statistical difference), 'Mv' = modern variety, 'Tv' = traditional variety, 'DSR' = direct-seeding, 'TPR' = transplanting, 'Hand' = hand-weeding, 'Herb' = herbicide use, 'highN' = high N applied, 'lowN' = low N applied. *Data source*: Central Luzon Loop Survey.

Wet season	**Rice yield (Ya) (t ha^{-1})**			**Labour productivity (kg ld^{-1})**			**Gross margin (kPhP ha^{-1})**			**N use efficiency (kg N kg^{-1} N)**		
	Mean	**SE**	**Signif.**	**Mean**	**SE**	**Signif.**	**Mean**	**SE**	**Signif.**	**Mean**	**SE**	**Signif.**
Irrigated, Mv												
DSR_Hand_highN	3.30	0.34	abcdef	70.45	7.68	abc	10.06	6.08	a	0.22	0.04	ab
DSR_Herb_highN	3.07	0.20	abde	69.00	4.40	ab	10.60	3.60	a	0.20	0.02	a
DSR_Herb_lowN	3.10	0.20	abcde	73.02	4.35	a	10.91	3.58	a	0.34	0.02	b
TPR_Hand_highN	3.79	0.13	c	47.99	2.80	cdef	9.14	2.40	a	0.25	0.01	a
TPR_Hand_lowN	3.02	0.13	ade	43.77	2.72	de	12.92	2.34	a	0.33	0.01	b
TPR_Herb_highN	3.64	0.10	bc	52.51	2.18	cdf	9.57	1.96	a	0.25	0.01	a
TPR_Herb_lowN	3.36	0.12	abce	54.14	2.51	bcf	11.95	2.19	a	0.35	0.01	b
Irrigated, Tv												
TPR_Hand_lowN	2.04	0.21	fg	31.76	4.67	eg	11.18	3.87	a	0.30	0.02	ab
Rainfed, Mv												
TPR_Hand_highN	3.38	0.24	abcde	47.65	5.22	bcdef	13.74	4.24	a	0.23	0.03	a
TPR_Hand_lowN	2.75	0.14	dfg	35.88	3.06	g	9.53	2.62	a	0.33	0.02	b
TPR_Herb_highN	3.71	0.19	abc	53.49	4.19	bcdf	13.60	3.43	a	0.24	0.02	a
TPR_Herb_lowN	2.94	0.16	de	43.72	3.41	cdef	8.22	2.88	a	0.33	0.02	b
Rainfed, Tv												
TPR_Hand_lowN	1.90	0.24	g	23.22	5.32	g	4.49	4.35	a	0.30	0.03	ab
TPR_Herb_lowN	2.33	0.35	defg	34.15	7.67	defg	3.32	6.13	a	0.34	0.04	ab

Table 4.2: Comparison of farm level indicators across different management packages used by rice farmers in Central Luzon, Philippines, 1966 - 2012. Details per management package are provided in Supplementary Material (Tables B3 and B4). Codes: 'Mean' = predicted means, 'SE' = standard error of the means, 'Signif.' = letter-based representation of pairwise comparisons per season at 5% significance level (no overlap in letters between pairs means a statistical difference), 'Mv' = modern variety, 'Tv' = traditional variety, 'DSR' = direct-seeding, 'TPR' = transplanting, 'Hand' = hand-weeding, 'Herb' = herbicide use, 'highN' = high N applied, 'lowN' = low N applied. *Data source*: Central Luzon Loop Survey. (*continued*)

Dry season	**Rice yield (Ya) (t ha^{-1})**			**Labour productivity (kg ld^{-1})**			**Gross margin (kPhP ha^{-1})**			**N use efficiency (kg N kg^{-1} N)**		
	Mean	**SE**	**Signif.**	**Mean**	**SE**	**Signif.**	**Mean**	**SE**	**Signif.**	**Mean**	**SE**	**Signif.**
Irrigated, Mv												
DSR_Hand_highN	3.96	0.30	abc	78.26	9.00	ab	7.56	5.27	a	0.23	0.02	ab
DSR_Hand_lowN	3.88	0.33	abc	96.31	9.65	ab	11.00	5.67	a	0.32	0.03	acd
DSR_Herb_highN	3.95	0.20	ab	92.35	5.30	ab	13.22	3.41	a	0.23	0.02	b
DSR_Herb_lowN	3.39	0.20	c	88.16	5.17	ab	10.01	3.36	a	0.29	0.02	ac
TPR_Hand_highN	4.16	0.21	ab	48.99	5.83	bcd	9.86	3.65	a	0.24	0.02	ab
TPR_Hand_lowN	3.29	0.21	c	43.87	5.63	cd	10.77	3.55	a	0.32	0.02	cd
TPR_Herb_highN	4.35	0.19	a	59.82	5.08	bc	13.59	3.28	a	0.25	0.02	ab
TPR_Herb_lowN	3.73	0.19	bc	59.29	5.02	bcd	10.95	3.25	a	0.34	0.02	d
Irrigated, Tv												
TPR_Hand_lowN	2.08	0.32	d	29.50	8.98	d	8.47	5.50	a	0.27	0.03	abcd

higher labour productivity than transplanting. No significant differences in gross margin were observed for the different management packages due to the large variability of this indicator. The driver of NUE was the N application rate associated with each management package: low N application (i.e. N application rate lower than the overall average N application rate) resulted in significantly higher NUE (> 0.3 kg N kg^{-1} N) than high N application in both seasons.

Overall, modern varieties and high N rates are required to achieve high Ya and these had no negative effects on labour productivity but significantly lowered NUE, and possibly had higher risk in the WS. High values of labour productivity were associated with the use of direct-seeding and herbicide use instead of transplanting and hand-weeding, respectively. Direct-seeding had a slight negative effect on Ya as compared to transplanting while there were no differences between hand-weeding and herbicide use. Further, there were no significant differences between management packages in gross margin per ha. These results suggest that intensification is best achieved with management packages using modern varieties and high N rates, independently of the weed control method. The choice of crop establishment method is key to balance production and drudgery objectives.

4.4.3 Step 2: Trade-offs between indicators

The yield gap between Yp and the yield of production maximising farms was on average 1.1 and 2.3 t ha^{-1} (or 17% and 24% of Yp in relative terms) in the WS and DS (Figure 4.4A and 4.4B), respectively. The yield gap between Yp and the yield that maximised gross margin, labour productivity and NUE was nearly double that reported for production maximising farms: on average 2.2 t ha^{-1} (34%) in the WS and 3.7 t ha^{-1} (40%) in the DS (Figure 4.4A and 4.4B). Production maximising farms achieved labour productivity that is significantly lower than that achieved by labour productivity maximizing farms in both seasons (Figure 4.4C and 4.4D) and obtained significantly lower gross margin than gross margin maximizing farms but only in the WS (Figure 4.4E and 4.4F). Moreover, production maximising farms achieved NUE that is significantly lower than that achieved by NUE maximizing farms but significantly higher than farms maximising labour productivity (Figure 4.4G and 4.4H).

A similar analysis was conducted using the MISTIG survey (2013 DS - 2014 WS) to cross-validate the historical results of Figure 4.4. The maximum labour productivity observed in the WS was 185 kg ld^{-1} and in the DS 243 kg ld^{-1}, which were obtained with yields of 4.9 and 6.0 t ha^{-1} (Figures 4.5A and 4.5B), respectively. The yields that maximised labour productivity are similar to the average Ya (4.5 and 5.5 t ha^{-1}

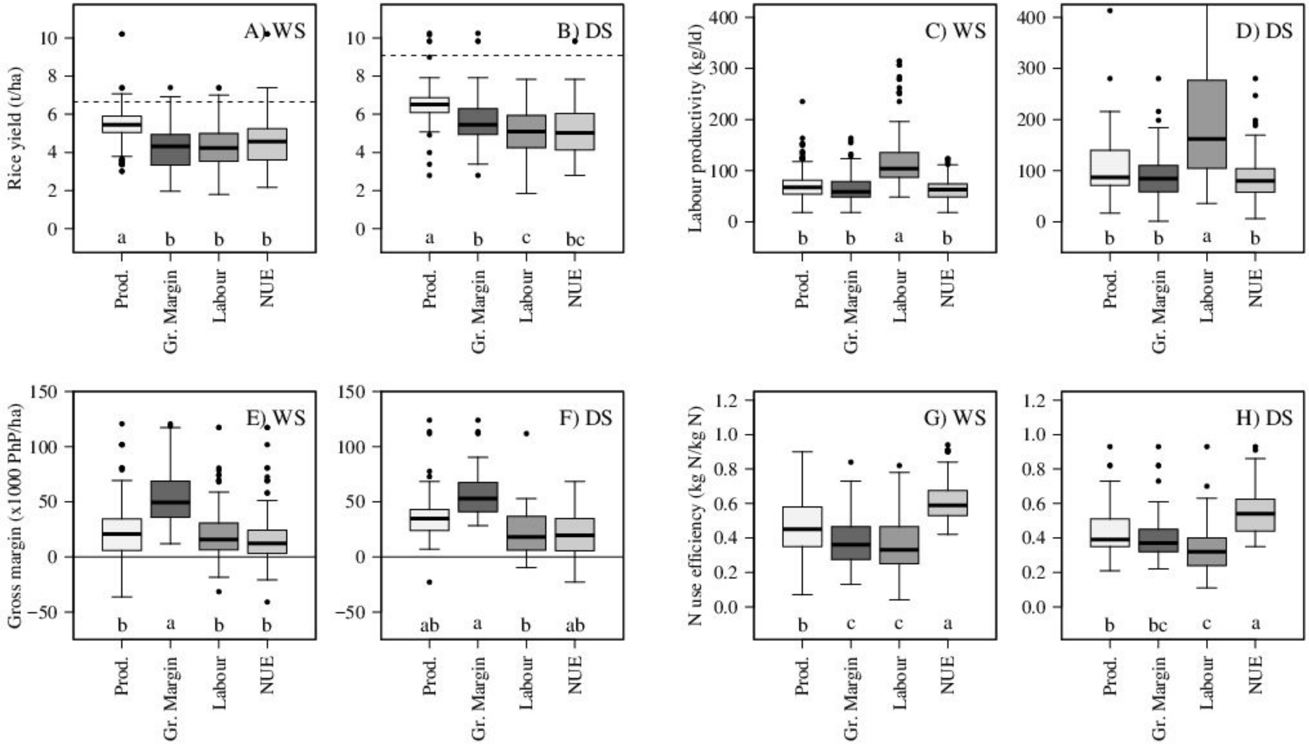

Figure 4.4: Variability in farm level indicators for different farmers' objectives (maximum production, maximum gross margin, maximum labour productivity and maximum NUE) in Central Luzon, Philippines. WS data are presented in A), C), E), G) and DS data in B), D), F), H). Mean differences at 5% significance level between farmers' objectives are indicated with different letters at the bottom of each boxplot. Data for the WS were pooled across 12 observation years and data for the DS were pooled across 11 observation years. Horizontal dashed lines in A) and B) indicate the climatic potential yield (Yp). *Data source*: Central Luzon Loop Survey.

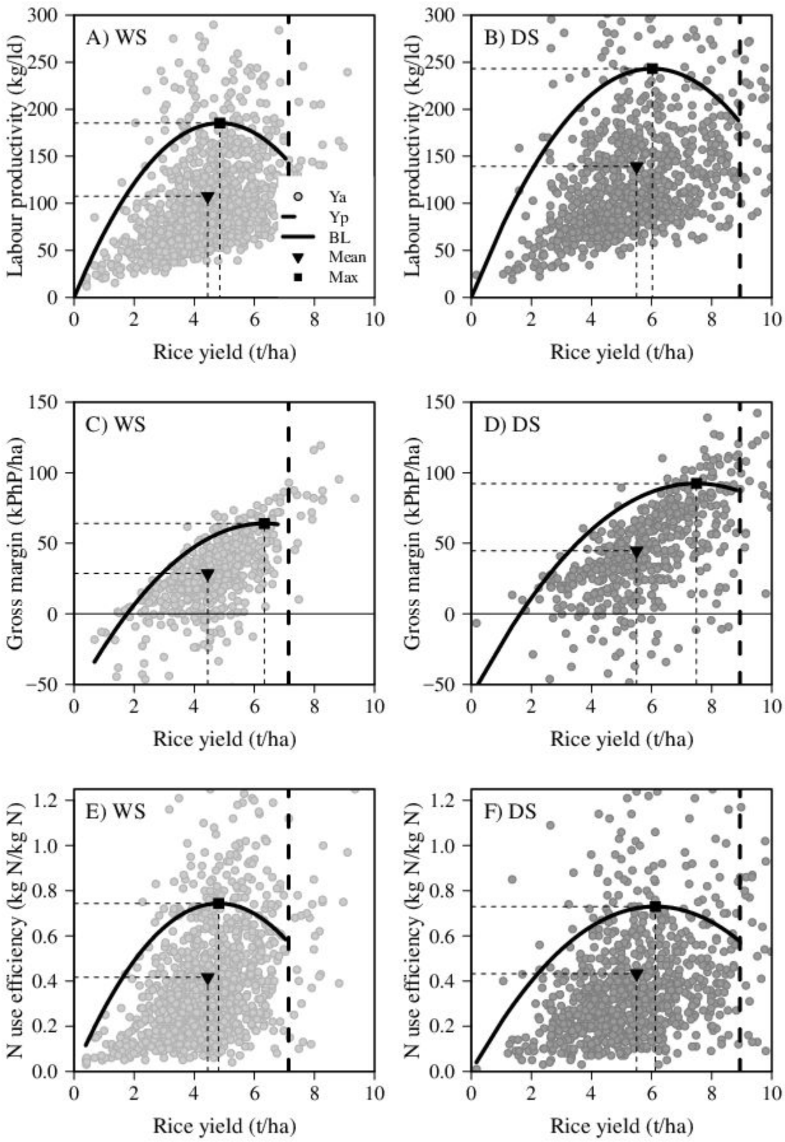

Figure 4.5: Relationship between rice yields and A - B) labour productivity, C - D) gross margin; E - F) N use efficiency for rice-based farming systems in Central Luzon, Philippines (2013 DS - 2014 WS). WS data include both irrigated and rainfed observations while DS data include irrigated observations only. Background data are provided in Figure B6 (Supplementary Material). Legend: 'Ya' = actual farmers' yields, 'Yp' = climatic potential yield, 'BL' = fitted boundary-line, 'Mean' = average across fields, 'Max' = maximum of predicted by the boundary line. *Data source*: MISTIG Survey.

in the WS and DS) but much lower than Yp (7.1 and 8.9 t ha^{-1} in the WS and DS). A similar Ya maximised NUE (Figures 4.5E and 4.5F), which was ca. 0.75 kg N kg^{-1} N. Maximum gross margin from rice farming was ca. 65 and 90 kPhP ha^{-1} in the WS and DS (Figures 4.5C and 4.5D), respectively, which was achieved with yields of 6.3 and 7.5 t ha^{-1}. The difference between Yp and the yield maximising gross margin was 0.8 (11%) and 1.4 t ha^{-1} (16%) in the WS and DS, respectively. These results indicate that yields that maximise labour productivity also maximise NUE but perform sub-optimally in gross margin (i.e. yield gap closure has a positive effect on gross margin). Hence, while the historical analysis (Figure 4.4) suggests lower yields for gross margin maximising farms, the analysis for the most recent year (Figure 4.5), which is not distorted by price changes, suggests that farmers with high yields can also achieve high gross margins.

4.4.4 Step 3: Farm and regional drivers of performance indicators

The most important factors associated with Ya (i.e. elasticity $\geq$ 15%) in the WS were farm size ($-$), number of parcels (+), proportion of hired labour (+), farm power (+), off-farm income ($-$) and province (Table 4.3). A 1% increase in farm size was associated with 0.25% decrease in Ya. A slightly lower positive relationship was observed for the proportion of hired labour (0.17%) and for the number of parcels (0.10 %). Moreover, the use of machinery was positively associated with Ya, there were negative associations between Ya and off-farm income and between Ya and education of the household head and higher Ya was observed on rented than owned land. During the DS, Ya was positively associated with the proportion of hired labour (+) and farms using machinery for land preparation had lower Ya than farms using animal draught power. Ya was consistently higher in Nueva Ecija (the province at the heart of the irrigation system) independently of the season.

There was a consistent association between labour productivity and proportion of hired labour (+) and paddy price (+) in both seasons (Table 4.3). The latter had a relatively large effect: 1% increase in paddy price was associated with ca. 0.70% increase in labour productivity. During the WS, there was a negative relationship between off-farm income and labour productivity, and a positive relationship between this indicator and labour-saving technologies (i.e. animal draught power and machinery). During the DS, labour productivity was further associated with farm size (+) and other income ($-$). In both seasons, lower labour productivity was observed in Pampanga and Pangasinan than in Nueva Ecija.

Table 4.3: Farm and regional drivers of farm performance indicators in the WS and DS. Significant coefficients are reported in bold. Variables were ln-transformed prior to the analysis so that coefficients can be interpreted as elasticities. The reference level of- the categorical variables is as follows: Province = 'Nueva Ecija'; Year_WS = '1970'; Year_DS = '1971'; Power = 'Animal'; Other income = 'No'; Off-farm income = 'No'; n.a. = not applicable. Significance codes: '***' 0.1% '**' 1% '*' 5% '#' 10%. *Data source*: Central Luzon Loop Survey.

	Rice yield (Ya) (t ha^{-1})		**Labour productivity (kg ld^{-1})**		**Gross margin (PhP ha^{-1})**		**N use efficiency (kg N kg $^{-1}$ N)**	
	WS	**DS**	**WS**	**DS**	**WS**	**DS**	**WS**	**DS**
Farm size (ha)	**-0.25** ***	-0.04	0.01	**0.20** ***	**-0.37** ***	**-0.44** *	**-0.17** ***	0.06
Parcels (#)	**0.10** *	0.09	-0.06	0.01	**1.24** ***	**1.96** ***	**0.11** *	0.03
Rented (%)	**0.01** ***	0.01	0.01	0.00	**-0.02** *	**-0.04** *	**0.02** ***	0.01
Shared (%)	0.00	-0.02 #	0.00	-0.01	**-0.05** ***	-0.06	0.01 #	-0.01
Hired labour (%)	**0.17** ***	**0.22** *	**0.11** *	**0.24** *	-0.01	0.42	0.09 #	0.16
Power_Animal+Tractor	**0.15** **	-0.29 #	**0.20** ***	-0.30 #	-0.02	0.29	0.02	**-0.64** ***
Power_Tractor	**0.15** **	**-0.38** *	**0.28** ***	-0.26	0.15	0.24	0.03	**-0.69** ***
Off-farm income_Yes	**-0.16** ***	0.01	**-0.12** *	0.06	-0.18	-0.23	**-0.13** **	0.05
Other income_Yes	-0.04	-0.02	-0.04	**-0.17** **	-0.04	-0.23	-0.03	0.03
Household size (#)	0.09 #	0.04	0.07	0.06	0.12	0.44	0.01	-0.00
Gender (% male)	0.07 #	0.03	0.03	-0.05	0.20	-0.05	0.01	-0.03
Education (yr)	**-0.05** ***	0.02	**-0.05** **	-0.00	-0.07	-0.13 #	**-0.08** ***	0.02
Members farming (#)	-0.02	-0.00	-0.01	0.00	-0.06	-0.06	0.01	-0.00
Members labourer (#)	-0.03 #	0.03	**-0.04** **	**0.06** *	-0.09	-0.02	-0.02	0.02
Members other (#)	**0.02** **	-0.01	0.01	-0.01	**0.05** *	**0.14** *	**0.02** *	0.01
Rice sold (%)	**0.02** ***	0.01	0.01	0.01	**0.03** **	0.03	**0.02** ***	0.00
Paddy price (PhP kg^{-1})	n.a.	n.a.	**0.69** ***	**0.72** **	n.a.	n.a.	**0.36** **	**0.54** *

Table 4.3: Farm and regional drivers of farm performance indicators in the WS and DS. Significant coefficients are reported in bold. Variables were ln-transformed prior to the analysis so that coefficients can be interpreted as elasticities. The reference level of- the categorical variables is as follows: Province = 'Nueva Ecija'; Year_WS = '1970'; Year_DS = '1971'; Power = 'Animal'; Other income = 'No'; Off-farm income = 'No'; n.a. = not applicable. Significance codes: '***' 0.1% '**' 1% '*' 5% '#' 10%. *Data source*: Central Luzon Loop Survey. *(continued)*

	Rice yield (Ya) (t ha^{-1})		**Labour productivity (kg ld^{-1})**		**Gross margin (PhP ha^{-1})**		**N use efficiency (kg N kg $^{-1}$ N)**	
	WS	**DS**	**WS**	**DS**	**WS**	**DS**	**WS**	**DS**
Province_Bulacan	-0.06	**-0.21** **	-0.04	-0.10	-0.30 #	-0.32	0.06	-0.05
Province_Pampanga	**-0.86** ***	**-0.33** **	**-0.72** ***	**-0.32** *	**-1.10** ***	**-1.15** **	**-0.53** ***	0.15
Province_Pangasinan	**-0.22** ***	**-0.25** **	**-0.12** *	**-0.44** ***	-0.05	0.14	**-0.23** ***	-0.16 #
Province_Tarlac	**-0.26** ***	0.18	-0.06	-0.24	0.07	0.48	**-0.35** ***	0.10
Year_1979	**0.25** *		**0.36** **		-0.36		0.16	
Year_1982	**0.26** *		**0.65** ***		**-0.65** *		0.27 #	
Year_1986	0.13		**0.51** ***		**-0.73** *		0.02	
Year_1990	0.11		0.28 #		-0.39		-0.07	
Year_1994	0.16		**0.57** ***		-0.41		-0.04	
Year_2008	0.22 #		**0.88** ***		**-1.56** ***		0.08	
Year_2011	0.03		**0.56** ***		**-1.19** **		-0.10	
Year_1987		0.06		**0.67** **		-1.18		0.12
Year_1991		0.15		**0.89** ***		**-2.05** **		0.19
Year_2004		0.18		**1.21** ***		**-1.55** *		0.23
Year_2007		0.35 #		**1.21** ***		-0.54		0.22
Year_2012		0.30		**1.28** ***		-1.28 #		0.33
Intercept	0.04	0.52	0.83	1.28	1.45	0.72	**-2.25** ***	**-2.69** **
Adjusted-R^2	0.32	0.18	0.27	0.50	0.31	0.27	0.21	0.13

Gross margin (expressed on a per ha basis) was positively associated with the number of parcels and a 1% increase in parcel number resulted in 1.2 and 1.9% increase in this indicator in the WS and DS (Table 4.3), respectively. The effects of farm size on gross margin were also significant in both seasons but of smaller magnitude and negative (ca. 0.4%). However, these effects should be interpreted with caution because farm size was positively correlated with the number of parcels (Pearson correlation coefficient equal to 0.32 and 0.58 in the WS and DS, respectively, but the impact of this was minor in all other models). There was also an association between gross margin and land tenure (−) and the number of household members with 'other' primary occupation (+) in both seasons, and with share of rice sold (−) in the WS. Lower gross margin was observed in Pampanga than in Nueva Ecija independently of the season. The negative coefficient for the most recent WS years was larger than for older years which partially indicates a decline in economic performance of rice farming during this season.

Finally, there were unclear associations between NUE and farm and regional factors and these models had the lowest R^2 in our analysis: 0.21 and 0.13 in the WS and DS (Table 4.3), respectively. As an example, there was a positive association between paddy price and NUE (0.36 - 0.54%) in both seasons which is difficult to explain. Moreover, there was a relationship between NUE and farm size (−), number of parcels (+), off-farm income (−) and, surprisingly, education (−) in the WS, and between NUE and the use of animals and machinery (−) in the DS. NUE was significantly greater in Nueva Ecija as compared to other provinces (e.g. Pampanga, Pangasinan and Tarlac) in the WS while there were no significant differences in the DS.

4.4.5 Step 4: Farm and regional drivers of management practices

In an earlier paper, we found that the management practices explaining the efficiency yield gap were crop establishment method and date, and timing of the first fertiliser application (Chapter 3). Here, we found an association between crop establishment method, farm size and hired labour (Figures 4.6A - 4.6D). For the WS periods between 1986 and 1999, farms using direct seeding had larger farm sizes than farms using transplanting. This is an expected result because direct-seeding requires less labour and it is then more suitable for larger farms. In addition, there were significant differences between the two practices in terms of hired labour during most time periods indicating that transplanting relies more on hired labour.

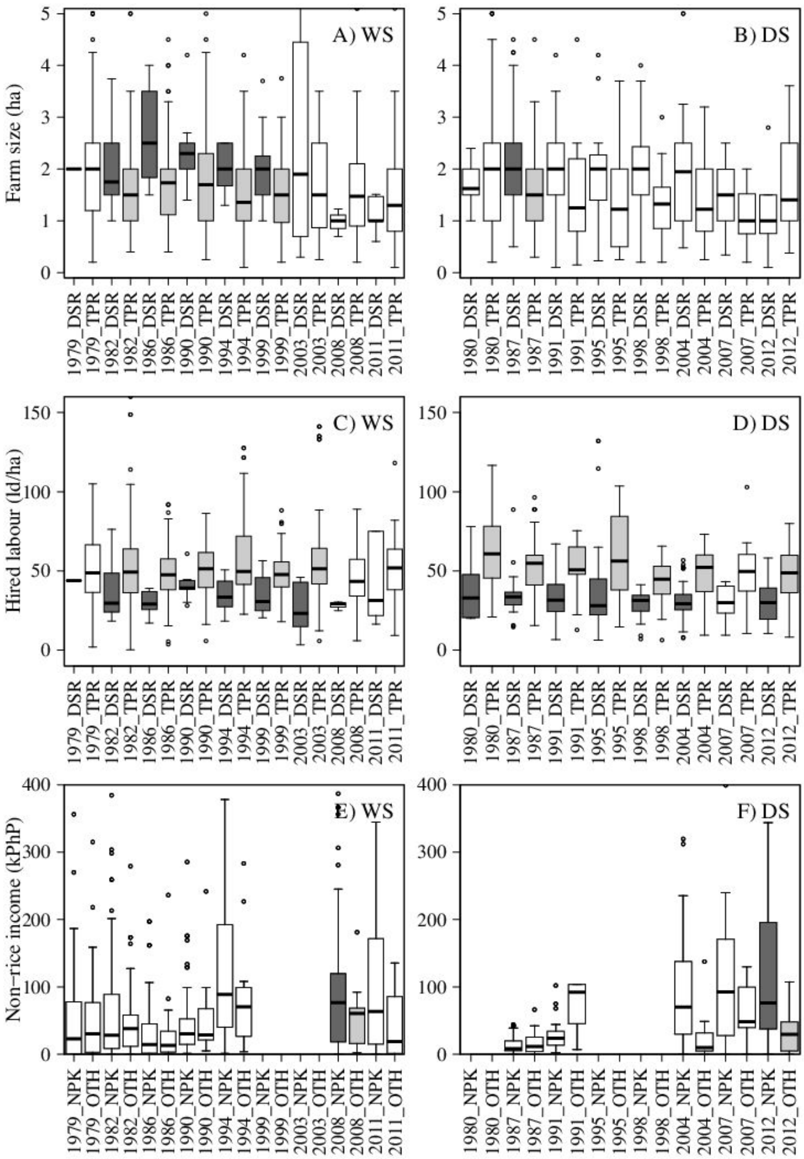

Figure 4.6: Differences in farm size (A - B), hired labour (C - D) and non-rice income (E - F) between management practices that explain rice yield gaps in Central Luzon, Philippines. Note that income data should not be compared between different years due to differences in data collection approaches. Light and dark grey colours indicate significant differences between contrasting management practices (i.e. direct-seeding 'DSR' vs. transplanting 'TPR' and use of N, P and K 'NPK' vs. use of N, P or K 'OTH') in a particular year; white colour means no significant difference. Background data are provided in Tables B5 and B6 (Supplementary Material). *Data source*: Central Luzon Loop Survey.

Table 4.4: Farm and regional drivers of crop management practices in the WS and DS. Significant coefficients are reported in bold. Independent and dependent variables were ln-transformed prior to analysis so that coefficients can be interpreted as elasticities. The reference level of the categorical variables is as follows: Province = 'Nueva Ecija'; Year_WS = '1970'; Year_DS = '1971'; Power = 'Animal'; Other income = 'No'; Off-farm income = 'No'; n.a. = not applicable. Significance codes: '***' 0.1% '**' 1% '*' 5% '#' 10%. *Data source*: Central Luzon Loop Survey.

	Timing 1st fertilisation (days after sowing, DAS)		**Crop establishment date (day of the year, DOY)**		**Nitrogen applied (kg N ha^{-1})**	
	WS	**DS**	**WS**	**DS**	**WS**	**DS**
Farm size (ha)	**-0.09** *	-0.02	-0.01	0.02	**-0.23** ***	**-0.12** **
Parcels (#)	-0.04	**0.26** **	**-0.03** *	-0.05 #	0.04	-0.06
Rented (%)	**0.01** **	**-0.01** *	0.00	0.00	0.00	0.00
Shared (%)	**0.02** ***	0.00	**0.01** ***	0.00 #	0.00	-0.01
Hired labour (%)	-0.10 #	0.03	0.01	-0.06 #	**0.16** **	-0.03
Power_Animal+Tractor	**-0.13** *	-0.35	0.00	n.a.	**0.28** ***	-0.03
Power_Tractor	-0.01	-0.40	-0.02	0.01	**0.33** ***	-0.08
Off-farm income_Yes	0.04	-0.02	-0.02 #	0.04 #	**-0.14** **	-0.05
Other income_Yes	0.01	0.00	-0.02	0.02	0.06	-0.07
Household size (#)	**0.18** ***	-0.05	-0.01	0.00	0.10	-0.01
Gender (% male)	-0.04	**-0.17** *	0.02 #	0.00	**0.13** **	0.04
Education (yr)	0.01	0.01	-0.01	-0.01	**0.04** *	-0.01
Members farming (#)	0.02	-0.02	0.00	**-0.02** *	-0.03	0.05
Members labourer (#)	-0.02	0.02	0.00	-0.01	-0.01	0.01
Members other (#)	-0.01	-0.01	**0.01** *	0.00	-0.01	-0.02
Rice sold (%)	0.01	0.01	0.00	0.00	0.00	0.00
Paddy price (PhP kg^{-1})	-0.23 #	0.19	**0.08** *	**0.23** ***	0.23	0.22

Table 4.4: Farm and regional drivers of crop management practices in the WS and DS. Significant coefficients are reported in bold. Independent and dependent variables were ln-transformed prior to analysis so that coefficients can be interpreted as elasticities. The reference level of the categorical variables is as follows: Province = 'Nueva Ecija'; Year_WS = '1970'; Year_DS = '1971'; Power = 'Animal'; Other income = 'No'; Off-farm income = 'No'; n.a. = not applicable. Significance codes: '***' 0.1% '**' 1% '*' 5% '#' 10%. *Data source*: Central Luzon Loop Survey. *(continued)*

	Timing 1st fertilisation (days after sowing, DAS)		**Crop establishment date (day of the year, DOY)**		**Nitrogen applied (kg N ha^{-1})**	
	WS	**DS**	**WS**	**DS**	**WS**	**DS**
Province_Bulacan	-0.01	**0.26** ***	**-0.07** ***	-0.03	**-0.20** ***	**-0.19** **
Province_Pampanga	-0.05	**0.44** ***	**0.09** ***	0.02	**-0.62** ***	**-0.73** ***
Province_Pangasinan	-0.08	-0.11	**-0.06** ***	**-0.09** ***	0.07	-0.04
Province_Tarlac	**-0.17** *	-0.09	**-0.08** ***	-0.05	0.07	0.08
Year_1979	n.a.		n.a.		**0.76** ***	
Year_1982	**0.40** ***		**0.07** ***		**0.77** ***	
Year_1986	0.12 #		0.02		**0.93** ***	
Year_1990	**0.60** ***		**-0.05** *		**0.85** ***	
Year_1994	**0.38** ***		0.02		**1.09** ***	
Year_2008	**0.25** **		0.01		**0.88** ***	
Year_2011	0.16		0.00		**0.94** ***	
Year_1987		n.a.		n.a.		**0.51** *
Year_1991		-0.02		n.a.		**0.60** **
Year_2004		-0.18		-0.02		**0.63** **
Year_2007		-0.09		**-0.11** ***		**0.46** *
Year_2012		-0.19		**-0.09** ***		**0.59** **
Intercept	**3.48** ***	**3.36** ***	**5.19** ***	**5.62** ***	**1.30** *	**3.83** ***
Adjusted-R^2	0.22	0.18	0.26	0.29	0.31	0.24

The factors associated with the timing of the first fertiliser application and crop establishment date are reported in Table 4.4, and their variability and relationship with performance indicators are provided in Figures B4 and B5 (Supplementary Material), respectively. The technical recommendation for the timing of fertiliser application proposes to apply ca. 30% at an early stage ($<$ 20 DAS) and the rest at critical growth stages such as panicle initiation (R. Buresh, personal communication). In the CLLS, most farmers used only one N application, which for ca. 60% of the cases occurred before 20 DAS (Figure B5; Supplementary Material). In the WS, fertiliser timing was done at later dates by larger households (household size can be used as proxy for family labour availability). Moreover, fertiliser timing was associated with farm size ($-$), farm power ($-$) and, to a less extent, land tenure (later application in rented and shared fields than in own fields). During the DS, fertiliser timing was positively associated with the number of parcels and negatively associated with the gender balance in the household meaning that fertiliser was applied later in more fragmented farms and earlier in households with greater proportion of males. In addition, the first fertiliser application was done earlier in Nueva Ecija than in Bulacan and Pampanga but only during the DS. There were no clear relationships between farm and regional factors and crop establishment date in the WS while there was a positive association with paddy prices (0.23%) in the DS.

Following our earlier paper (Chapter 3), the management practices explaining the resource yield gap were fertiliser (N and K) application as well as weed control. In the most recent periods, non-rice income tended to be greater on farms using N, P and K than on farms using either N, P or K but results were only significant for 2008 WS and 2012 DS (Figures 4.6E and 4.6F). The driver of N use in both seasons was farm size ($-$, Table 4.4) suggesting that smaller N rates were applied on larger farms. During the WS, there were also relationships between N use and proportion of hired labour ($+$), farm power ($+$), off-farm income ($-$), gender balance ($+$) and, to a lower extent, education ($+$). Finally, significantly lower N amounts were applied in Bulacan and Pampanga than in Nueva Ecija in both seasons. For weed control, there were no significant differences between using herbicide and hand-weeding for most variables analysed (Tables B5 and B6; Supplementary Material).

4.5 Discussion

We conducted an integrated assessment of rice yield gaps and associated management packages used by rice farmers in Central Luzon at field, farm and regional levels. We targeted to identify constraints or stimuli for intensification of rice farming based on two hypothesis: 1) there is a trade-off between closing yield gaps and optimising performance of other indicators (labour productivity, gross margin and NUE) and 2) intensification of management practices used by farmers (determining farm performance) can be explained by farm and regional conditions. Below, we discuss our main findings in relation to these hypothesis and to the effects of farm structural change and (hired) labour at different levels.

4.5.1 Yield gap closure and farm performance

An exploitable yield potential of 80% Yp has been proposed and used as a benchmark for yield gap analysis (van Ittersum et al., 2013; Cassman et al., 2003). The rationale behind this threshold is that it is not economically viable, nor environmentally/socially desirable, for farmers to produce above this threshold. Our analysis show that for the most recent period (year 2014), the relative yield gap closure of farms maximising labour productivity or maximising NUE was ca. 68% and that of farms maximising gross margin was ca. 86% (Figure 4.5). These results partially confirm our first hypothesis: an exploitable yield potential of 80% Yp may be a rather high benchmark to avoid trade-off between production and social, or environmental, performance but adequate to avoid trade-off between production and economic performance.

Labour productivity was the indicator whose performance increased most over time (Figure 4.3F) especially due to the adoption of direct-seeding as a crop establishment method instead of transplanting (Table 4.2 and Figure 4.3B). The key factors associated with labour productivity were hired labour (+) and paddy price (+, Table 4.3), which is similar to the findings of Erguiza et al. (1990). No clear trends were observed over time for NUE (Figure 4.3H), although N application rates increased over time especially in the DS, and the large variability of this indicator (0.1 - 1.2 kg N kg^{-1} N) was mostly explained by the amount of N used in different management packages (Table 4.2). In general, lower amounts of N applied contributed to higher NUE and NUE was higher in Nueva Ecija than in other provinces during the WS (Table 4.3), possibly due to the proximity to PhilRice which is promoting site-specific nutrient management.

A synergy between high Ya and high gross margin was observed in the data from the most recent time period (Figure 4.5) but not when data from a period of 50 years were pooled (Figure 4.4). The former may be explained by the high paddy prices obtained by Philippine farmers relative to other regions in Southeast Asia (Moya et al., 2004) while the latter may be distorted by different paddy prices pooled across different years (cf. Figure 4.3I). High paddy prices in the Philippines are imposed through government intervention to prevent imports of cheap rice from exporting countries such as Vietnam and Thailand. This rice price policy has a negative effect for consumers and a marginal positive effect for producers as gross margins from rice are only ca. 16.2 kPhP ha^{-1} $season^{-1}$ (ca. 300 € ha^{-1} $season^{-1}$). Improved crop management does not seem to contribute to considerable improvement in gross margin either (Table 4.2; Lampayan et al., 2015; Rejesus et al., 2011; Pampolino et al., 2007; Moya et al., 2004).

4.5.2 Constraints and stimuli to intensification

Although there was some association between factors at farm and regional levels and management practices such as crop establishment date and method, and N application rate and timing (Figure 4.7), the results of our analysis were not very conclusive (Tables 4.3 and 4.4). We focus the discussion below on management practices related to crop establishment (method and date) and nutrient management (quantity and timing) as these were found to relate to the efficiency and resource yield gaps (Chapter 3). We were not able to obtain further insights into the technology yield gap, which was ca. 28% - 46% of the total yield gap (Chapter 3), as that requires other types of data, e.g. agronomic trials conducted on-farm or on-station.

Crop establishment

There were clear differences in farm size and hired labour between farms using transplanting or direct-seeding as a crop establishment method (Figure 4.6). Transplanting is a labour-intensive practice which tends to be used in smaller farms and with greater amounts of hired labour compared to direct-seeding (see also Pandey and Velasco, 2002; Erguiza et al., 1990). In addition, we found no differences between the two establishment methods in terms of labour use for weed control (data not shown) even though it has been shown empirically that direct-seeding has higher labour requirements for weeding than transplanting (Lantican et al., 1999; de Datta, 1986). This may explain why farms using direct-seeding exhibit greater labour productivity (Table 4.2), but slightly lower Ya (Chapter 3), than farms using transplanting. Although no major yield differences between the two crop establishment methods are generally reported in controlled conditions (e.g. Peng et al., 1995), it has been shown that indeed the yield of direct-seeded rice tends to be lower in farmers' field conditions due to uneven land levelling and inadequate weed control (Pandey and Velasco, 2002; Lantican et al., 1999; de Datta, 1986).

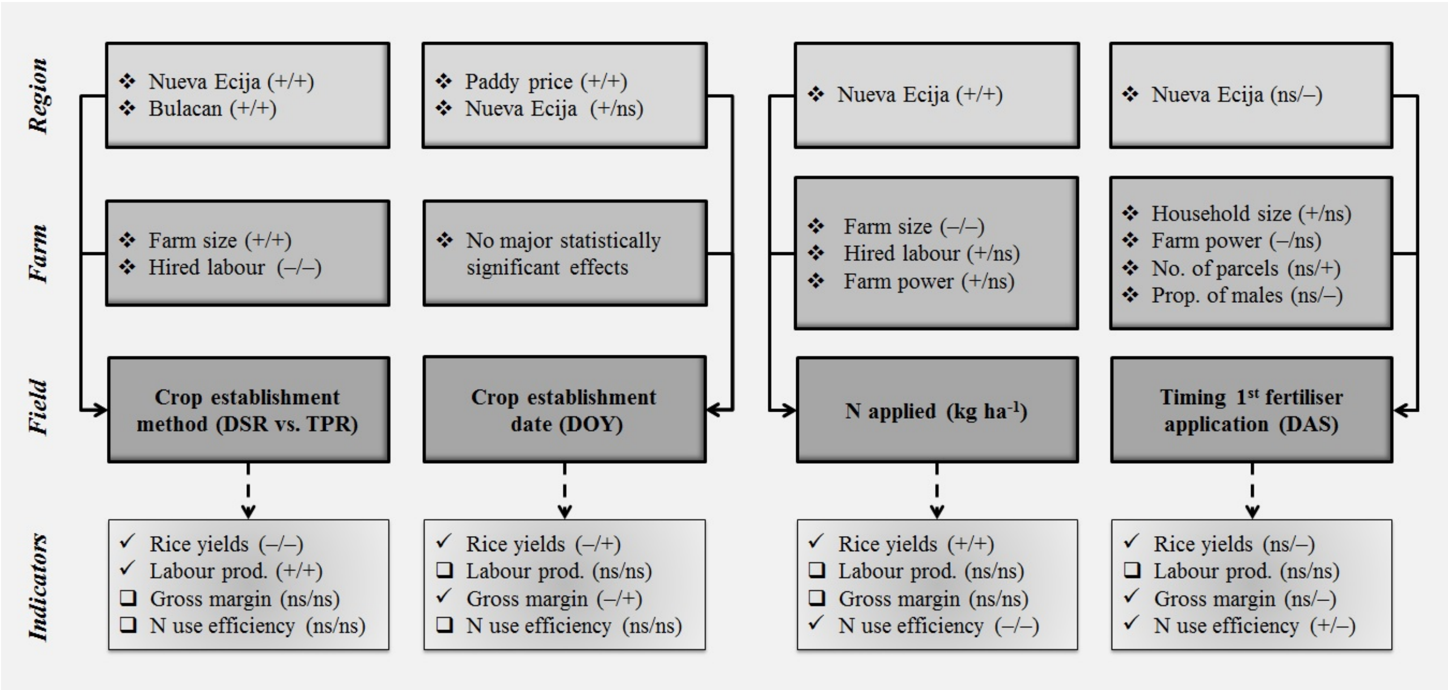

Figure 4.7: Farm and regional drivers of crop establishment method, crop establishment date, N applied and timing of the 1st fertiliser application, and implications for farm performance, in rice-based farming systems in Central Luzon, Philippines. The sign of the effects is reported for both seasons in the following order WS/DS. Regional effects are expressed as differences in management practices between Nueva Ecija (or Bulacan) and other provinces. For crop establishment method, effects of direct-seeding over transplanting are reported. Codes: 'DSR' = direct-seeding, 'TPR' = transplanting, 'DOY' = day of the year, 'DAS' = days after sowing. *Data source*: Central Luzon Loop Survey.

Crop establishment was performed at earlier dates in Nueva Ecija than in other provinces. This may be related to more water stress in provinces not covered by UPRIIS (see Figure 4.2) and by better water accessibility, and proximity to research institutions, in Nueva Ecija which may favour more timely crop establishment (Loevinsohn et al., 1993). Institutional factors at regional level may also affect water accessibility, and establishment dates, as an uneven distribution of water has been recorded within UPRIIS even if that compromises rice production and rice cultivated area across the region (Barker and Levine, 2012). In addition, there was a positive association in the WS between establishing rice at earlier dates on the one hand and rice yields and gross margins on the other hand, while the opposite was observed in the DS, but in either case the evidence was not very strong (Figure B4; Supplementary Material). We would expect that earlier crop establishment would allow farmers to harvest the crop earlier while the price is still high, but from our data this association was only evident in the WS.

Nutrient management

There was a negative association between farm size and N application rate in both seasons while households with off-farm income applied lower amounts of N in the WS (Table 4.4). Lower N application rates may be a strategy to minimise the occurrence of lodging of crop stands due to typhoons but its relationship with off-farm income is unclear. The implications of low N application rates in the short-term are low Ya, but high NUE, while in the long-term this may lead to a decline in soil fertility due to mining of soil nutrient stocks (Pampolino et al., 2008; Dobermann et al., 2000; Cassman et al., 1995). We also found that there were no major significant differences in non-rice income between farms using N, P and K and using N, P or K, even though non-rice income tended to be higher for the former group in the most recent periods (Figures 4.6E and 4.6F). This indicates that fertiliser use is not closely associated with the availability of non-rice income during the growing season and that other factors not studied in this paper such as access to credit or fertiliser prices may be more important.

Increases in N use efficiency are still possible for many farms (Figure 4.5). This can be realised if farmers minimise the amount of N applied before/at sowing, use one or two split applications at panicle initiation and at flag leaf extrusion and apply a total N amount consistent with the expected yield levels and the amount of indigenous soil N supply (Cassman et al., 1998). Adjusting timing and using split applications are simpler to implement than to predict indigenous soil N supply due to its highly variable nature in flooded soils (Olk et al., 1998; Cassman et al., 1996). However, Moya et al. (2004) noted that farmers tend to limit the number of applications to once or twice in locations where wage rates, i.e. opportunity costs for labour, are higher (such as in the Philippines) and this is likely to have important implications for fertiliser timing as well. In our analysis, farms with larger households tended to fertilise later than farms with smaller households (but we cannot assess the cause of this) and farmers prioritized owned fields over rented or shared fields (but the evidence is not so strong. Moreover, farmers who performed the first fertiliser application at later dates in DS obtained lower yields, gross margins and N use efficiency than farmers who applied fertiliser at earlier dates and an optimum for the aforementioned indicators was observed at ca. 10 DAT (Figure B5), which is consistent with the current technical recommendation (R. Buresh, personal communication).

4.5.3 Farm structural change and hired labour

Over the past half-century there has been a gradual transition from subsistence to commercial rice-based farming systems in Central Luzon (Figure 4.3; Erenstein, 2006; Pingali, 1997). For example, there was a structural shift of household income away from land-saving and towards labour-saving technologies due to increased access to urban labour markets (Estudillo and Otsuka, 1999) and an increasing importance of non-rice income for rural livelihoods (Takahashi and Otsuka, 2009). Moreover, land reform programs were unable to consolidate farm size and secure land ownership (Koirala et al., 2016; Otsuka, 1991; Ledesma, 1980). Underlying these changes were social relationships between landowners and landless households, captured by anthropological research on everyday politics in Central Luzon (Kerkvliet, 1990).

The aforementioned dynamics can be seen in the effects of (hired) labour at different levels. At field level, there was a reduction in labour use due to adoption of labour-saving technologies. At farm level, saving of labour increased labour productivity, at the expense of actual yields for some farms (Figures 4.4 and 4.5), and allowed family members to engage in off-farm activities (Figure 4.3L; Estudillo and Otsuka, 1999). Together with favourable labour markets, these contributed to an increasing importance of hired labour over family labour (Figure 4.3J) for e.g. crop establishment operations (Figure 4.6). The proportion of hired labour was also positively associated with N use in the WS (Table 4.4). These changes in crop and farm management had repercussions at regional level because they created further employment opportunities for landless households (Otsuka, 2000; Kerkvliet, 1990). Despite their positive effect for the rural economy, these labour dynamics suggest that rice farming has become a secondary activity for many landowner households in Central Luzon, which in addition to the riskiness of crop failure in the WS, may not be conducive to further intensification and yield gap closure.

4.6 Conclusions

Major changes occurred in rice-based farming systems of Central Luzon during the past half-century. In terms of crop management, there was an increase in the proportion of irrigated fields and a wide adoption of improved varieties. Moreover, direct-seeding substituted transplanting as a crop establishment method, N application rates increased particularly in the DS and hand-weeding was largely replaced by herbicides. These changes point towards gradual transition from subsistence to commercial production systems, which is further supported by an increasing importance of hired labour and rice sold. As a result, there was an increase in rice yields and, especially, in labour productivity while there were no clear trends in gross margin and N use efficiency.

Our results partially confirm the first hypothesis of this study because there was a trade-off between closing yield gaps and maximising N use efficiency or labour productivity but not much between closing yield gaps and maximising gross margin. At the field level, highest actual yields were achieved with high amounts of N and with transplanting as crop establishment method. However, high N application resulted in sub-optimal N use efficiency and transplanting resulted in lower labour productivity compared to direct-seeding. In contrast, there were no significant differences in gross margin across management packages indicating no trade-off between production and economic objectives under the prevailing price ratios in the study region. These conclusions are also true at the farm level as actual yields maximising N use efficiency or labour productivity were ca. 32% lower, and actual yields maximising gross margin were only ca. 15% lower, than climatic potential yields.

Even though there was some association between factors at farm level and management practices, or farm performance indicators, results regarding the second hypothesis were not very conclusive. For nutrient management, higher N application rates were applied on smaller farms while the first fertiliser application was done at later dates by larger households and farms with more parcels in the wet and dry season, respectively. In contrast, the adoption of direct-seeding as a crop establishment method was observed in larger farms and used lower amounts of hired labour than transplanting. The variation in sowing dates was mostly related to different provinces in the WS (a proxy for e.g. water accessibility) but not to other factors at farm or regional levels. Management practices used by farmers in Central Luzon seem to be affected, but not constrained, by farm and regional conditions.

4.7 Acknowledgements

We are grateful to Ms. Piedad Moya (IRRI SSD) and Ms. Mary-Rose San Valentin (IRRI SSD) for making the Central Luzon Loop Survey readily available for analysis and discussing preliminary findings. We thank Prof. Ken Cassman (University of Nebraska-Lincoln) for his insightful comments on the interpretation of rice yield responses to N and variability in N use efficiency and Dr. Joost van Heerwaarden (WUR PPS) for his help in the estimation of the linear mixed model. This research was part of the Wageningen University strategic programme 'Mapping for Sustainable Intensification' and was funded by Wageningen University and the Global Rice Science Partnership (GRiSP).

4.8 Appendix

The steps followed to estimate the boundary lines used in this chapter were:

1. The average Ya ($\overline{\text{Y}}$a) for the observations above the 90$^{\text{th}}$ percentile for each indicator was estimated;

2a. Farms with Ya lower than $\overline{\text{Y}}$a were ordered in ascending order of Ya after excluding observations above the 90$^{\text{th}}$ percentile for each indicator (dependent variable);

2b. Farms with Ya greater than $\overline{\text{Y}}$a were ordered in descending order of Ya after excluding observations above the 90$^{\text{th}}$ percentile for each indicator (dependent variable) and observations with Ya greater or equal than the long-term average of maximum Yp (6.7 and 9.1 ton ha^{-1} for WS and DS, respectively);

2c. The maximum response of each indicator for a given Ya (BLx) was estimated for each field i as follows: if Ya$_i$ < BLx_{i-1} then assign BLx_{i-1}, else assign Ya$_i$;

3. Quadratic functions were fitted with the *lm* function of R software for each indicator using BLx as dependent variable and Ya as independent variable.

CHAPTER 5

Yield gaps in Dutch arable farming systems: Analysis at crop and crop rotation level

This chapter is published as:
Silva, J.V.; Reidsma, P.; van Ittersum, M.K. 2017. Yield gaps in Dutch arable farming systems: Analysis at crop and crop rotation level. *Agricultural Systems*. 158, 78 - 92.

Abstract

Arable farming systems in the Netherlands are characterized by crop rotations in which potato, sugar beet, spring onion, winter wheat and spring barley are the most important crops. The objectives of this study were to decompose crop yield gaps within such rotations into efficiency, resource and technology yield gaps and to explain those yield gaps based on observed cropping frequencies and alternative farmers' objectives. Data from specialized Dutch arable farms between 2008 and 2012 were used. Production frontiers and efficiency yield gaps were estimated using the stochastic frontier framework. The resource yield gap was quantified through the estimation of highest farmers' yields (Y_{HF}, average across farms with actual yields above the 90th percentile). Crop model simulations and variety trials were compiled to assess climatic potential yields (Yp) and technology yield gaps. The contribution of crop area shares and farmers' objectives to actual yields were assessed using regression analysis and based on five different farm level indicators (N production, energy production, gross margin, nitrogen-use efficiency and labour use), respectively.

The average yield gap per crop (as percentage of Yp which is given in parentheses) was: 29.2% (of 72.6 t ha^{-1}) for ware potato, 39.7% (of 71.6 t ha^{-1}) for starch potato, 26.4% (of 107.1 t ha^{-1}) for sugar beet, 32.3% (of 88.3 t ha^{-1}) for spring onion, 25.2% (of 12.3 t ha^{-1}) for winter wheat and 37.5% (of 10.4 t ha^{-1}) for spring barley. The efficiency yield gap ranged between 6.6% (starch potato) and 18.1% of Yp (spring onion). The resource yield gap was lower than 10% of Yp for all the crops and the technology yield gap ranged between 7.1% (ware potato) and 30.7% of Yp (starch potato). There were statistically significant effects of potato (positive quadratic) and onion (positive) area shares on ware potato, sugar beet and winter wheat yields, of sugar beet area share (positive quadratic) on winter wheat yield and of cereal area share (negative) on sugar beet and winter wheat yields. Farmers' objectives explain part of the variability observed in crop yields, i.e. yields were 7 - 24%, 13 - 24% and 12 - 32% lower than Y_{HF}, respectively, for gross margin maximising, labour minimising and N use efficiency maximising farms. In addition, there was a significant positive relationship between gross margin and the yield of ware potato, sugar beet and winter wheat. By contrast, no significant relationships were found between crop yields and NUE or labour use.

We conclude that most of the yield gap was explained by the efficiency yield gap for ware potato and spring onion and by both the efficiency and technology yield gaps for sugar beet and cereals. The resource yield gap explained most of the yield gap of seed potato, and the technology yield gap of starch potato. The results regarding the effects of cropping frequency and crop rotations to crop yields were not very conclusive which suggest that agronomic principles become less evident at 'systems level' given the number of interacting factors at crop rotation level. Finally, although N and energy production were lower for gross margin maximising farms, most crop yields were not significantly different between farms with the highest N and energy production compared to farms performing best on economic (gross margin) objectives.

5.1 Introduction

Crop yield gaps can be estimated and explained at different spatial scales using a wide range of methodologies (Beza et al., 2017; van Ittersum et al., 2013). For instance, yield gap analysis at field (crop) level is usually performed using field trials and/or farm surveys in combination with crop growth simulation models (e.g. Affholder et al., 2012; Subedi and Ma, 2009; Abeledo et al., 2008) and with multivariate statistics (e.g. Delmotte et al., 2011; Fermont et al., 2009; Tittonell et al., 2008a). Such type of analyses provide good insights about the limiting factors to crop growth but they fail to capture the multi-dimensional aspects of crop production occurring at farm and farming systems level.

Understanding the scope for sustainable intensification of current farming systems requires an in-depth, and integrated assessment, of crop yield gaps at the farm level for three main reasons. First, farmers make decisions about which activities to pursue and how to allocate the available resources given their personal objectives and circumstances (Kanellopoulos et al., 2014; Mandryk et al., 2014). Second, there can be incompatibilities or synergies between different activities performed within the same farm (Hochman et al., 2014; Dogliotti et al., 2003; Struik and Bonciarelli, 1997). Third, the farm integrates both biophysical and socio-economic components of agricultural systems. Therefore, farm level analysis using individual farm data are important to expose interactions between different activities as well as the potential limitations and consequences of different management and livelihood strategies (Reidsma et al., 2015b; Kanellopoulos et al., 2012; Tittonell et al., 2009).

Arable farming systems in the Netherlands provide a good case study to test a suite of methodologies aiming at explaining yield gaps at both crop and farm level. Dutch arable farms are organized into crop rotations in which a succession of different crops is repeated every certain number of years. The most important crops are ware, seed and starch potato, sugar beet, spring onion, winter wheat and spring barley. In 2015, approximately 155.000 ha (21% of the total arable area) of potato were harvested in the Netherlands, followed by 130.000 ha of winter wheat, 70.000 ha of sugar beet, 35.000 ha of spring barley and 20.000 ha of spring onion (CBS, 2015). In addition, farms operate close to the climatic potential yield (Yp, www.yieldgap.org) and resource use efficiencies are strongly influenced by economic performance (Mandryk et al., 2014), environmental legislation limiting fertiliser and pesticide use (Boatman et al., 1999) or market regulations (e.g. sugar beet quota).

The objectives of this study are twofold: 1) to disentangle crop yield gaps within Dutch arable farming systems using a standard methodological approach and 2) to explain those yield gaps based on observed cropping frequencies and alternative farmers' objectives. For this purpose, we applied the theoretical framework introduced in Chapter 3 to analyse yield gaps for the most important crops cultivated in arable farming systems in the Netherlands. We hypothesized that yield gaps of the main crops (ware potato, sugar beet and winter wheat) are relatively small (80% of Yp) and that much of this yield gap can be explained by farm and crop rotation factors rather than field and crop level conditions.

5.2 Theoretical framework

A generic arable farm system with a four-year crop rotation composed of potato (*Solanum tuberosum* L.), winter wheat (*Triticum aestivum* L.), sugar beet (*Beta vulgaris* L.) and spring barley (*Hordeum vulgare* L.) is depicted in Figure 5.1. This rotation is a typical example of how rotations looked like in The Netherlands traditionally but they have become more diversified. In addition, there are also distinct regional differences with more (lighter soils) or less (heavy soils) root and tuber crops depending on the soil type. Following Ewert et al. (2011), in this system it is important to differentiate processes and flows occurring at crop rotation level from the ones occurring at crop level as these two levels are nested and have different spatial (i.e. farm area vs crop area) and temporal scales (i.e. length of crop rotation vs crop growing season). The concepts developed to disentangle and explain yield gaps at crop and crop rotation level in this study are described in this section.

In this paper, yield refers to the land productivity of an individual crop and is expressed in ton fresh matter (FM) ha^{-1} whereas production refers to the total production at farm level calculated as the sum of the different crop yields in kg N ha^{-1} or MJ ha^{-1}. Non-substitutable (i.e. water and nutrients) and substitutable inputs (e.g. herbicides and nematicides) for crop growth are referred to as inputs and those can be aggregated at crop or crop rotation level.

5.2.1 Disentangling yield gaps at crop level

Yield gap analysis is useful to understand the relative contribution of growth-defining, -limiting and -reducing factors to actual yields. A framework integrating concepts of production ecology (van Ittersum and Rabbinge, 1997) and methods of frontier analysis (Farrell, 1957) was introduced in Chapter 3 to explain crop yield gaps using individual crop and/or farm data. Crop yield gaps refer to the difference between Yp and actual yields (Ya), and can be further decomposed into efficiency, resource and technology yield gaps (Figure 5.1A).

Five different yield levels are required to decompose yield gaps at crop level. Actual yields (Ya) are the yields currently achieved by farmers and can be compiled through for example farm surveys. Technical efficient yields (Y_{TEx}) refer to the max-

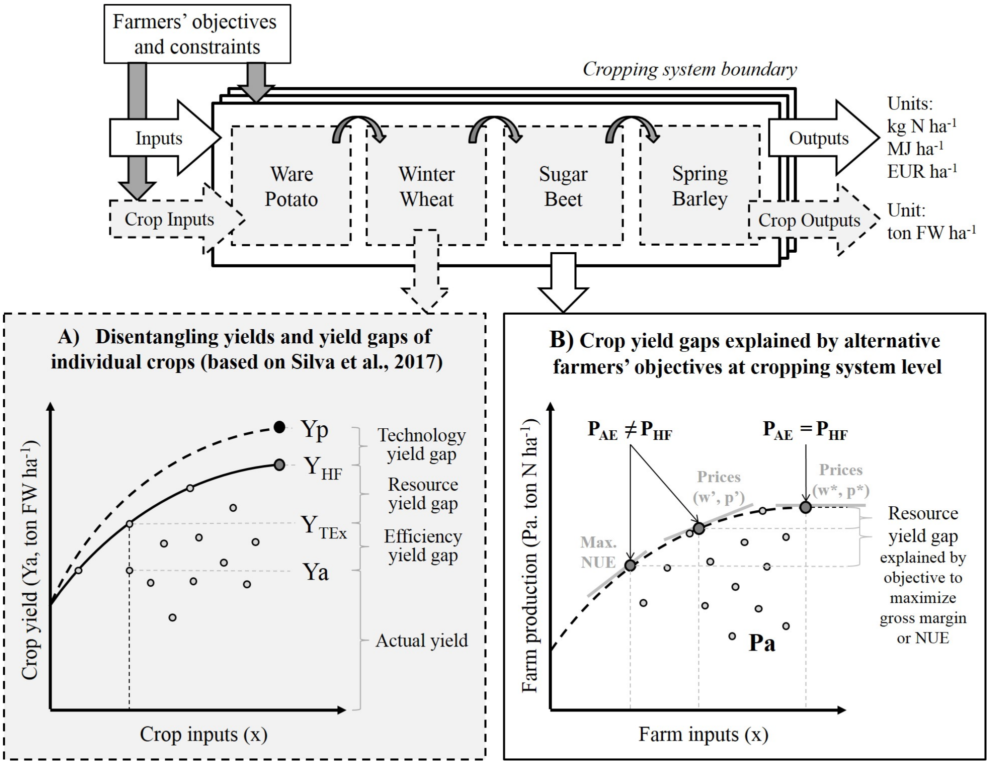

Figure 5.1: Representation of an arable crop rotation and theoretical framework to A) disentangle crop yield gaps and B) explain those considering alternative farmers' objectives. Yp is the climatic potential yield as defined by van Ittersum and Rabbinge (1997). Y_{HF}, Y_{TEx} and Ya are abbreviations for highest farmers' yield, technical efficient yield at a specific input level and actual yield of a specific crop across multiple farms, respectively. P_{HF}, P_{AE} and Pa stand for highest farmers', allocative efficient and actual production at crop rotation level. Solid grey lines in B) represent the maximum nitrogen use efficiency (NUE) and gross margin for two input-output price ratios, (w', p') and (w*, p*). Single input-output relationships are shown for illustration purposes only. See text for further explanation.

imum yield which can be achieved with current input use and can be estimated using methods of frontier analysis (Farrell, 1957). Allocative efficient yields (Y_{AE}) can be defined as the Y_{TEx}, which optimise levels of crop production given farmers' objectives and resource constraints (similar to P_{AE} in Figure 5.1B). Highest farmers' yields (Y_{HF}) provide an indication of the maximum yields currently achieved by farmers and can be estimated as the mean of Ya above the 90th percentile. Finally, the climatic potential yield (Yp) is the maximum theoretical yield which a genotype can achieve in a well-defined biophysical environment (van Ittersum and Rabbinge, 1997).

The efficiency yield gap is defined as the difference between Y_{TEx} and Ya and expresses by how much yield can be increased with current levels of inputs in a particular environment. Yield differences between farms using similar inputs can then be explained by differences in timing, spacing and form of the inputs applied, observed variation in sowing dates as well as rotational effects due to interactions between crops (see below), while controlling for differences in biophysical conditions. The resource yield gap can be estimated as the difference between Y_{HF} and Y_{TEx} and it indicates the additional yield that can be obtained in case input use is increased to the level used to achieve Y_{HF}. Finally, the technology yield gap refers to the difference between Yp (or Yw in rainfed conditions) and Y_{HF} and can be explained by existing limiting factors to production (i.e. von Liebig's law of the minimum) and/or the lack of precision agriculture practices and new varieties able to exploit Yp. Rotational effects may also explain the technology yield gap in case the farms included in the sample share a similar crop rotation plan and hence show little variation in this factor.

5.2.2 Explaining yield gaps at crop rotation level

Understanding crop yield gaps requires looking beyond the field scale and individual season. Below, we frame the importance of rotational effects over time and alternative farmers' objectives when allocating resources to multiple activities within the theoretical framework proposed in Figure 5.1.

Rotational effects on crop yields

A crop rotation can be defined as an ordered succession of crops which are cultivated repetitively every certain number of years (cf. Wijnands et al., 2002). Crop rotations are particularly important to preserve soil fertility and to control pests, diseases and weeds. However, their 'efficacy' depends on a number of factors including the species of crops cultivated, their frequency and sequence, the length of the complete cycle and the number of different crops, among others. Further information about the importance of these factors for the productive, economic and environmental performance of Dutch arable crop rotations can be found in Dogliotti et al. (2003) and Vereijken (1997).

The aforementioned rotational effects need to be taken into account when trying to understand yield variability and yield gaps from a crop rotation perspective. In general, it can be argued that rotational effects explain part of the efficiency, resource and technology yield gap (Figure 5.1A). Different farms sharing similar crop management practices and biophysical environment may still exhibit different efficiency yield gaps in case they cultivate a different number of crop species or cultivate the same crop species in different frequencies and sequences (Mazzilli et al., 2016; Rijk et al., 2013). Crop rotations may also affect the resource yield gap as follows: including legumes as previous crop reduces the need for N application (Reckling et al., 2016; Plaza-Bonilla et al., 2017) and increasing crop diversity helps reducing the pressure of biotic factors and the use of pesticides (Andert et al., 2016; Landis et al., 2008; Struik and Bonciarelli, 1997). Both strategies can explain why different farms require different input levels to achieve similar yield levels. Finally, rotational effects can also explain the technology yield gap in case all farms in the sample use sub-optimal crop rotations. It is important to note that individual effects are generally hard to isolate because they are dynamic and can be confounded with other factors (e.g. renting extra land allows to grow the same specialised crop on this extra land if the rented land is different from year to year).

Farmers' objectives and constraints

Farmers' objectives and resource constraints determine the intensity with which inputs are used and hence, the production level of different activities. We define the difference between maximum production at crop rotation level (P_{HF}, kg N ha^{-1}) and the production observed under different farmers' objectives (e.g. maximum gross margin, P_{AEecon} in kg N ha^{-1}) as 'production trade-off'. This is expected to occur, as generally farmers prioritize economic (or environmental) performance over production maximization (Monjardino et al., 2015; Mandryk et al., 2014). Production trade-offs have socio-economic causes and may arise if different objectives have different optimal production levels.

Production trade-offs can be quantified using individual farm data as the difference in farm production between production and, for instance, gross margin maximizing farmers (Figure 5.1B). Maximum production cannot be based on yield levels, as yields of different crops cannot be directly compared. Instead, N or energy production can be used as proxies. The top 10th percentile farmers of N or energy production and gross margin can then be compared. Other important objectives which can be assessed following the same rationale are minimum labour use ($P_{AElabour}$) or maximum nitrogen use efficiency (NUE) at crop rotation level as an indicator for environmental performance (P_{AEenv}).

5.3 Material and methods

5.3.1 Farm accountancy data

Individual farm data from specialized arable farms in the Netherlands between 2008 - 2012 were used in this study (Figure C1; Supplementary Material). Such data are collected every year by the Wageningen Economic Research with the purpose of monitoring the income and economic performance of agricultural holdings in the Netherlands (van der Veen et al., 2014). The farms monitored are selected from the most recent agricultural census based on a disproportional stratified random sample, i.e. a random sample of farms selected from the agricultural census is assigned to a specific strata defined by the type of farming (e.g. arable farms) and the economic size class (greater than 25000 €).

The Netherlands is a relatively homogeneous country with a temperate maritime climate. The variation in temperature and radiation as well as the distribution of rainfall

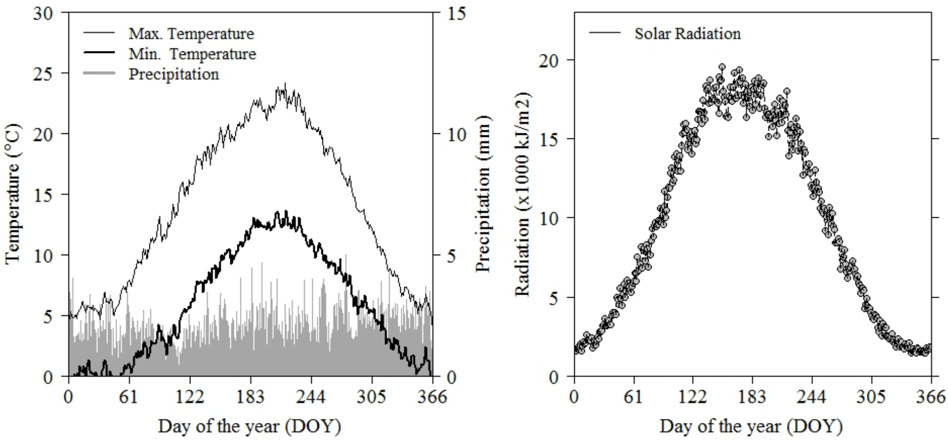

Figure 5.2: Mean daily maximum and minimum temperature (°C), rainfall (mm) and solar radiation (kJ m^{-2}) for the period 1970 - 2012 in De Bilt, The Netherlands. *Source*: Royal Netherlands Meteorological Institute (KNMI).

is well captured by the meteorological station located in De Bilt (Figure 5.2). Annual precipitation is on average 810 mm yr^{-1} (1970 - 2012) and it falls uniformly throughout the year. Minimum temperatures are around 0°C during the winter months (December - February) and have a maximum of about 12.3°C in July/August. Maximum temperatures follow a similar temporal pattern with a minimum value of 5.4°C during the winter and a maximum of about 22.3°C during the summer. Solar radiation increases during the first half of the year to a value of about 17.8 MJ m^{-2} in the summer after which it steadily declines during the second half of the year. In terms of edaphic conditions, three main soil types can be identified in the country: marine clay soils in the north and south-western areas, sandy soils in most of the south- and central-eastern areas and, somehow less representative, peat soils in the north and center-western areas.

Arable crops are mostly cultivated between March and October with the exception of winter wheat, which is cultivated between November and August. Most farms grow a succession of root and tuber crops and cereals over multiple years, in particular sugar beet, potato and winter wheat. Spring onion and spring barley are also important crops but they are cultivated by a smaller number of farms. Generally, potato is the crop with largest gross margin but it is important to differentiate between three crop rotations: 1) starch potato are cultivated once every two years and in relatively marginal soils with the objective of maximising starch production, 2) seed potato are cultivated for high quality medium size tubers which depend on initial seed quality, crop management and haulm killing date, and 3) ware potato are mostly cultivated for the french fries industry with specific quality and size restrictions. Descriptive statistics of the crop level data used are provided in Table C1 (Supplementary Material). The sample size, i.e. total number of farms in the dataset, between 2008 - 2012 was ca. 175 farms per year. We assumed one crop rotation per farm as the data does not allow for a more detailed analysis, but this may not fully represent actual farmers' practice.

5.3.2 Yield gap analysis at crop level

Efficiency yield gap

Stochastic frontier analysis (Kumbhakar and Lovell, 2000; Aigner et al., 1977; Meeusen and van den Broeck, 1977) was employed to estimate crop-specific efficiency yield gaps. The translog functional form was chosen for this analysis because it explicitly considers the first- and second-order terms, including interactions, across inputs and outputs and it does not pose restrictive assumptions about the shape of the production frontier. Its mathematical formulation as well as the estimation of the efficiency yield gap and Y_{TEx} are as follows:

$$\begin{aligned} ln\, y_{it} &= \alpha_0 + \sum_{k}^{K} \beta_k\, ln\, x_{kit} + \frac{1}{2} \sum_{k}^{K} \sum_{j}^{K} \theta_{kj}\, ln\, x_{kit} \times ln\, x_{jit} + \\ &\quad + \sum_{t}^{T} \delta_t\, DT_t + v_{it} - u_{it} \qquad (5.1) \\ \text{EffYg}_{it} &= 1 - exp(-u_{it}) \qquad (5.2) \\ \text{Y}_{\text{TEx}_{it}} &= \text{Ya}_{it} \times exp(-u_{it})^{-1} \qquad (5.3) \end{aligned}$$

where y_{it} is the yield of a specific crop in farm i and period t, x_{it} a vector of input k and j defined according to principles of production ecology (van Ittersum and Rabbinge, 1997) and DT_t a year dummy. The random error term v_{it} is assumed to be N(0, σ_v^2), and independent of the u_{it}, which accounts for technical inefficiency in production and is assumed to be independently distributed following N(μ, σ_u^2). The parameters to be estimated are α_0 (intercept), β_k (first-order terms), θ_{kj} (second-order terms) and δ_t (year effects). Y_{TEx} and the efficiency yield gap are based on u_{it} and defined as in Section 5.1. All parameters and error terms were estimated with maximum likelihood using the R package *frontier* (Coelli and Henningsen, 2013).

Quantitative input-output data for a set of unique combinations of farm $\times$ crop $\times$ year were used to estimate the stochastic frontier models (Equation 1) and the efficiency yield gap (Equation 2). Fresh matter yield (y, kg FM ha^{-1}) of potato (ware, seed and starch), sugar beet, spring onion, winter wheat and spring barley were defined as dependent variables. Independent variables (x) were selected to capture the role of growth-defining, -limiting and -reducing factors to actual yields.

Growth-defining factors were included in the model as a year-dummy in order to control for year-specific climatic conditions referring to e.g. temperature and solar radiation. No regional dummies were included because the Netherlands is considered a homogeneous climatic region (van Wart et al., 2013). We did not control for different varieties, planting/sowing dates and densities and previous crop, as data were not available, so these factors are captured in the efficiency yield gap. The variables considered as growth-limiting factors were N and P use (expressed in kg ha^{-1}) and soil type expressed as a dummy variable with four levels namely, clay (clay content $>$ 80% of the farm area), sand (sand content $>$ 80%), loess (loess content $>$ 80%) and mixed (clay, sand and loess content $<$ 80%). The N use variable corresponds to plant available N applied, i.e. the N applied with mineral fertilisers together with the N fertiliser replacement value of organic amendments (which can be understood as the equivalent effect of organic amendments expressed in mineral fertiliser-N). Fungicide, herbicide, insecticide and nematicide variables (expressed in kg a.i. ha^{-1}) were used to control for growth-reducing factors in the crop production systems to which they are relevant. Finally, a dummy variable was included to differentiate between conventional (n = 169) and organic (n = 16) production systems to control for potential differences in nutrient, pest, disease and weed management. The y and x variables included in the production frontier were mean-scaled and ln-transformed prior to the analysis.

Resource yield gap

The resource yield gap indicates the potential to increase yields due to increases in input use and it can be quantified as the difference between Y_{HF}, i.e. the mean across observations of Ya above the 90^{th} percentile, and Y_{TEx} estimated using stochastic frontier analysis. We did not differentiate between soil types due to non-statistically significance of yield differences between soil types for most crops (see Table 5.1). In order to explain the resource yield gap, we further grouped the observations into lowest yielding farms (Y_{LF}, i.e. mean across observations of Ya below the 10^{th} percentile) and average yielding farms (Y_{AF}, i.e. mean across observations with Ya between the 10^{th} and 90^{th} percentiles) and compared the application of plant available N across the three groups. Furthermore, the statistical significance of the parameters β_k and θ_k in Equation 5.1 provide an indication of the most important input(s) explaining the resource yield gap. Results are presented for 2012 only because no major differences in the resource yield gap were observed during the period analysed (2008 - 2012).

Technology yield gap

Climatic potential yields (Yp) set the biophysical boundaries of yield gap analysis and therefore it is important to assess the difference between Yp and Y_{HF}. We opted to use estimates of Yp instead of water-limited yields (Yw) due to the high uncertainty of the latter estimates related to the depth of water table and frequent occurrence of capillary rise. However, we cannot rule out water stress because irrigation is not a default in the Netherlands. Furthermore, we assumed there is no spatial variability in Yp for the arable crops cultivated in the Netherlands (www.yieldgap.org) since the climate is so homogeneous across the country. Estimates of Yp for ware potato, winter wheat and spring barley were simulated with WOFOST (Boogaard et al., 2013) for an average sowing date and during the period 1992 - 2008 (see Reidsma et al., 2015b). Further details about model calibration and validation can be found in Wolf et al. (2011). In case of starch potato, sugar beet and spring onion, Yp was obtained from variety trials (Rijk et al., 2013). No yield benchmark was used for seed potato due to the peculiarities of the seed potato production system.

Standard moisture contents were used to convert Yp from dry to fresh matter to ensure these estimates are comparable with the Ya values observed in the accountancy database. The standard dry matter contents considered per crop were 21.5% for ware potato and for seed potato, 20.0% for sugar beet and spring onion and 84.0% winter wheat and spring barley. In case of sugar beet and starch potato, we further assumed a standard sugar content of 17% (Suiker Unie, personal communication) and an underwater weight of 524 g (Avebe, personal communication), respectively.

5.3.3 Resource and crop allocation at farm level

Influence of cropping frequency

'Cropping frequency' was the only rotational effect we could analyse using the available data as we had no information about 'cropping sequence'. Cropping frequency was computed as the ratio between the area (in ha) of an individual crop and the total cultivated land (in ha). We had no means to exclude the area of crops cultivated on rented land as that information was not available for each individual crop and the succession of crops grown on rented land was unknown. This is particularly relevant in the case of potato cultivation as it is common practice that farmers rent extra land to grow this crop (Figure C2; Supplementary Material). Therefore, we did the analysis for potato using the subset of farms with less than 33.3% of rented land and less than 40.0% of potato in the total area. We considered these thresholds to restrict the analysis to farms with a relatively low share of rented land (because higher potato area shares of potato point at the importance of potato cultivated on rented land, where rotations are likely wider) and to account for the legal norm of potato cultivation in the Netherlands.

The relationship between crop yields in year t and crop area share in year $t-1$ was studied using regression analysis. We allowed for linear and quadratic effects in order to test whether the increasing share of one crop leads to an increase (or decrease) of crop yields or if a maximum (or minimum) yield level occurs for a particular crop at a specific crop area share. We focused our analysis on the proportion of potato (ware, seed and starch potato), sugar beet, spring onion and cereals (winter wheat and spring barley) and assessed the contribution of each of them to the yield of ware potato, sugar beet and winter wheat. Regressions were fitted to the entire sample of farms available (except for potato, see above) pooling data from four different time periods: 2008 - 2009, 2009 - 2010, 2010 - 2011 and 2011 - 2012. The lm function of R was used for the analysis.

Trade-offs between farmers' objectives

The impacts of farmers' objectives on crop yield gaps were assessed for clay soils and the period 2008 - 2012 (n = 97 farms) in order to average environmental conditions and rotational effects across years. The crops at stake in this analysis were ware and seed potato, sugar beet, spring onion, winter wheat and, spring barley as there were no observations of starch potato in clay soils. Actual farm production, Pa, was calculated in N (kg N ha^{-1}) and energy terms (MJ ha^{-1}) for farm j and crop i as follows:

$$\text{Pa\_N}_j = \frac{\sum_i \text{Yield}_{ij} \text{ (kg FM ha}^{-1}) \times \text{N content}_i \text{ (\%)} \times \text{Area}_{ij} \text{ (ha)}}{\sum_i \text{Area}_{ij} \text{ (ha)}} \tag{5.4}$$

$$\text{Pa\_Energy}_j = \frac{\sum_i \text{Yield}_{ij} \text{ (kg FM ha}^{-1}) \times \text{Energy}_i \text{ (MJ kg FM}^{-1}) \times \text{Area}_{ij} \text{ (ha)}}{\sum_i \text{Area}_{ij} \text{ (ha)}} \tag{5.5}$$

N concentration and energy content in the crop products were compiled from de Haan and van Geel (2013)[†] and Meul et al. (2007)[†], respectively. Maximum farm production, P_{AEprod}, was quantified as the mean Pa in terms of N and energy across the farms above the 90th percentile for each of these indicators.

Gross margin was quantified from the farm accountancy data as the difference between economic returns from crop activities and material costs (i.e. seeds, fertilisers, pesticides, energy and other). Input and output prices were deflated to 2005 prices using the consumer price index available from CBS (2015). Maximum gross margin was calculated as the mean gross margin of the farms above the 90th percentile of this indicator, and P_{AEecon} (in N or energy per ha) was calculated for the same group of farms. A similar approach was followed to identify the maximum NUE at

[†]3.3 kg N t^{-1} FM of ware potato tubers, 3.0 kg N t^{-1} FM of seed potato tubers, 1.8 kg N t^{-1} FM of sugar beets, 2.2 kg N t^{-1} FM of onion bulbs, 16.8 kg N t^{-1} FM of wheat grains and 12.6 kg N t^{-1} FM of barley grains.

[†]3.4 MJ kg FM^{-1} for potato tubers, 5.3 MJ kg FM^{-1} for sugar beet roots, 1.3 MJ kg FM^{-1} for onion bulbs, 15.5 MJ kg FM^{-1} for winter wheat grains and 15.8 MJ kg FM^{-1} for spring barley grains.

crop rotation level (kg N kg^{-1} N) which was estimated for farm j based on crop i as follows:

$$\text{N output}_j \text{ (kg N)} = \sum_i \text{Yield}_{ij} \text{ (kg FM ha}^{-1}) \times \text{N content}_i \text{ (\%)} \times \text{Area}_{ij} \text{ (ha)} \tag{5.6}$$

$$\text{N applied}_j \text{ (kg N)} = \sum_i \big(\text{Fertiliser}_{ij} + \text{Manure}_{ij}\big) \text{ (kg N ha}^{-1}) \times \text{Area}_{ij} \text{ (ha)} \tag{5.7}$$

$$\text{N input}_j \text{ (kg N)} = \text{N applied}_j \text{ (kg N)} \; + \; \text{N indigenous}_j \text{ (kg N)} \tag{5.8}$$

$$\text{NUE}_j \text{ (kg N kg}^{-1} \text{ N)} = \frac{\text{N output}_j \text{ (kg N)}}{\text{N input}_j \text{ (kg N)}} \tag{5.9}$$

P_{AEenv} was calculated based on the farms above the 90th percentile for NUE. The amount of N applied per farm (Equation 5.7) includes the amount of mineral N applied with fertilisers and the amount of mineral and organic N applied with organic manures; hence, higher NUE can be expected among farmers using only mineral fertilisers as compared to farmers using both mineral and organic N sources (Schröder, 2014). The variable 'N indigenous' in Equation 5.8 includes N inputs from seed, atmospheric deposition and biological N fixation and these were retrieved directly from the dataset. Finally, the minimum labour use (hr ha^{-1}) was calculated as the mean labour use of the farms below the 10th percentile of this indicator, and $P_{AElabour}$ was based on this group of farms. The Mann-Whitney *U* test was used to check for significant differences between the level of each indicator across different farmers' objectives and between crop yields achieved by production maximising farms (N or energy) and by farms with a different objective (gross margin, labour use or NUE). This non-parametric test was used because for most indicators the hypothesis that the sample comes from a population which has a normal distribution was rejected (results of the Shapiro-Wilk test not shown).

5.4 Results

5.4.1 Yields and yield gaps at crop level

The parameter estimates characterizing the crop-specific production frontiers are provided in Table 5.1. The gamma values indicate that the inclusion of two random errors is appropriate for all crops, except for seed potato (γ = 0.602), because most of the unexplained variability in crop yields can be attributed to the efficiency yield gap (u_{it}) rather than to statistical noise (v_{it}). Results for starch potato should be interpreted with caution as the statistical tests may be biased when $\gamma = 1$.

The importance of environmental conditions for yield variability were analysed in terms of year and soil type effects (Table 5.1). Statistically significant year effects during the period 2008 - 2012 are observed for all crops except spring onion. Crop yields in clay soils are significantly higher than in mixed and/or sandy soils only for seed potato and cereals and no statistically significant yield differences across different soil types are observed for the other crops. As for certification type, actual yields of ware, seed and starch potato are significantly lower in organic than in conventional production systems while no such difference is observed for the other crops.

Variability in crop yields is mostly associated with the management of growth-reducing rather than growth-limiting factors (Table 5.1). There is a significant relationship (first and second-order terms) between fungicide use and ware potato yields and between insecticide use and spring onion yields. Moreover, there is a significant linear relationship between fungicide use and spring onion yield (negative) and winter wheat yield (positive). The use efficiency of fungicides is positively associated with herbicide and nematicide use for starch potato and negatively associated with herbicide use for sugar beet. For starch potato, the efficiency in the use of herbicides is positively associated with nematicide use. For seed potato, yield responses to fungicide are lower in mixed than in clay soils while yield responses to herbicide are greater in mixed than in clay soils. Finally, the use efficiency of fungicides increased over time for seed potato and spring onion but declined for starch potato. A decline in the efficiency of herbicide use is observed for ware and seed potato and in insecticide use for spring onion.

Management of growth-limiting factors affects the use efficiency of inputs associated with the control of growth-reducing factors, particularly for potato production systems. For example, greater responses to herbicides are observed with greater amounts of plant available N applied for ware and seed potato. Plant available N applied also affects the use efficiency of fungicides, negatively for ware potato and positively for starch potato. Other examples include increasing N use efficiency over time for ware potato, negative quadratic relationship between yield and nutrients (N and P) for starch potato, negative linear relationship between yield and P and greater yield response to P in mixed than clay soils for spring barley, and lower yield response to N in mixed than in clay soils for winter wheat.

Table 5.1: Parameter estimates of the crop-specific stochastic frontiers estimated for Dutch arable farming (2008 - 2012). The default soil type is 'clay', except for starch potato which is 'mixed'. Significance codes: '***' 0.1% '**' 1% '*' 5%.

Variables	**Ware potato**	**Seed potato**	**Starch potato**	**Sugar beet**	**Spring onion**	**Winter wheat**	**Spring barley**
Intercept	**0.335*** **	**0.228*** **	**0.118*** **	0.046	**0.288*** **	**0.184*** **	**0.204*** **
Nitrogen	-0.045	0.001	0.231	-0.003	0.061	0.016	0.043
Nitrogen^2	-0.008	-0.003	**-1.119** **	-0.003	0.007	-0.002	0.005
Phosphorus	0.003	-0.015	0.058	-0.009	0.012	0.001	**-0.030** **
Phosphorus^2	-0.014	-0.005	**-0.336** **	-0.001	0.001	-0.001	-0.002
Fungicide	**0.058** **	-0.024	0.089	0.005	**-0.128***	**0.024***	0.021
Fungicide^2	**0.006***	0.003	-0.122	0.001	-0.006	0.002	0.001
Herbicide	0.046	0.060	0.074	-0.023	0.060	0.004	-0.018
Herbicide^2	-0.004	-0.002	-0.002	-0.004	-0.019	-0.001	0.000
Insecticide	0.011	-0.006	-0.001	-0.003	**0.041*** **		
Insecticide^2	-0.001	-0.003	0.000	0.000	**0.007***		
Nematicide	0.003	0.002	0.012				
Nematicide^2	-0.001	-0.001	0.002				
Certification_Organic	**-0.389*** **	**-0.172***	**-0.132** **	0.031	-1.753	-0.026	0.057
Year_2009	-0.002	**0.074** **	-0.002	**0.094*** **	0.019	**0.038***	**0.124*** **
Year_2010	**-0.085***	**0.071***	-0.038	0.032	-0.024	0.023	0.013
Year_2011	-0.004	0.009	0.023	**0.082*** **	0.025	**-0.112*** **	0.018
Year_2012	-0.034	0.084	**0.065***	**0.061***	-0.008	**-0.073*** **	**0.078***
Soil_Sand	-0.042	**-0.416*** **	0.032	-0.039			**-0.112*** **
Soil_Mixed		**-0.540*** **		-0.044		**-0.123*** **	-0.053
Soil_Loess				0.065			

Table 5.1: Parameter estimates of the crop-specific stochastic frontiers estimated for Dutch arable farming (2008 - 2012). The default soil type is 'clay', except for starch potato which is 'mixed'. Significance codes: '***' 0.1% '**' 1% '*' 5%. *(continued)*

Variables	**Ware potato**	**Seed potato**	**Starch potato**	**Sugar beet**	**Spring onion**	**Winter wheat**	**Spring barley**
Nitrogen × Phosphorus	0.007	0.006	0.103	0.002	0.000	0.007	0.001
Nitrogen × Fungicide	**-0.006****	-0.016	**0.466*****	0.000	0.000	0.005	-0.001
Nitrogen × Herbicide	**0.006***	**0.013****	-0.045	-0.021	-0.001	-0.002	-0.001
Nitrogen × Insecticide	-0.002	-0.004	-0.027	0.002	0.001		
Nitrogen × Nematicide	0.001	-0.003	0.001				
Phosphorus × Fungicide	0.002	**0.029****	**-0.202****	0.000	0.000	0.001	0.001
Phosphorus × Herbicide	0.004	**-0.029*****	-0.067	-0.003	0.000	0.001	-0.002
Phosphorus × Insecticide	**0.004***	0.002	**0.034****	-0.001	0.001		
Phosphorus × Nematicide	0.002	-0.002	**0.048*****				
Fungicide × Herbicide	0.000	0.000	**0.066***	**-0.011*****	0.026	-0.002	-0.001
Fungicide × Insecticide	0.001	0.001	0.000	0.000	-0.013		
Fungicide × Nematicide	0.003	0.001	**0.017****				
Herbicide × Insecticide	0.002	0.004	-0.005	0.002	-0.001		
Herbicide × Nematicide	0.003	0.001	**0.008***				
Insecticide × Nematicide	0.000	0.000	0.001				
Year × Nitrogen	**0.016*****	0.002	-0.037	0.006	-0.007	0.004	0.001
Year × Phosphorus	0.004	0.000	-0.005	0.000	0.001	-0.001	0.004
Year × Fungicide	0.000	**0.025*****	**-0.029***	0.000	**0.024****	-0.003	-0.001
Year × Herbicide	**-0.006***	**-0.017***	0.001	-0.008	-0.016	-0.002	0.005
Year × Insecticide	0.000	0.005	0.002	0.000	**-0.006***		
Year × Nematicide	-0.001	0.000	0.003				

Table 5.1: Parameter estimates of the crop-specific stochastic frontiers estimated for Dutch arable farming (2008 - 2012). The default soil type is 'clay', except for starch potato which is 'mixed'. Significance codes: '***' 0.1% '**' 1% '*' 5%. *(continued)*

Variables	**Ware potato**	**Seed potato**	**Starch potato**	**Sugar beet**	**Spring onion**	**Winter wheat**	**Spring barley**
Soil_Sand × Nitrogen	0.089	-0.046	-0.220	-0.033			-0.047
Soil_Sand × Phosphorus	0.004	0.025	-0.002	-0.002			-0.001
Soil_Sand × Fungicide	0.027	0.016	0.040	0.000			-0.004
Soil_Sand × Herbicide	0.013	0.001	-0.044	0.037			0.030
Soil_Sand × Insecticide	-0.005	-0.006	0.006	0.007			
Soil_Sand × Nematicide	0.003	-0.007	-0.003				
Soil_Mixed × Nitrogen		-0.018		-0.114		**-0.371*****	0.021
Soil_Mixed × Phosphorus		-0.005		0.004		0.040	**0.077****
Soil_Mixed × Fungicide		**-0.290*****		0.010		-0.006	-0.023
Soil_Mixed × Herbicide		**0.160***		0.002		0.025	-0.003
Soil_Mixed × Insecticide		-0.028		0.007			
Soil_Mixed × Nematicide		0.008					
Soil_Loess × Nitrogen				-0.080			
Soil_Loess × Phosphorus				0.003			
Soil_Loess × Fungicide				0.000			
Soil_Loess × Herbicide				-0.002			
Soil_Loess × Insecticide				0.028			
Model evaluation							
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	**0.104*****	**0.043*****	**0.018*****	**0.031*****	**0.113*****	**0.036*****	**0.079*****
$\gamma = \sigma_u^2 / \sigma^2$	**0.959*****	**0.602*****	**1.000*****	**0.752*****	**0.969*****	**0.936*****	**0.947*****
TE scores (%)	79.8	88.4	90.5	89.0	78.8	87.1	81.6
Sample size (n)	406	461	134	756	286	506	336

'TE scores' refer to the average of the field-specific Technical Efficiency scores (i.e. 100 minus the efficiency yield gap expressed in %) obtained for each stochastic frontier model.

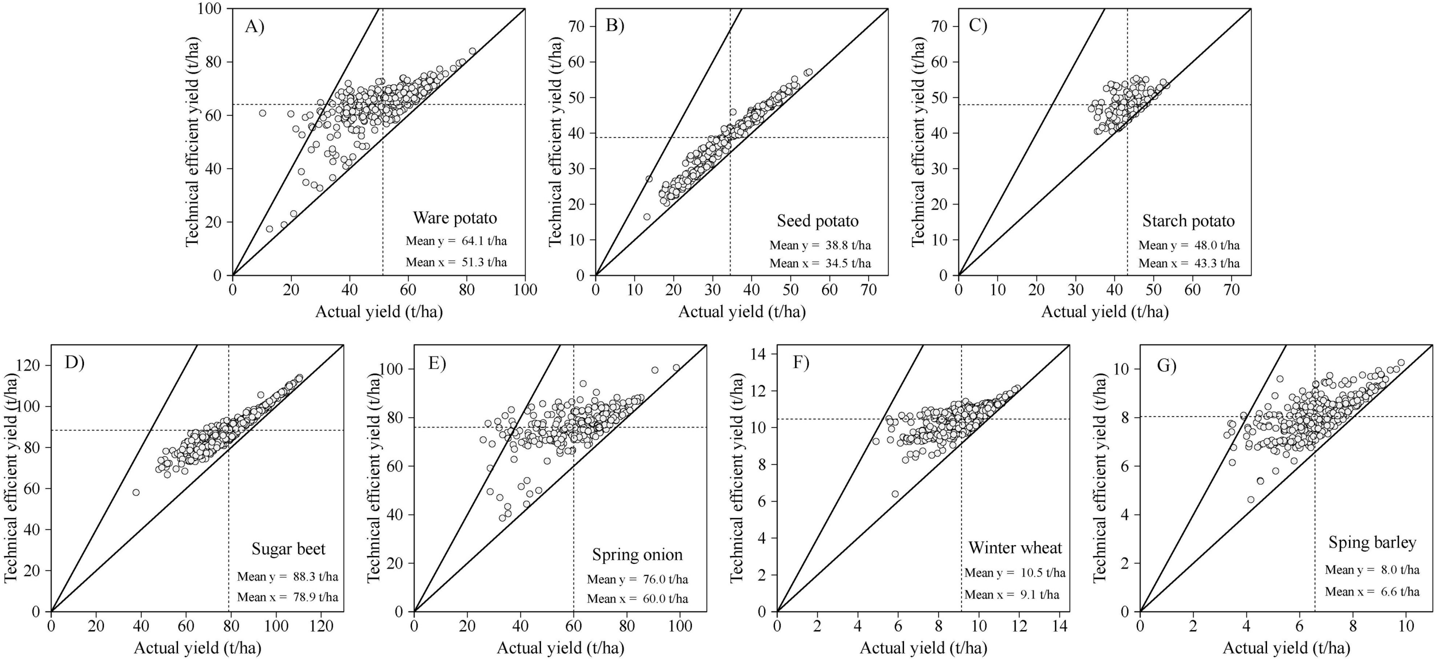

Figure 5.3: Efficiency yield gaps of A) ware potato, B) seed potato, C) starch potato, D) sugar beet, E) spring onion, F) winter wheat and F) spring barley in arable farming systems in the Netherlands. Data refers to the period 2008 - 2012; solid lines represent the 1:1 and 1:2 lines while the dashed lines represent the mean Ya (x-axis) and Y_{TEx} (y-axis).

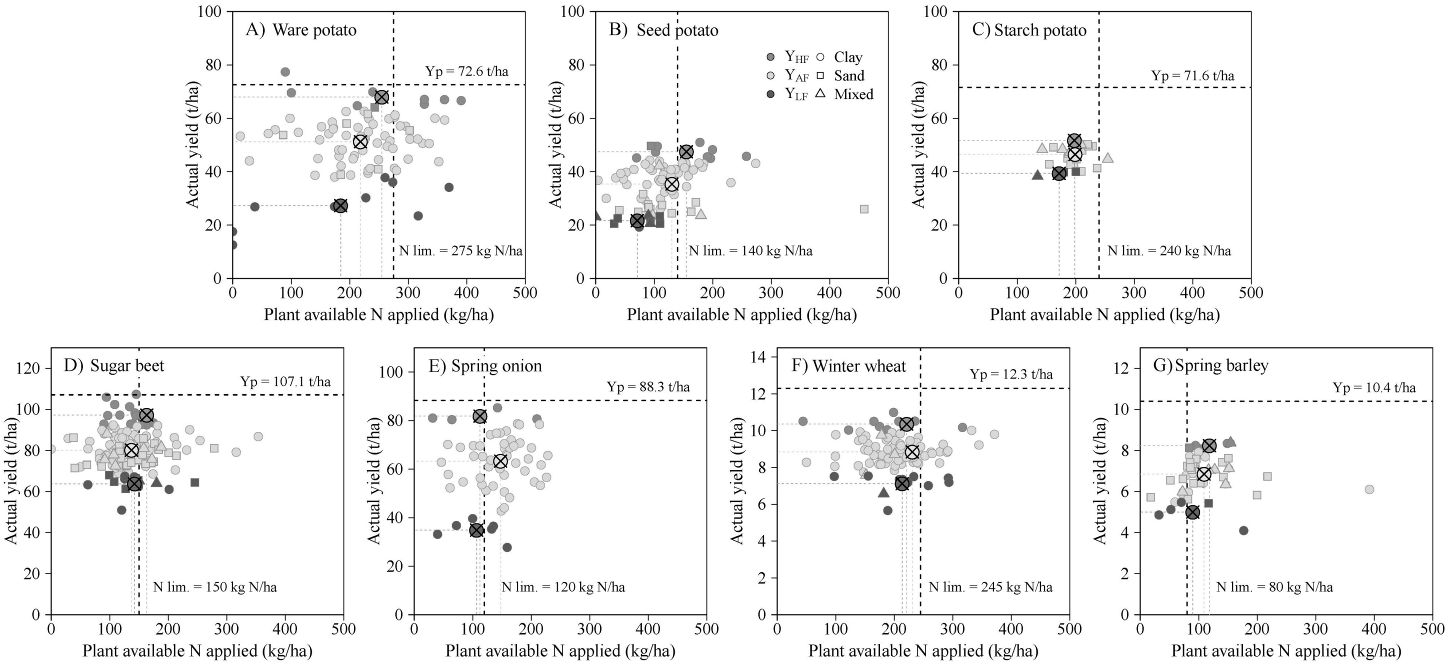

Figure 5.4: Resource (of N) and technology yield gaps of A) ware potato, B) seed potato, C) starch potato, D) sugar beet, E) spring onion, F) winter wheat and G) spring barley in arable farming systems in the Netherlands. Plant available N applied refers to the N supplied with mineral fertilisers and the N fertiliser replacement value of organic amendments. Each dot refers to a single farm in the year 2012 and soil types are differentiated into clay, sand and mixed. Horizontal dashed lines present the crop-specific Yp estimates while vertical dashed lines present the N application standards on clay soils for the year 2012 (LNV, 2009). Y_{HF} = highest farmers' yield; Y_{AF} = average farmers' yield and Y_{LF} = lowest farmers' yield.

Table 5.2: Descriptive statistics of Ya and yield gap, decomposed into the efficiency, resource and technology yield gaps (t ha^{-1} and % of Yp) for arable crops in the Netherlands during the period 2008 - 2012. Standard deviations are presented between brackets. Data of the resource yield gap refer to the stochastic frontier (Table 5.1). n.a. = not applicable

	Actual yield (Ya)		**Efficiency Yg**		**Resource Yg**		**Technology Yg**	
	(t ha^{-1})	**(%)**	**(t ha^{-1})**	**(%)**	**(t ha^{-1})**	**(%)**	**(t ha^{-1})**	**(%)**
Ware potato	51.3 (± 11.3)	70.6	12.7 (± 8.0)	17.5	3.4 (± 7.8)	4.6	5.2 (± 1.5)	7.1
Seed potato	34.5 (± 8.2)	n.a.	4.3 (± 1.5)	n.a.	8.3 (± 8.0)	n.a.	n.a. (± n.a.)	n.a.
Starch potato	43.3 (± 4.2)	60.5	4.7 (± 3.9)	6.6	1.7 (± 3.7)	2.4	22.0 (± 1.9)	30.7
Sugar beet	78.9 (± 10.8)	73.7	9.4 (± 4.4)	8.8	9.3 (± 7.8)	8.7	9.5 (± 3.8)	8.9
Spring onion	60.0 (± 13.1)	68.0	16.0 (± 10.6)	18.1	3.8 (± 8.0)	4.3	8.7 (± 2.9)	9.9
Winter wheat	9.1 (± 1.2)	74.0	1.3 (± 0.9)	10.6	0.3 (± 0.7)	2.4	1.5 (± 0.5)	12.2
Spring barley	6.6 (± 1.3)	63.5	1.5 (± 0.9)	14.4	0.6 (± 0.9)	5.8	1.8 (± 0.4)	17.3

Yield gaps in Dutch arable farming are about 30% of Yp for all crops, except starch potato and spring barley for which the gap is ca. 40% (Table 5.2, which provides data in absolute terms as well). This yield gap is mostly explained by the efficiency yield gap (around 20% of Yp) for crops such as ware potato and spring onion while its contribution is much lower (around 10%) for starch potato, sugar beet and cereals. The resource yield gap is small for all crops ranging between 2% of Yp for winter wheat and starch potato to 9% for sugar beet. Finally, the technology yield gap is as high as 31% for starch potato, 12-17% for cereals and slightly lower (less than 10%) for ware potato, sugar beet and spring onion.

The magnitude and variability of the efficiency yield gap is presented in Figure 6.4 (2008 - 2012) and that of the resource (using N as an example) and technology yield gaps in Figure 5.4 (2012 data is shown for illustration purposes). A large variability in the efficiency yield gap is observed for ware potato, spring onion, winter wheat and spring barley (Figures 6.4A, 6.4E, 6.4F and 6.4G, respectively): for these crops efficiency yield gaps as large as 50% are still observed for a number of farms and at different Ya levels. Data presented in Figure 5.4 for 2012 corroborate these results. The resource (N) yield gap is relatively small for all crops because Y_{HF}, Y_{AF} and Y_{LF} are obtained with similar applications of plant available N, which comply with the N application standards at crop level for most farms except for spring onion and spring barley (Figure 5.4). However, yield responses to plant available N applied are still evident for ware and seed potato (Figures 5.4A and 5.4B). Finally, the technology yield gap is larger for starch potato and cereals than for the other crops (see difference between Yp and Y_{HF} in Figure 5.4).

5.4.2 Insights at crop rotation level

Based on our framework, the efficiency yield gap must be explained by the timing, placing and form of the inputs applied and by rotational effects between different crops. Due to lack of data, we cannot assess empirically the contribution of the first set of factors to the efficiency yield gap but we do not expect crop management to be optimal due to biophysical and socio-economic constraints. Alternatively, we focused our analysis on the relationship between cropping frequencies and crop yields, and explored the contribution of different farmers' objectives for crop yield gaps and farm performance.

Cropping frequency and yield variability

The distribution of the proportion of the farm area allocated to potato is slightly right-skewed because few farms cultivate a large share of their land with ware, seed or starch potato (Figures 5.5A - 5.5C). However, most of the farms have a share of potato area lower than 40% which corresponds to the cultivation of one potato crop every 2.5 years. It is important to mention that there is a significant positive linear relationship between the proportion of potato and the proportion of rented land in the farm (Figure C2; Supplementary Material) meaning that effects of potato areas on crop yields are confounded with the effects of rented land. Nonetheless, and despite the low R^2 estimated (less than 20%), statistically significant (quadratic) relationships are found between the proportion of potato and the yield of ware potato, sugar beet and winter wheat. Maximum yield levels are observed at ca. 30% of potato share for ware potato and winter wheat and 25% for sugar beet.

The share of sugar beet per farm follows a normal distribution, with most farms cultivating 10 - 15% of their farm area with this crop (Figures 5.5D - 5.5F). No statistically significant relationships are found between the share of sugar beet and the yield of ware potato and sugar beet while a significant quadratic effect is observed with the yield of winter wheat, though with very low R^2 (about 2%).

Similarly to the proportion of potato, the share of spring onion in total farm size follows a distribution which is right-skewed (Figures 5.5G - 5.5I). The majority of farms cultivate spring onion on 5 - 10% of their total area and only very few farms have an area share of spring onion greater than 20%. There is a statistically significant relationship between crop yield and area share of spring onion, i.e. the yields of ware potato, sugar beet and winter wheat increase linearly with an increasing proportion of spring onion in the farm plan. However, the fitted regressions have a R^2 lower than 13%.

There is a large variation in the proportion of cereals (i.e. winter wheat and spring barley) cultivated across farms, with very few farms not cultivating cereals and other farms very specialized in cereal production (Figures 5.5J - 5.5L). Overall, the proportion of cereals follows a distribution which is right-skewed due to few specialized cereal farms. There is a significant negative relationship between the proportion of cereals and sugar beet and winter wheat yield, but again, the R^2 is very low (3%).

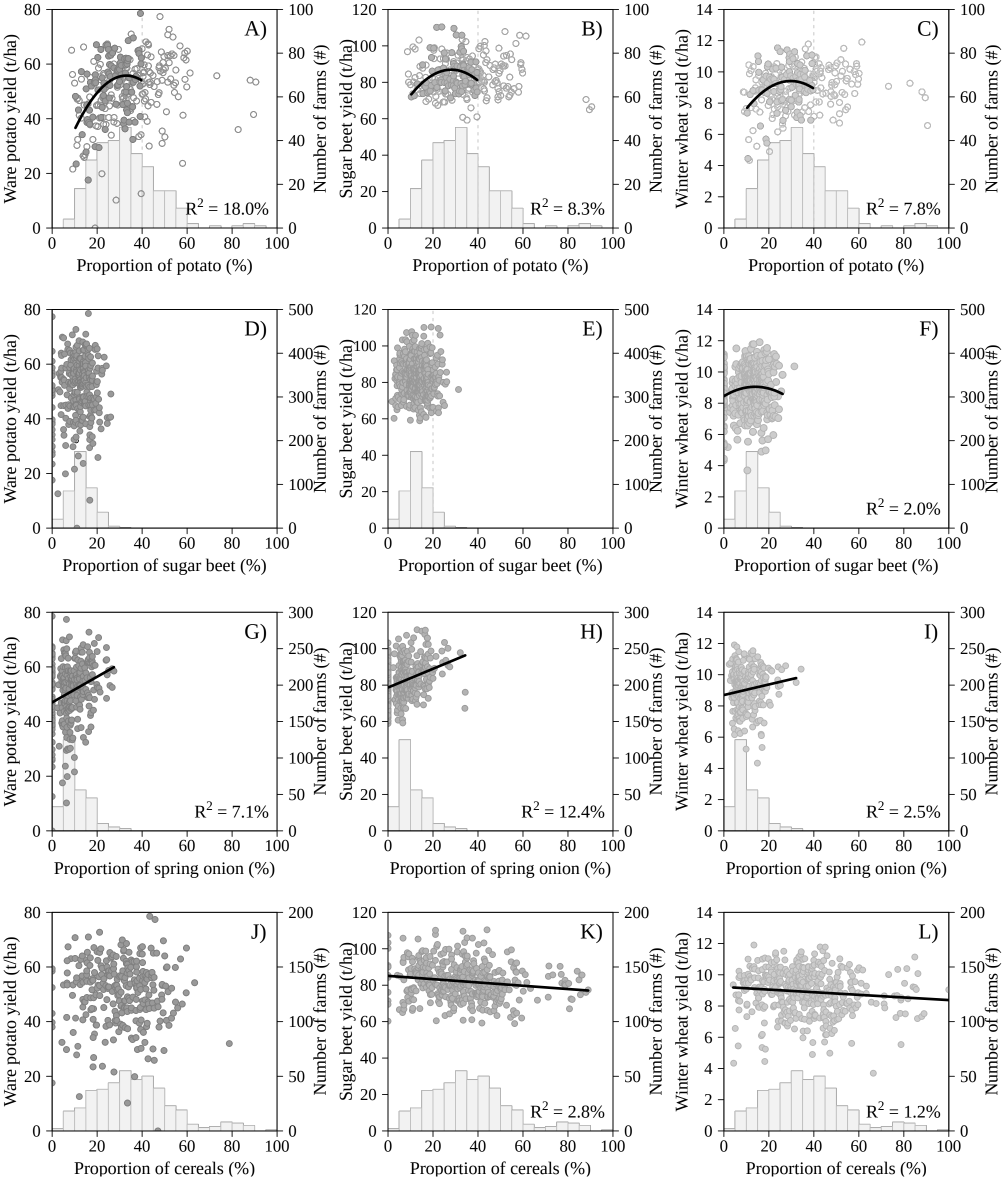

Figure 5.5: Relationship between Ya of ware potato, sugar beet and winter wheat in year t and the proportion of potato, sugar beet, spring onion and cereals in year $t-1$. Each dot represents an individual farm with area share and crop yield observed in the following periods: 2008 - 2009, 2009 - 2010, 2010 - 2011 and 2011 - 2012. Light grey dots in A), B) and C) were not included in the regression analysis due to the high proportion of rented land (more than 33.3%). The histogram represents the distribution of the share of specific crop(s). Statistically significant linear and/or quadratic regressions are presented as solid lines (parameter estimates are provided in Table C2; Supplementary Material).

Farmers' objectives and yield variability

The maximum farm production is 166 kg N ha^{-1} (Figure 5.6A) to which ware potato, sugar beet and winter wheat contribute ca. 82% (Figure 5.6B). This level of production is observed for the group of farms with highest N and energy production whereas gross margin and NUE maximisation as well as labour minimisation result in lower farm N production: about 20% less for maximum gross margin and NUE and 35% less for minimum labour use (Figure 5.6A). In terms of NUE, the maximum value observed is ca. 0.90 kg N kg^{-1} N (Figure 5.6B). Farms with maximum N production, energy production and gross margin have NUE values of ca. 0.65 - 0.69 kg N kg^{-1} N while NUE is ca. 0.55 kg N kg^{-1} N in farms with minimum labour use.

There is a large variation in gross margin between different farmers' objectives (Figure 5.6D). The maximum value of gross margin observed is 5056 € ha^{-1} for gross margin maximizing farms. Compared to gross margin maximising farms, production maximizing farms obtain 37% (N production) or 30% (energy production) lower gross margin while labour minimizing farms obtain 66% lower and NUE maximizing farms 23% lower gross margin. Moreover, only gross margin maximising farms reach the maximum gross margin observed (see error bars). Spring onion, ware potato and seed potato contribute to about 74% of the maximum gross margin, with seed potato contributing to nearly 50% (Figure 5.6E). The average labour use observed for labour minimizing farms is 21.5 hr ha^{-1} (Figure 5.6F). Farms maximizing production, economic or NUE objectives use on average much more labour (23 to 47 hr ha^{-1}) than labour minimizing farms.

There are few statistically significant yield differences for most crops between farms maximising N production and farms maximising energy production and gross margin (Figure 5.6G - 5.6L) which indicates that differences in N production across these objectives are mostly due to differences in crop area shares (Figures 5.6B and 5.6E). As compared to N production maximising farms, significantly lower yields are observed in case of ware potato, sugar beet, spring onion and spring barley for labour minimising farms and lower ware potato and sugar beet yields for NUE maximising farms. In general, labour minimising and NUE maximising farms tend to have lower yields for all crops compared to production, energy or gross margin maximizing farms. Some farms with maximum N production achieve Y_{HF} for ware potato, sugar beet and winter wheat. In contrast, gross margin maximizing farms achieve yields reaching Y_{HF} for seed potato and spring onion. Yields of seed potato, sugar beet and winter wheat close to Y_{HF} are achieved by farms with maximum NUE and Y_{HF} is not achieved by labour minimising farms for any crop.

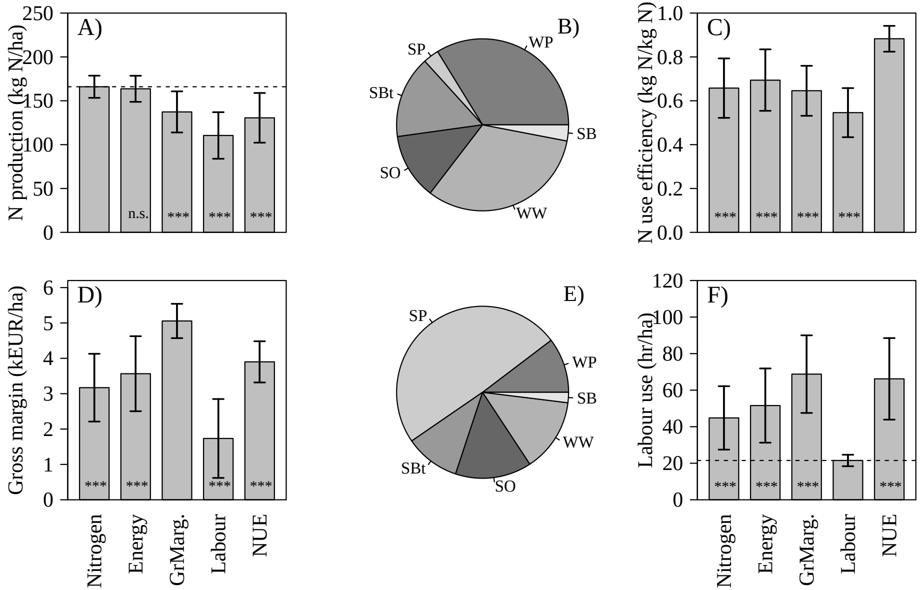

Figure 5.6: Production trade-offs for Dutch arable farms in clay soils during the period 2008 - 2012. The top four barplots show the performance of the farmers' objectives maximising N and energy production, maximising gross margin, minimising labour use and maximising NUE at crop rotation level in terms of A) N production, C) NUE, D) gross margin and F) labour use, respectively; dashed lines show the mean observed across the farms above the 90th percentile of each indicator, except for NUE which has a benchmark of 1 kg N kg^{-1} N. The six barplots presented below (G - L) show differences in Ya across different objectives; in these figures the dashed and solid lines represent the crop-specific Y_{HF} and Yp, respectively. Error bars are standard deviations of the mean. Crop area shares underlying the maximum N production and maximum gross margin are presented in the pie charts B) and E), respectively (codes: WP = ware potato; SP = seed potato; SBt = sugar beet; SO = spring onion; WW = winter wheat; SB = spring barley). The underlying data are presented in Table C3 (Supplementary Material). Significance codes: 'n.s.' = not significant; '***' significant at 5%.

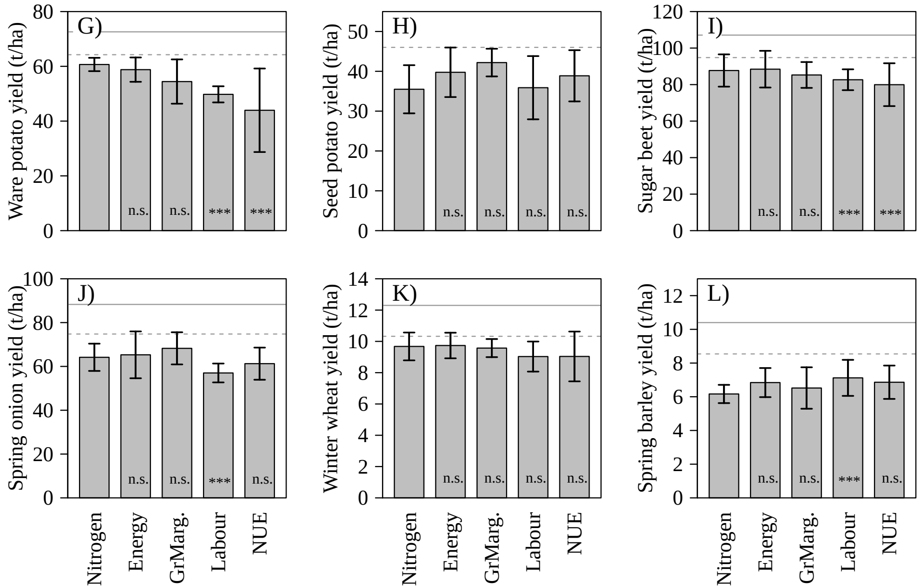

Figure 5.6: Production trade-offs for Dutch arable farms in clay soils during the period 2008 - 2012. The top four barplots show the performance of the farmers' objectives maximising N and energy production, maximising gross margin, minimising labour use and maximising NUE at crop rotation level in terms of A) N production, C) NUE, D) gross margin and F) labour use, respectively; dashed lines show the mean observed across the farms above the 90th percentile of each indicator, except for NUE which has a benchmark of 1 kg N kg^{-1} N. The six barplots presented below (G - L) show differences in Ya across different objectives; in these figures the dashed and solid lines represent the crop-specific Y_{HF} and Yp, respectively. Error bars are standard deviations of the mean. Crop area shares underlying the maximum N production and maximum gross margin are presented in the pie charts B) and E), respectively (codes: WP = ware potato; SP = seed potato; SBt = sugar beet; SO = spring onion; WW = winter wheat; SB = spring barley). The underlying data are presented in Table C3 (Supplementary Material). Significance codes: 'n.s.' = not significant; '***' significant at 5%. *(continued)*

The relationship between different indicators and yields of ware potato, sugar beet and winter wheat is provided in Figure 5.7. There is a significantly positive relationship between N production and gross margin (Figure 5.7A), whereas no significant relationship is found between N production on the one hand and NUE and labour use on the other hand (Figures 5.7B and 5.7C). Despite the large variability and low R^2 of the fitted linear regressions (maximum of 22% in Figure 5.7A), there is a positive and statistically significant relation between gross margin and the yield of ware potato, sugar beet and winter wheat (Figure 5.7D). No significant relationship is found neither between NUE and crop yield (Figure 5.7E) nor between labour use and crop yield (Figure 5.7F).

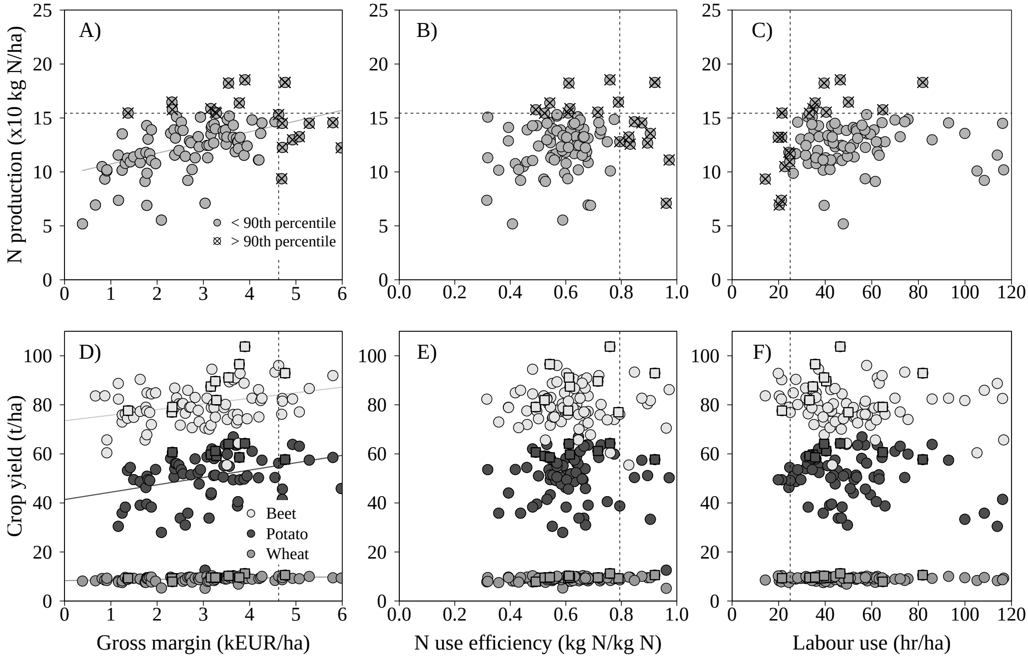

Figure 5.7: Relationship between N production at farm level (kg N ha^{-1}) and A) gross margin in k€ ha^{-1}, B) NUE in kg N kg^{-1} N and C) labour use in hr ha^{-1}; and between yield of ware potato, sugar beet and winter wheat (t FM ha^{-1}) and D) gross margin, E) NUE and F) labour use. Dashed lines represent the 90^{th} percentile value of each indicator while full lines are fitted, and statistically significant, linear regressions. Crossed symbols in A), B) and C) highlight the observations above the 90^{th} percentile for the indicators presented in the x-axis and/or the y-axis. Square symbols in D), E) and F) identify the farms above the 90^{th} percentile of N production. Parameter estimates are provided in Table C4 (Supplementary Material).

5.5 Discussion

5.5.1 Crop yield gaps in Dutch arable farming

Crop yield gaps in Dutch arable farming systems ranged between 25% (winter wheat) and 40% (starch potato) of Yp (Table 5.2). For ware potato and spring onion, the yield gap was mostly explained by the efficiency yield gap while the resource yield gap was particularly important for seed potato, and the technology yield gap for starch potato. The efficiency and the technology yield gaps are equally important for cereals (winter wheat and spring barley) and, the yield gap of sugar beet is equally explained by the three intermediate yield gaps. The small resource and technology yield gaps estimated for most crops indicate that farmers are applying input levels high enough to achieve Yp, but the high application of plant available N (sometimes above the N application standard particularly for spring onion and spring barley) is likely to induce losses and environmental impacts.

The efficiency yield gap can be explained by sub-optimal crop management in terms of timing, placing and form of inputs (not tested empirically due to lack of data, Figure 5.4) and/or to a minor extent by sub-optimal cropping frequencies from a production perspective (Figure 5.5). The timing of operational activities may be sub-optimal because of delayed harvest of the previous crop (Mazzilli et al., 2016), poor soil trafficability in key periods (Droogers et al., 1996), unfavourable weather conditions (van Oort et al., 2012) and, labour and machinery constraints on large farms (Reidsma et al., 2015a). Narrow crop rotations may also explain the relatively large efficiency yield gap as they exacerbate negative effects of biotic factors to crop yields (Scholte and s'Jacob, 1990). The strategic decision of increasing the share of tuber and root crops in the crop rotation, at the expense of cereals, is likely to be explained by farmers' objectives of maximising economic performance instead of crop production (Figures 5.6B and 5.6E). However, this might not always be possible due to production quotas in place as it is the case for sugar beet in the European Union (Figures 5.5D - 5.5F). Compliance with a particular mode of production may introduce further constraints to reduce the efficiency yield gap. We found lower crop yields in organic than in conventional production systems, similarly to de Ponti et al. (2012) and Seufert et al. (2012), but statistically significant differences were observed only for ware, seed and starch potato (Table 5.1). Nonetheless, these results must be interpreted cautiously because of the small sample size available for organic farms.

The technology yield gap is partly explained by water-limitations (Yp - Yw) during the growing season, particularly for starch potato and cereals, as irrigation is not default in the Netherlands, and certainly not for cereals. Despite the relatively humid climate and the presence of capillary rise in many areas, it is likely that water stress plays a role in some parts of the growing season, in particular on sandy and mixed soils where most of the starch potato is cultivated. In fact, higher yields were observed in clay than in sandy and mixed soil types for seed potato, winter wheat and spring barley which may be explained by less water stress in clay soils (Table 5.1). In fact, this points at Yp not being the correct benchmark for yield gaps. Differently from the other crops, rotational effects explain the technology rather than the efficiency yield gap of starch potato because most farmers cultivate this crop in a narrow crop rotation which increases the pressure from cyst nematodes. No other technologies seem to explain the technology yield gap because most farmers use up-to-date and high-yielding technologies (e.g. varieties; Rijk et al., 2013).

Beneficial effects of cropping frequencies on crop yields are difficult to show in our study (Figure 5.5), which challenges statements about the effectiveness of crop diversity and crop rotations for crop productivity (Ponisio et al., 2014). This discrepancy may be partly due to data problems because we were not able to take into account factors such as cropping sequence nor to disentangle confounding effects of rented land due to lack of data. In addition, the high input levels used by farmers (Figure 5.4) may mask negative effects of narrow crop rotations which lead to weaker and more inefficient root systems in terms of water and nutrient uptake. However, we would expect to capture any beneficial effects associated with crop rotations with our methodology if such effects would be outstanding.

Lobell et al. (2009) and van Ittersum et al. (2013) suggested that the exploitable yield for farmers is ca. 80% of Yp (or Yw), since the remaining 20% of the yield gap will not be closed because of economic and environmental considerations and management constraints of the farmer. Our analysis partly supports the aforementioned statement because the difference between crop yield across the different farmers' objectives and crop-specific Yp ranged between 20 and 40% (Figures 5.6G - 5.6L). Ware potato, sugar beet, spring onion and winter wheat yields were on average 20 to 25% lower than Yp while yields of spring barley were on average 35% lower than Yp for gross margin and NUE maximising farms and labour minimising farms (Figures 5.6G - 5.6L). However, most crop yields are not significantly different between farms with the highest production compared to farms performing best in economic (gross margin) objectives. Also Figure 5.7D suggests little trade-off between crop yields and gross margin. Differences in N production across farmers' objectives (Figure 5.6A) are mainly due to differences in crop area shares (e.g. Figures 5.6B and 5.6E) which indicates that maximisation of N (or energy) production at crop rotation level requires a particular cropping pattern rather than closing yield gaps at crop level.

5.5.2 Methodological considerations

The methodology applied builds largely on the framework introduced in Chapter 3. The most obvious improvement to the current analysis would be to estimate the inefficiency effects model (Battese and Coelli, 1995) in order to explain quantitatively the efficiency yield gap for each crop (e.g. Mazzilli et al., 2016). This was not done in this paper for two reasons: 1) the dataset lacked detailed field level data referring to time, form and space of inputs as well as preceeding crop and field history and 2) for most crops there were not enough degrees of freedom to analyse the influence of individual crop shares in year $t-1$ on the efficiency yield gap in year t. Therefore, further improvement of the current crop level analysis requires a detailed crop-specific dataset of different farms containing both input-output data and field specific information (e.g. previous crop, cropping sequence, tenure).

The general consensus that in developing countries an 'intensification' pathway is necessary to achieve food security and that in developed countries the (environmental) sustainability aspects are most relevant requires a more comprehensive understanding of the concept of resource use efficiency. One important aspect for Dutch arable farming not addressed in this article is the scope to improve resource use efficiency, and hence environmental and economic performance, through the reduction of input use without compromising crop yields (e.g. Ma et al., 2012; de Koeijer et al., 2002; Reinhard et al., 1999). Although most farms comply with the N application standards at crop level (Figure 5.4) and at farm level (Figure C3; Supplementary Material), our results may suggest a potential to decrease the application of plant available N for many farms without reducing crop yields. The only two exceptions are spring onion and spring barley, crops for which the current N application standards may be too restrictive to allow farmers to achieve high yields and close yield gaps.

Further methodological improvements could be achieved through the application of linear programming to compute production trade-offs for alternative farmers' objectives and under different scenarios. As an example, this could be implemented from a economic perspective through an optimization analysis in which P_{AEecon} is estimated under different input/output price ratios (e.g. Kanellopoulos et al., 2014;

Jatoe et al., 2008) or from an environmental perspective in which the impact of different policies on P_{AEecon} is explored (building upon Lassaletta et al., 2014). The construction of trade-off curves is important to make explicit the yield penalty and environmental consequences associated with different management practices under specific objectives (Bos et al., 2016). Moreover, output distance functions considering multiple outputs and inputs at the farm level (Emvalomatis et al., 2011; Reidsma et al., 2009b) can be estimated to study interactions between crops and between land, labour and capital and crop production, and to assess the magnitude and determinants of a 'system-wide yield gap' (Henderson et al., 2016).

5.6 Conclusion

Yield gaps of the most important arable crops cultivated in the Netherlands were decomposed into efficiency, resource and technology yield gaps. The crop level analysis focused on ware, seed and starch potato as well as sugar beet, spring onion, winter wheat and spring barley. Yield gaps were ca. 30% of potential yield (Yp) for all crops, apart from starch potato and spring barley for which it was ca. 40%. We found that most of the yield gap is explained by the efficiency yield gap for ware potato and spring onion, by the resource yield gap for seed potato, by the technology yield gap for starch potato and by both the efficiency and technology yield gaps for sugar beet and cereals. The efficiency yield gap ranged between 9% (sugar beet) and 18% (spring onion and ware potato) of Yp. The resource yield gap was less than 10% for all the crops and the technology yield gap ranged between 7% (ware potato) and 31% (starch potato). These results show that farmers are generally using technologies and input levels that allow them to reach Yp and that further yield improvements, especially for ware and seed potato and spring onion, must be derived through better crop management practices as well as through improved crop rotations (and water management in case of starch potato and cereals). However, rotational effects tend to disappear when high input levels are applied, except when some pests (e.g. nematodes) cannot be controlled with plant protection agents.

The results regarding the effects of cropping frequency to crop yields were not very conclusive, partly due to scant information on crop rotations (e.g. cropping sequence) and confounding factors such as rented land. Yet, we consider the lack of clear effects of cropping frequency an interesting result because it shows that agronomic principles become obscured/muddled at 'systems level' given the interacting factors at crop rotation and farm level (e.g. economies of scope, financial situation of the farm, access to rented land). Significant differences in farm performance, but not in crop yields, were found across different farmers' objectives. For instance, gross margin maximising, labour minimising and N use efficiency maximising farms can improve N production by about 17, 34 and 21%, respectively. Increasing N production at farm level needs to be achieved through increases in actual yields and through changes in the crop rotation plan. Further, there is potential to increase N use efficiency at crop and crop rotation level by reducing N inputs while maintaining crop yields.

Further research about the crop and farm level determinants of the efficiency yield gap is necessary to understand the underlying limiting factors to crop production in Dutch arable farms. For that purpose, a smaller sample of farms (e.g. Mandryk et al., 2014) must be monitored in-depth and over time. Improving the understanding of the efficiency yield gap may also provide operational insights for precision agriculture (e.g. Reidsma et al., 2015a). In addition, it seems necessary to develop and test suitable methodologies for the analysis of resource use efficiency at crop level to assess by how much input levels could be decreased without incurring yield losses as well as for disentangling the interactions between different crops and available resources at farm level.

5.7 Acknowledgements

Micro economic data of the Dutch FADN system (Wageningen Economic Research) from the years 2008 - 2012 was used in this research. Results shown are and remain entirely the responsibility of the authors and they do not represent Wageningen Economic Research views. We acknowledge Ing. Bert Rijk, Dr. René Schils and MSc. Anton Bartelen (Plant Production Systems, Wageningen University) for stimulating discussions during the development of this manuscript and Dr. Hennie van der Veen and Ir. Ruud van der Meer (Wageningen Economic Research) for organizing and sharing the farm data used in this study. Prof. Dr. Alfons Oude Lansink (Business Economics, Wageningen University) is acknowledged for the clarifications and suggestions provided on the productivity and efficiency analysis. Dr. Jaap Schröder, Dr. Hein ten Berge (Plant Research International, Wageningen University), Dr. Tanja de Koeijer and Ing. Henri Prins (Wageningen Economic Research) are acknowledged for their important insights about nutrient management, and regulations, in the Netherlands. This study was carried out within the project 'Mapping for Sustainable Intensification' which is part of a strategic programme funded by Wageningen University.

CHAPTER 6

Is labour a major determinant of yield gaps in sub-Saharan Africa? A case study for cereals in Southern Ethiopia

This chapter will be submitted as:
Silva, J.V.; Baudron, F.; Reidsma, P.; Giller, K.E. 2017. Is labour a major determinant of yield gaps in sub-Saharan Africa? A case study for cereals in Southern Ethiopia.

Abstract

The objective of this paper is to understand the role of labour in explaining the yield gap of cereals at both crop and farm levels across smallholder farms in Southern Ethiopia. A household survey containing detailed information of labour use at crop and farm level of ca. 100 farms in a maize-based system around Hawassa and ca. 100 farms in a wheat-based system around Asella was used for this purpose. Stochastic frontier analysis was combined with principles of production ecology to decompose maize and wheat yield gaps. For both crops, nearly half of the yield gap was attributed to the technology yield gap, indicating sub-optimal crop management to achieve the water-limited yield (Yw). The efficiency yield gap was ca. 20% of Yw for both crops and it was negatively associated with sowing date and with the proportion of women labour for sowing in case of maize and with the proportion of hired labour used for sowing and weed control in case of wheat. The resource yield gap was less than 10% of Yw for both crops due to small differences in input use between highest and lowest yielding farms. The contribution of capital and farm power availability to crop yields, input use and labour use was analysed at the farm level. Crops cultivated in Hawassa showed complementary demand for labour throughout the year while crops cultivated in Asella showed strong competition for labour, and potential trade-offs at farm level, during sowing, hand-weeding and harvesting months. Oxen ownership was a proxy of capital availability, but not farm power, in Hawassa and a proxy for both capital availability and farm power in Asella. Farms with more oxen used more N for maize in Hawassa and cultivated more land in Asella, which is indicative of an intensification pathway in the former and an extensification pathway in the latter. The lower land:labour ratio and types of crops cultivated in Hawassa compared to Asella explained these different strategies in the two sites. In both sites, while gross margin per unit area increased with increasing crop yield and farm N production, gross margin per labour unit decreased after an optimal level of crop yield and farm N production. This implies that from a labour productivity perspective, aiming for increased yield and farm N production levels is not economically rational. We conclude that labour (and farm power) is not a major determinant of maize yield gaps in Hawassa but is a major determinant of wheat yield gaps in Asella.

6.1 Introduction

Despite the availability of many agricultural technologies to increase crop yields, productivity in sub-Saharan Africa has remained stagnant (FAO, 2014). This is reflected in the large yield gaps for almost all crops in all regions (Tittonell and Giller, 2013). The main resources available to smallholders for farming are their land and labour, and the ability to invest in technology depends largely on access to capital. In this context, yield gaps are a consequence of poor soil fertility and nutrient availability (Vanlauwe et al., 2014) and possibly a consequence of trade-offs at farm scale, regarding where and how labour is invested (e.g. Kamanga et al., 2014). This is particularly important given the stagnation, or even decline in some regions, of available farm power across the continent (Baudron et al., 2015).

Current discourses on agricultural development focus on land productivity but seldom consider labour productivity (e.g. Woodhouse, 2010). However, it is unclear whether land or labour is the most limiting factor to smallholder production. The land:labour ratio and the seasonality of labour demand determine whether land or labour is the 'binding constraint' (Erenstein, 2006). Farmers are more likely to increase production per unit area through the use of more intensive production practices, i.e. intensification, when the land:labour ratio is low (Jayne et al., 2014) and when there is little temporal overlap between labour activities. Conversely, increases in production through area expansion while maintaining or reducing input levels per unit area, i.e. extensification, are more likely to occur when the land:labour ratio is high and there is strong competition for labour in specific periods (e.g. Leonardo et al., 2015; Baudron et al., 2012).

Most households in rural areas of Ethiopia cultivate less than two ha of land and possess cattle (CSA and WB, 2013). Moreover, about 34% of the population lives with less than 1.90 US\$ person^{-1} day^{-1} (WB, 2017). Households are often capital constrained and most labour operations are performed either manually or with animal traction (Baudron et al., 2015). Indeed, a recent analysis for Ethiopia based on the Living Standards Measurement Study (LSMS; Sheahan and Barrett, 2017) indicates that only ca. 30% of the sampled farms used herbicides and there was no record of households owning or renting a tractor. Ploughing one ha of land in the Ethiopian highlands with animal traction takes up to 50 hours 'per pass' (Aune et al., 2001) and labour remains an important productive factor for operations such as hand-weeding (Amare, 2014; Workayehu and Wortmann, 2011).

The main objective of this study is to understand the role of labour in explaining yield gaps in cereal-based farming systems in Southern Ethiopia. Understanding whether, how and when labour can be a major constraint to smallholder production requires that yield gaps are analysed at the farm level. This is important because farmers make decisions on resource allocation and prioritization of crop management across the entire farm. A detailed household survey conducted in 2012 across smallholder maize- and wheat-based farms in Hawassa and Asella, respectively, was analysed for this purpose.

6.2 Framework of analysis

The main concepts for explaining yield gaps at crop level have been documented elsewhere (Chapters 3 and 5). In short, the yield gap, i.e. the difference between water-limited (Yw) and actual farmers' yields (Ya), can be decomposed into efficiency, resource and technology yield gaps if methods of frontier analysis are combined with concepts of production ecology and applied to individual farm level data (Figure 6.1A). The efficiency yield gap indicates by how much output can be increased for a given input level and can be explained by differences in timing, frequency, space and form of inputs applied, which in turn are affected by prioritization of crop management, availability of farm power and/or labour quality. The resource yield gap captures the contribution of sub-optimal input quantities required to achieve highest farmers' yields (Y_{HF}). Finally, the technology yield gap captures the difference between Yw and Y_{HF}, which can be attributed to resource yield gaps of specific inputs and/or differences in resource use efficiency between technologies used by farmers and agronomic 'best practices'.

In this paper, we apply these concepts to a production system in which resource availability, and allocation, may lead to trade-offs at farm level when closing the yield gap of multiple crops. A first step of analysis is to understand whether the availability of animal traction limits the area cultivated (Figure 6.1B), which is still the case in many African smallholder farming systems (Ollenburger et al., 2016; Leonardo et al., 2015; Baudron, 2011). Second, the impact of alternative resource allocation strategies used by farmers to crop yield gaps is explored at farm level (Figure 6.1C). These may differ for different crops both in the quantity of input used (Tittonell et al., 2007) and in timeliness of the operations performed (Kamanga et al., 2014). Limited availability of labour, capital and land, and their prioritization to different activities, induces trade-offs at farm level (Figure 6.1D). Capital constraints may be the driver of production trade-offs when land:labour ratios are low while labour may limit crop production when the land:labour ratio is high and when there is a lack of capital to adopt labour-saving technologies.

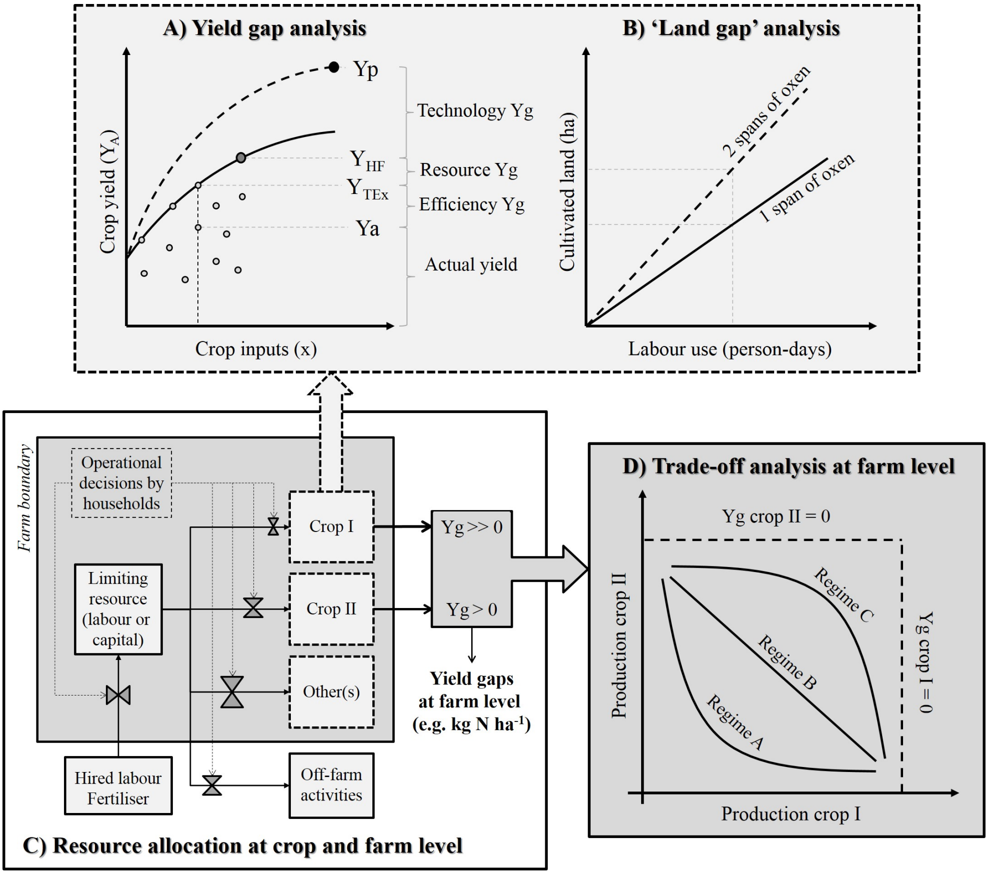

Figure 6.1: Framework of the analysis conducted to understand the role of labour to crop production across smallholder farms in Southern Ethiopia: A) explaining yield gaps at crop level (Chapter 3), B) effects of labour use and spans of oxen available on cultivated area, C) labour and capital allocation at crop and farm levels, and D) trade-offs due to resource allocation at farm level. In D), regime A depicts a situation of strong competition, regime B a situation of substitutability through inverse proportionality and regime C a situation of possible complementarity between the production of crops I and II (Tittonell et al., 2015). Yp = climatic potential yield; Y_{HF} = highest farmers' yield; Y_{TEx} = technical efficient yield at a specific input level; Ya = actual yield; Yg = yield gap.

6.3 Material and methods

6.3.1 Household survey

Individual farm data was collected in 2012 within the project 'Farm Mechanisation and Conservation Agriculture for Sustainable Intensification' (FACASI, www.facasi.act-africa.org). The purpose of the farm survey was to map the potential demand for mechanisation in Eastern and Southern Africa. A total of 200 farms were interviewed in Southern Ethiopia (100 interviews per site: Hawassa and Asella) using a semi-structured questionnaire requesting detailed information on labour use at crop and farm level. Households were selected using a systematic sampling procedure in each village based on transect routes across the village in which every fourth household, on alternate sides of the track, was sampled. In case one of the selected households was not available, the next one was selected.

Hawassa is located in the Rift Valley 1708 m above sea level, while Asella is located in the Southern highlands 2430 m above sea level. This difference in elevation affects climatic conditions (Figure 6.2) and the type of crops cultivated in each site. Minimum and maximum temperatures are higher throughout the year in Hawassa compared to Asella, while the amount and intra-annual variation of solar radition is rather similar between both sites. The average precipitation is about 900 mm yr^{-1} in both sites, ca. 80% of which falls between April and October in Hawassa and between March and September in Asella. Fertile fluvisols and luvisols are the dominant soil types around Hawassa while luvisols and vertisols, which have problems of workability when wet, are the major soil types around Asella (Dewitte et al., 2013). Yield responses to N, but not P, are often observed in Hawassa (TAMASA, unpublished data) and to both N and P in Asella due to the P-fixing soils in this site (Habte et al., 2014).

Farm systems differ across sites in the number and types of crops (Table 6.1 and Figure D1; Supplementary Material) and in livestock ownership, in particular the number of oxen owned (Figure D2; Supplementary Material). The main crops in Hawassa are maize (*Zea mays* L.), bean (*Phaseolus vulgaris* L.) and enset (*Ensete ventricosum* Bruce), mostly for home consumption. In contrast, small grains such as wheat (*Triticum aestivum* L.), barley (*Hordeum vulgare* L.) and tef (*Eragrostis tef* (Zuccagni) Trotter) are common in Asella due to favourable agroecology (cool weather and clay soils) and markets (e.g. presence of breweries and large national demand for malt barley). Legumes such as pea (*Pisum sativum* L.) and faba bean (*Vicia faba* L.) are also common and mostly used for home consumption.

Table 6.1: Descriptive statistics of yields, mineral N and P applied, area cultivated and labour use per crop in Hawassa and Asella, Southern Ethiopia (year 2012). No mineral N or P is applied to enset, instead farmers apply manure and/or compost to this crop. Standard deviations are presented between brackets.

	Farms (*n*)	Crop yield ($t\ ha^{-1}$)		N applied ($kg\ N\ ha^{-1}$)		P applied ($kg\ P\ ha^{-1}$)		Crop area (ha)		Total labour (person-days ha^{-1})	
Hawassa											
Maize	93	1.6	(1.1)	65.5	(76.3)	26.6	(67.3)	0.5	(0.4)	92.8	(73.1)
Bean	34	1.0	(0.8)	25.5	(40.5)	12.3	(15.5)	0.2	(0.2)	100.1	(78.0)
Enset	51	0.6	(0.8)	0.0	–	0.0	–	0.3	(0.3)	92.2	(81.0)
Asella											
Wheat	100	2.6	(1.7)	46.2	(24.5)	32.4	*(15.5)*	1.1	(1.6)	72.5	(38.0)
Barley	60	2.1	(0.9)	30.9	(20.3)	27.6	(12.9)	0.7	(0.6)	74.5	(46.2)
Tef	28	1.2	(0.4)	23.8	(13.3)	20.4	(8.8)	0.5	(0.2)	91.1	(52.6)
Sorghum	16	2.1	(0.8)	6.1	(10.8)	6.1	(11.9)	0.4	(0.2)	66.6	(37.8)
Pea	36	1.5	(0.7)	18.3	(11.1)	18.4	(11.4)	0.5	(0.3)	52.5	(22.1)
Faba bean	26	1.8	(0.8)	23.9	(27.9)	22.4	(26.2)	0.3	(0.1)	82.4	(40.6)

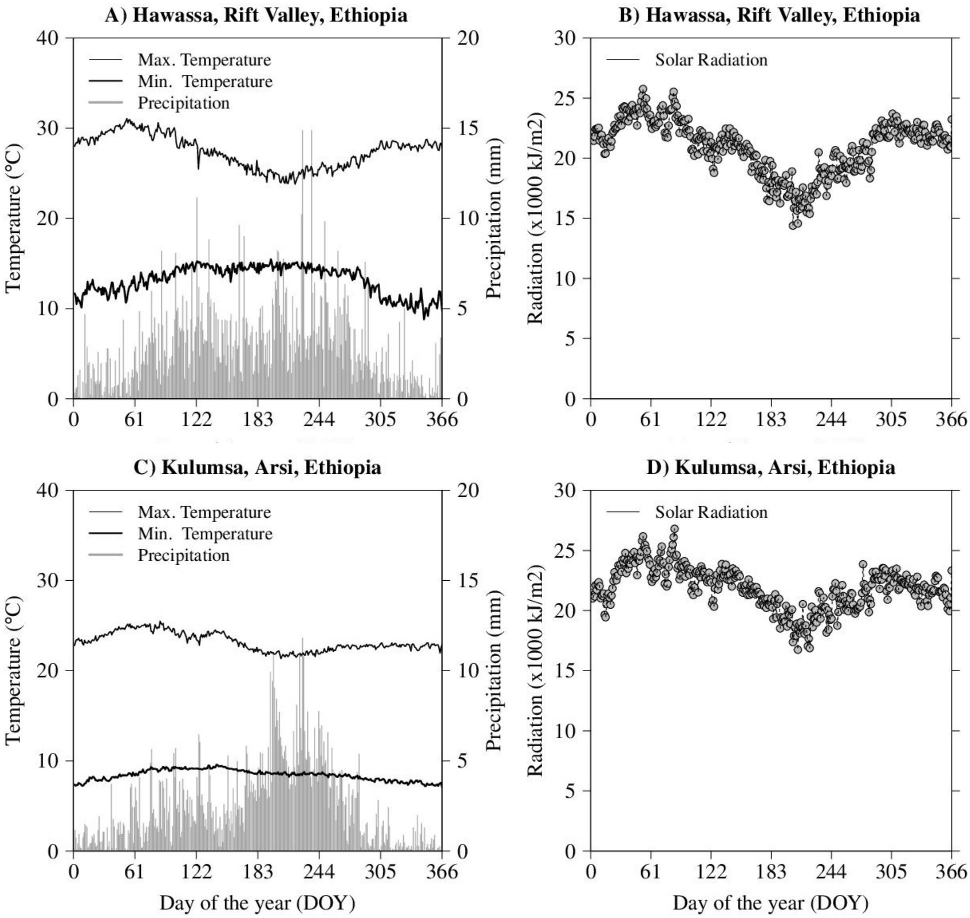

Figure 6.2: Mean daily maximum and minimum temperatures (°C), rainfall (mm) and solar radiation (kJ m^{-2}) for the period 1998 - 2012 in A-B) Hawassa (solar radiation data was taken from Arsi-Negele station), Rift Valley and C-D) Kulumsa (the nearest station to Asella), Arsi, Ethiopia. *Source*: Southern Agricultural Research Institute (SARI) and Global Yield Gap Atlas (www.yieldgap.org).

Labour calendars were developed to identify periods of labour peaks at farm level and the seasonality of labour demand (e.g. Stone et al., 1990). For each crop, labour and animal power used per month ($1 \leq t \leq 12$) were assessed as follows:

$$\text{Labour use}_{ot} = \frac{\sum_i \left[\text{Family labour}_{iot} \text{ (person-day)} + \text{Hired labour}_{iot} \text{ (person-day)} \right]}{\sum_i \text{Area}_{io} \text{ (ha)}} \quad (6.1)$$

where 'Family labour' and 'Hired labour' refer to the person-days (standardized to an 8 hour working day) used for management operation o in month t by household i and 'Area' to the area cultivated with management operation o by household i. The management operations considered were land preparation, crop establishment (i.e. sowing), fertiliser application (both mineral and organic), hand-weeding and harvesting. Total labour use per month was further estimated by summing the monthly labour use per ha for the different management operations. The same approach was used to analyse the seasonality of animal power.

6.3.2 Estimating and explaining yield gaps

Stochastic frontier analysis

The yield gap analysis focused on maize in Hawassa and wheat in Asella as these were the main crops in each site (Table 6.1). We assumed the relationship between crop yield and a vector of inputs defined according to principles of production ecology to be approximated by a translog functional form and regressed the efficiency yield gap on a set of explanatory variables. The generic formulation of the inefficiency effects stochastic frontier model used is as follows:

$$ln\, y_i = \alpha_0 + \sum_{k}^{K} \beta_k \, ln\, x_{ki} + \frac{1}{2} \sum_{k}^{K} \sum_{j}^{K} \theta_{kj} \, ln\, x_{ki} \times ln\, x_{ji} + v_i - u_i \tag{6.2}$$

$$u_i = \beta' \, z_i + \epsilon_i \tag{6.3}$$

where y_i denotes the yield of maize or wheat reported in farm i, x_i a vector of agronomic inputs k and j (both first- and second-order terms) assumed to explain variability in crop yields and z_i a vector of management and labour quality variables assumed to explain the efficiency yield gap. Moreover, two independently distributed random errors are included in Equation 6.2 to capture random noise, $v_i \sim \text{N}(0, \sigma_v^2)$ and technical inefficiency, $u_i \sim \text{N}^+(\beta' z_i,\, \sigma_u^2)$ and one error term is included in Equation 6.3, $\epsilon_i \sim \text{N}(0, \sigma_\epsilon^2)$, with the distribution of W_i being bounded below by the variable truncation point $-\beta' z_i$ (Kumbhakar and Lovell, 2000; Battese and Coelli, 1995). The parameters α_0, β_k, θ_{kj}, b_0, β', v_i and u_i were estimated in a single-step procedure using maximum likelihood (R package *frontier*; Coelli and Henningsen, 2013). The efficiency yield gap, and Y_{TEx}, were estimated as follows:

$$\text{EffYg}_i = 1 - exp(-u_i) \tag{6.4}$$

$$\text{Y}_{\text{TEx}_i} = \text{Ya}_i \times exp(-u_i)^{-1} \quad (6.5)$$

Four different stochastic frontier models (Equations 6.2 and 6.3) were estimated for maize and wheat. Model I was estimated without inefficiency effects and it was used to assess the sign of the parameter estimates and to quantify the efficiency yield gap (cf. Figure 6.4). Models II and III included inefficiency effects related to the frequency and timing of management operations, respectively, and model IV included inefficiency effects related to the quality of labour used (source and gender).

The stochastic frontier models were estimated using farmer reported data on crop yields (y_i) and on growth-defining, -limiting and -reducing factors (x_i) for unique farm × crop combinations. Dry matter yields of maize and wheat (kg DM ha^{-1}) were used as dependent variables. These were calculated by assuming a standard moisture content of 15.5 and 13.5% for respectively maize and wheat (www.yieldgap.org). Differences in growth-defining factors were accounted for by using categorical variables for different communities (maize: Wondo Genet vs. Hawassa Zuria; wheat: Haro Bilalo vs. Gara Silingo). The only growth-limiting factor considered for both crops was the rate of N applied (kg N ha^{-1}). We were not able to account for the effects of P applied due to collinearity between this variable and N applied (as the main fertiliser used in Ethiopia is diammonium phosphate, DAP), and of organic amendments due to the low number of observations. Growth-reducing factors considered for maize were labour used for the first hand-weeding (person-days ha^{-1}), labour used for the second hand-weeding (person-days ha^{-1}) and inter-cropping with bean (yes or no) while for wheat these included labour used for the first hand-weeding (person-days ha^{-1}) and labour used for herbicide application (person-days ha^{-1}). The latter is assumed to be a proxy for the actual amount herbicide applied. Our analysis overestimates the efficiency yield gap because we were not able to control for differences in crop varieties, sowing densities and soil conditions. Input-output variables were mean-scaled and ln-transformed prior to the analysis.

The determinants of the efficiency yield gap (z_i) related to the frequency of management operations were the number of ploughing (#), sowing (#), fertiliser (#) and weed-control operations (#). The determinants related to the timing of management operations were sowing date (month), date of first fertiliser application (month after sowing) and date of first weed control (month after sowing). The number of sowing operations and sowing date was only considered for maize due to collinearity effects between these variables and the date of first fertiliser application and first weed control for wheat. The sowing window for wheat in Asella is narrow and most farmers perform a basal fertilizer application at sowing and a first weeding up to two weeks after sowing. Finally, the determinants related to labour quality were the proportion of hired labour used for land preparation (%), the input of animal power used for land preparation (animal-days ha^{-1}) and the proportion of hired labour (%), female labour (%) and child labour (%) used for sowing and for weed control. No data transformations were applied to the z_i variables.

Yield distribution and response to inputs

The resource yield gap due to N, P and labour was studied by comparing yields and input use among farmers with the highest, average and lowest yields. Farmers with highest yields were identified as the observations above the 90th percentile of Ya and the highest farmers' yield (Y_{HF}) was calculated as the mean Ya for these observations. A similar approach was used to identify farmers with smallest yields (Y_{LF}, mean Ya for observations below the 10$_{th}$ percentile of Ya) and average farmers' yields (Y_{AF}, mean Ya for observations between 10th and 90th percentile of Ya). Significant differences in crop yields, input use and labour use between Y_{HF}, Y_{AF} and Y_{LF} were tested using analysis of variance (ANOVA) followed by a Tukey HSD post-hoc test (p-value $\leq$ 0.05; *agricolae* R package de Mendiburu, 2015). Observations with organic fertiliser were not considered to avoid confounding.

The univariate relationships between crop yield and input and labour use was further studied using boundary line analysis (Shatar and McBratney, 2004; Webb, 1972). For this purpose, the independent variables were sorted in ascending order and the following model (cf. Fermont et al., 2009) was fitted to the observations with largest yield per unit input (boundary points; Schnug et al., 1996):

$$y_{\text{boundary}} = \frac{y_{max}}{1 + (\text{K} \times exp(-\text{R} \times x))} \tag{6.6}$$

where y_{max} is the Y_{HF} estimated for each crop, x is the independent variable and, K and R are constants to be estimated using nonlinear least squares (*nls* function in R; R Core Team, 2013). Two observations with maize yields greater than 8 t ha^{-1}, one observation with more than 600 kg N ha^{-1} applied for maize and another one with ca. 200 kg N ha^{-1} applied for wheat were excluded from the analysis. Results for maize and wheat are presented in Figure 6.5 and results for other crops are presented in Supplementary Material (Figures D7 and D8).

Simulated Yw and Yp

Yw was used as the biophysical benchmark for Ya, and to quantify the technology yield gap, because farmers in both sites operate under rainfed conditions. Yw represents the maximum yield that can be achieved under rainfed conditions with the use of modern varieties. Estimates of Yp were used as well to report the yield penalty associated with sub-optimal water supply (Yp - Yw). The source of Yp and Yw used in this study was the Global Yield Gap Atlas (GYGA, 2016) and information was only available for maize and wheat. Simulations for maize and wheat refer to the period 1999 - 2012 and 1998 - 2011, respectively.

6.3.3 Embedding yield gaps within farm dynamics

Further insights into the drivers of yield gaps and their crop level determinants were obtained by conducting different analyses at farm level. These captured the contribution of resource availability, and allocation, to management practices explaining the yield gap as well as the relationship between crop yield and farm performance indicators.

Resource availability at farm level

Farms were classified according to the number of pairs of oxen owned. These resulted in three different groups in Hawassa (none, one pair and two pairs or more) and only two groups in Asella (one pair and more than two pairs). This classification was done to test whether oxen ownership relates to labour, land and capital availability, and to analyse its relationship with input use and crop yield. Significant differences between groups of farms with different pairs of oxen were tested for farm assets (index), farm size, maize and wheat yields and, input and labour use for maize and wheat using ANOVA and the Tukey HSD post-hoc test, as specified above. An index estimating the value of farm assets was developed based on the number of productive assets (e.g. ploughs, hoes and water pumps but not the number livestock) possessed by each household and their relative economic value (ILRI, 2011; BMGF, 2010). Cultivated area refers to the total area (in ha) cultivated by each household, input use includes N application rates to maize and wheat, and labour use includes animal power used for ploughing and sowing and human labour used for the 1st weeding also for maize and wheat.

Resource allocation to different crops

Production trade-offs due to prioritization of resources were identified using the individual farm data. In what follows, the term 'main crop' denotes maize in Hawassa and wheat in Asella, while the term 'other crops' denotes bean and enset in Hawassa and small grains (i.e. barley, tef and sorghum) and pulses (i.e. pea and faba bean) in Asella. The production of the main crop was compared to the production of other crops to assess whether there is competition, substitution or complementary use of resources between the different crops (cf. Figure 6.1D). Furthermore, the amount of land allocated to the main crop was compared with the amount of land allocated to other crops. A similar approach was followed to compare the amount of labour and animal power used for the main crop vis-à-vis the amount of labour and animal power used for other crops in the months in which land preparation/crop establishment, hand-weeding and harvesting were performed. We were not able to control for differences farm specialization or farmers' preferences towards specific crops, which may influence prioritization of resource allocation to specific crops over others.

Trade-off curves between farm production, crop area and labour use for wheat and other crops were estimated using boundary line analysis. These were fitted to the observations with the maximum wheat production and wheat area per unit production and area of other crops, respectively. For labour use, observations above the third quartile plus two times the interquartile range (i.e. Q3 + $2\times$IQR) were excluded prior to the analysis. Boundary lines were then fitted to the observations with the maximum labour used for wheat for each given labour used for other crops using the entire dataset and a subset of data below the 90th percentile of labour use for wheat and other crops. The models were estimated with ordinary least squares (lm function in R) and assuming a quadratic functional form of the type y = ax^2 + bx + c.

Yield gaps and farm performance

Farm performance was evaluated based on the principles of food production (N production), economic viability (gross margin) and labour drudgery (returns to labour). The contribution of livestock production to these indicators was not assessed, which means that our calculations may be slightly under-estimated.

The principle of food production was assessed with the indicator N production, which was calculated (in kg N ha^{-1}, Chapter 5) as follows:

$$\text{N production}_i = \frac{\sum_j \text{Yield}_{ji}\ (\text{kg DM ha}^{-1}) \times \text{N content}_j\ (\%) \times \text{Area}_{ji}\ (\text{ha})}{\sum_j \text{Area}_{ji}\ (\text{ha})} \tag{6.7}$$

where 'Yield$_{ji}$' and 'Area$_{ji}$' refer to the dry matter yield and area, respectively, of crop j in farm i as reported in the household survey. N concentration in harvested products ('N content$_j$' on a dry matter basis) was taken from Mellisse et al. (2017):

1.13% for maize, 0.75% for *kocho* (enset), 3.78% for bean and faba bean, 1.25% for barley and 1.76% for tef. The N contents of wheat (2.15%), sorghum (2.10%) and pea (4.15%) were taken from Nijhof (1987).

Economic viability and labour drudgery were quantified with the indicators gross margin and returns to labour, respectively. Gross margin (ETB ha^{-1}) was estimated as the difference between revenues on the one hand and fertiliser and labour costs per unit area on the other hand. Returns to labour (RTL, ETB person-day^{-1}) were calculated as gross margin per unit labour used. The calculation of these indicators was done as follows:

$$\text{Revenue}_i = \sum_j \text{Yield}_{ji} \text{ (kg DM ha}^{-1}\text{)} \times \text{Price}_j \text{ (ETB kg}^{-1}\text{)} \tag{6.8}$$

$$\begin{aligned}\text{Total Cost}_i = & \sum_j \text{Fertiliser used}_{ji} \text{ (kg ha}^{-1}\text{)} \times \text{Fertiliser price (ETB kg}^{-1}\text{)} + \\ & + \sum_j \text{Labour hired}_{ji} \text{ (person-day ha}^{-1}\text{)} \times \text{Wage (ETB person-day}^{-1}\text{)}\end{aligned} \tag{6.9}$$

$$\text{Gr. Marg.}_i = \text{Revenue}_i \text{ (ETB ha}^{-1}\text{)} - \text{Total Cost}_i \text{ (ETB ha}^{-1}\text{)} \tag{6.10}$$

$$\text{RTL}_i = \frac{\text{Revenue}_i \text{ (ETB ha}^{-1}\text{)}}{\sum_j \text{Labour used}_{ji} \text{ (person-day ha}^{-1}\text{)}} \tag{6.11}$$

where j denotes a specific crop and i a specific farm. The underlying crop data used to calculate gross margin and returns to labour were: 'Yield' which refers to farmer self-reported dry matter yield, 'Fertiliser used' which refers to the amount of urea or DAP applied, 'Labour hired' which corresponds to the amount of labour hired from land preparation to harvesting and 'Labour used' which is the total amount of family and hired labour used between land preparation and harvesting. Input-output prices were obtained from key informants in each site and included the market price of harvested crop product ('Price'), the unit price of urea and DAP ('Fertiliser price') and the unit price paid for hiring labour for different operations ('Labour wage'). Farm revenue assumes all crop production is sold, and thus corresponds to the 'best-case scenario' in terms of economic performance. The total cost of production is slightly underestimated due to lack of data on the amount of e.g. seed and herbicide used. Family labour costs were not considered in the calculations.

The relationship between the aforementioned indicators, and between those and crop yield, was studied to identify trade-offs or synergies between different principles and between closing the yield gap and optimising different indicators, respectively. These were studied with the same boundary line approach used to study resource allocation to different crops but applied to the observations with largest gross margin or returns to labour on the one hand per unit farm N production and crop yield on the other hand.

6.4 Results

6.4.1 Labour calendars for crop production

The seasonality of labour used for crop production differed strongly between the two sites (Figures 6.3 and D3; Supplementary Material). In short, the three main crops cultivated in Hawassa showed a complementary use of labour throughout the year while the crops cultivated in Asella competed for labour during the months of sowing, hand-weeding or harvesting. This was true for both labour and animal draught power, which was more important in Asella. Although religious holidays and fasting periods of the Orthodox church often conflict with busy periods of the cropping calendar in Ethiopia, this has minor effects on labour availability in the study sites as most of the population is either Protestant or Muslim. Further details on labour dynamics in terms of timing (Figure D4), amount (Figure D5) and quality (Table D1) for management operations are provided in Supplementary Material.

The complementary use of labour observed in Hawassa can be explained by the type of crops cultivated in this site (Figures 6.3A, 6.3C and 6.3E). Land preparation for maize cultivation is performed, both manually and with animal traction, between January and April, depending on rainfall and pest abundance. Most farmers apply basal mineral fertiliser, and some apply a top dressing up to two months after sowing. Animal traction is often used up to one month after sowing to break surface crusts between rows of maize and control weeds, a ridging practice locally known as *shilshalo*, and two hand-weedings are common between May and June. The maize growing season ends in October - November. Labour requirements for maize are minimum between July and September, when labour is used to grow a bean intercrop with maize (land preparation and sowing in July, hand-weeding in August and harvesting slightly before the maize crop in October). Bean can also be cultivated in synchrony with maize between April and July but this seems to be less preferred. As enset is a perennial crop with no critical harvest time farmers post-pone operations

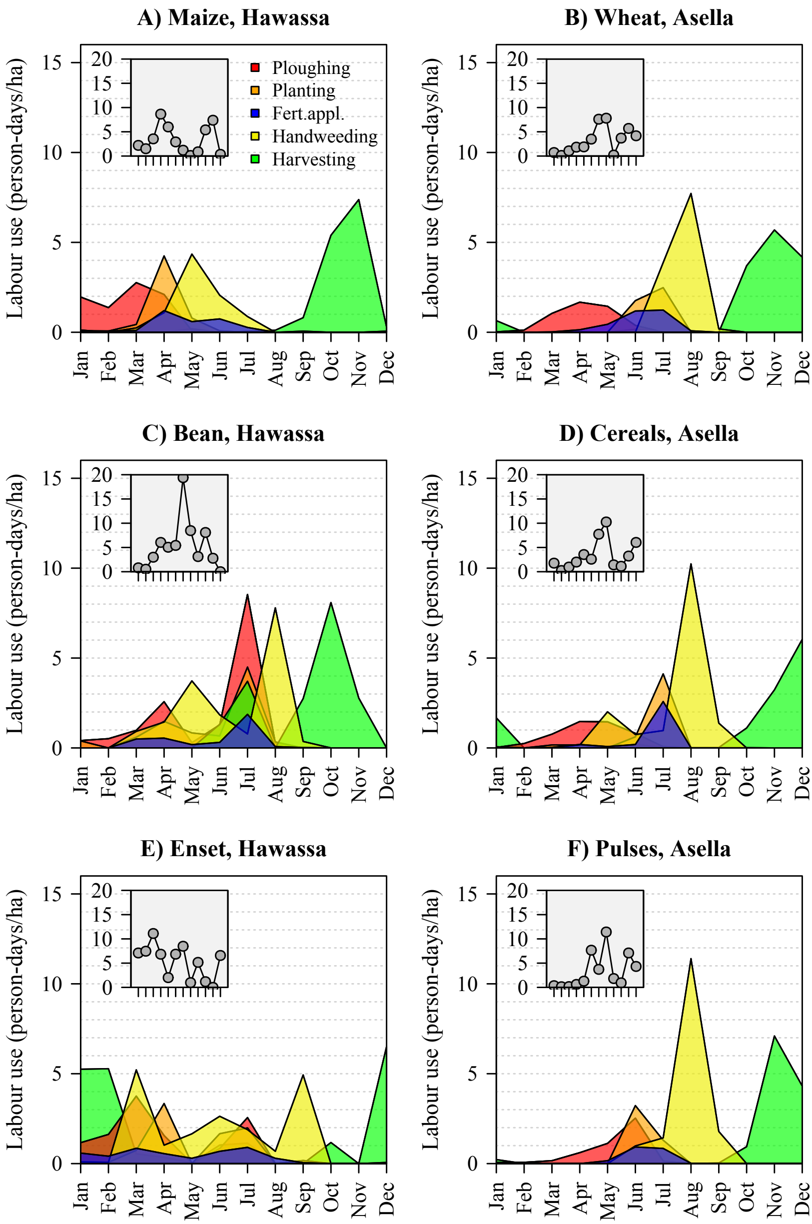

Figure 6.3: Labour calendar for crop-specific management operations across smallholder farms in Hawassa and Asella, Southern Ethiopia. The main crops in Hawassa are A) maize, C) bean and E) enset, and in Asella B) wheat, D) other cereals such as barley and tef and F) pulses such as pea and faba bean. The inset figures show the total labour used per month for each crop.

like hand-weeding and harvesting, and as it is cultivated in home gardens, compost is applied throughout the year. Thus, management operations for enset are mostly performed in the months with low labour use for maize and bean such as January - March, September and December.

Competition for labour was much more pronounced in Asella because labour peaks for small grains and pulses coincide with peaks of labour demand for wheat (Figures 6.3B, 6.3D and 6.3F). Similar sources of power are used to cultivate wheat and other small grains. Animal traction is the preferred source of power for land preparation, which is performed between February and June using a traditional plough locally known as a *maresha*. Up to three or four 'passes' are performed for cereals, which require a fine seed-bed for successful and uniform emergence. The sowing window for cereals is between June and July and most farmers use one basal application of mineral fertiliser. Weeding is laborious because cereals are broadcast. For this reason, pre-emergence herbicides are widely used and one hand-weeding is done up to 3 months after sowing (i.e. August). All crops are harvested using sickles between October and December.

6.4.2 Yield gaps and yield variability

Efficiency yield gap

In Hawassa, there was a positive linear and quadratic effect of N application on maize yields and yield response to N increased with increasing amounts of labour for the second hand-weeding (Model I without inefficiency effects; Table 6.2). In addition, labour used for the first hand-weeding had a positive effect on maize yields and maize yield responses to labour used for the first hand-weeding increased with increasing amounts of labour used for the second hand-weeding, when the number of management operations were included as inefficiency effects (Model II). The effects of N application on maize yields were 'diluted' when the timing of management operations were included as inefficiency effects (Model III) and sowing at later months resulted in significantly greater efficiency yield gaps. Finally, there was a significant effect of N (positive) and N $\times$ labour used for the first hand-weeding (negative) on maize yields when the effects of labour quality on the efficiency yield gap were considered (Model IV). In that model, a greater proportion of women labour used for sowing was associated with significantly greater efficiency yield gaps. Across all models, significantly smaller maize yields were reported for Wondo Genet than for Hawassa Zuria, but in Model IV the difference was only significant at 10% significance level.

Table 6.2: Parameter estimates of the stochastic frontier models estimated for maize in Hawassa (year 2012). The variables 'Weeding I' and 'Weeding II' refer to hand-weeding. Significance is indicated by the codes: '***' 0.1% '**' 1% '*' 5% '#' 10%.

	Maize I	Maize II	Maize III	Maize IV
Production frontier				
Intercept	**0.779** ***	**0.613** **	**0.873** ***	**0.507** ***
Nitrogen	**0.362** ***	**0.318** ***	-0.050	0.184 #
Nitrogen^2	**0.037** **	**0.031** *	0.269 #	**0.038** *
Weeding I	0.126	**0.162** *	**0.385** **	0.182
Weeding I^2	-0.004	-0.004	0.019	0.065
Weeding II	0.088	0.019	0.039	-0.138
Weeding II^2	0.011	-0.006	-0.018	-0.046
Nitrogen × Weeding I	-0.008	-0.006	**-0.215** ***	**-0.107** *
Nitrogen × Weeding II	**0.018** **	**0.016** *	**-0.070** **	0.001
Weeding I × Weeding II	0.016	**0.022** **	**0.044** *	0.004
Intercrop_Yes	0.070	0.065	0.048	-0.057
Location_Wondo Genet	**-0.467** *	**-0.413** *	**-0.527** **	-0.387 #
Inefficiency effects				
Ploughing operations		-1.873		
Sowing operations		2.409		
Fertiliser applications		0.755		
Weeding operations		-0.347		
Sowing time			**1.387** **	
1st fertiliser timing			0.835	
1st weed control			-0.846 #	
Ploughing hired				-2.097
Ploughing draught				-0.166 #
Sowing hired				2.780
Sowing female				**2.267** *
Sowing child				1.269
Weed control hired				-0.171
Weed control female				-2.282
Weed control child				-0.805
Model evaluation				
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	**1.290** ***	**2.784** *	**1.411** ***	**1.044** ***
$\gamma = \sigma_u^2 / \sigma^2$	**0.989** ***	**0.988** ***	**0.979** ***	**0.934** ***
TE scores	51.5	57.2	56.9	59.0
Sample size (n)	83	83	71	77

Table 6.3: Parameter estimates of the stochastic frontier models estimated for wheat in Asella (year 2012). The variables 'Weeding I' refers to hand-weeding and 'Weeding II' to herbicide application. Significance is indicated by the codes: '***' 0.1% '**' 1% '*' 5% '#' 10%. n.a. = not applicable.

	Wheat I	Wheat II	Wheat III	Wheat IV
Production frontier				
Intercept	**1.019** ***	**0.766** ***	**0.971** ***	**0.599** ***
Nitrogen	0.195	0.236 #	0.072	0.157
Nitrogen^2	0.423	0.363	0.611 #	**0.696** **
Weeding I	**0.238** **	**0.125** *	**0.208** **	0.083
Weeding I^2	**0.056** *	0.025 #	**0.048** **	0.018
Weeding II	**0.177** *	**0.178** *	**0.203** *	**0.190** *
Weeding II^2	0.030	0.040 #	0.044 #	0.041
Nitrogen × Weeding I	0.009	0.018	-0.003	0.016
Nitrogen × Weeding II	-0.011	0.003	-0.031	-0.130
Weeding I × Weeding II	0.005	-0.001	0.011	0.010
Location_Haro Bilalo	-0.021	-0.072	-0.107	-0.109
Inefficiency effects				
Ploughing operations		-10.513		
Sowing operations		n.a.		
Fertiliser applications		6.222		
Weeding operations		-4.436		
Herbicide applications		-3.454		
Sowing time			n.a.	
1st fertiliser timing			-1.215	
1st weed control			-1.141 #	
Ploughing hired				-6.915 #
Ploughing draught				-0.102
Sowing hired				**6.886** *
Sowing female				-15.475
Sowing child				-7.398
Weed control hired				**3.939** *
Weed control female				-10.159
Weed control child				-3.172
Model evaluation				
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	**0.676** ***	5.246	**0.980** ***	**0.932** *
$\gamma = \sigma_u^2 / \sigma^2$	**0.976** ***	**0.989** ***	**0.974** ***	**0.924** ***
TE scores	61.4	76.1	65.8	81.8
Sample size (n)	96	96	91	90

The drivers of the production frontier and efficiency yield gaps for wheat in Asella (Table 6.3) were considerably different from maize in Hawassa. In Asella, there was a statistically significant positive (linear) effect of herbicide use on wheat yields, which was consistent across all models. Moreover, significantly positive (linear and quadratic) effects of labour used for the first hand-weeding on wheat yields were observed in all but Model IV. Significant effects of N application on wheat yields (quadratic) were only observed when labour quality variables were included as inefficiency effects (Model IV). The proportion of hired labour used for ploughing (at 10% significance level only), sowing and weed control were positively associated with the efficiency yield gap, which indicate that wheat yield gaps in Asella are closely linked to labour quality for some management operations.

Based on Model I, the efficiency yield gap was on average 1.6 and 1.7 t ha^{-1} (or 49% and 38% of Y_{TEx}) for maize and wheat (Figures 6.4A and 6.4B), respectively. The variability of Y_{TEx} was large for maize in Hawassa (standard deviation ca. 1.5 t ha^{-1}) and for many farms the efficiency yield gap was well above 50% of Y_{TEx}, especially at low Ya levels. Other factors not included in our analysis may also explain the large magnitude of the efficiency yield gap as its average reduced by only 10% when more inefficiency effects were included (from 49% to 41%; Table 6.2). Conversely, the variability of Ya and Y_{TEx} of wheat in Asella was smaller compared to maize

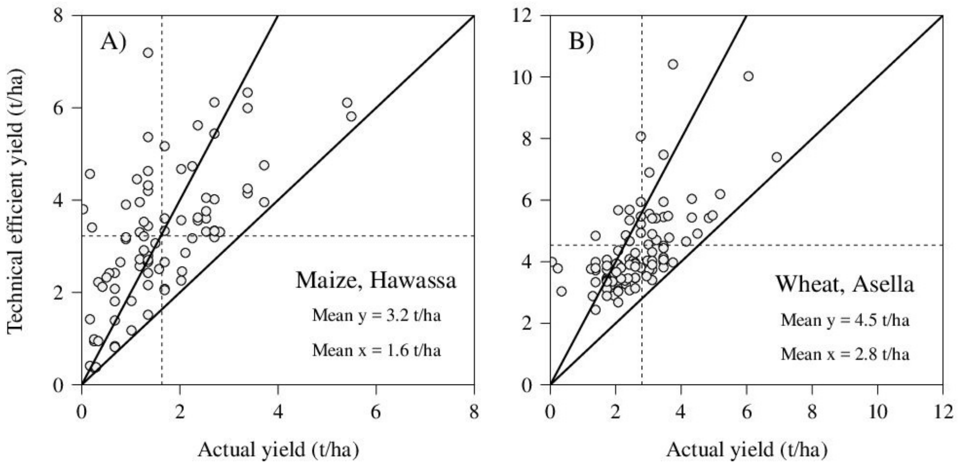

Figure 6.4: Efficiency yield gaps of A) maize and B) wheat across smallholder farms in Hawassa and Asella, Southern Ethiopia. Data refers to the year 2012; solid lines represent the 1:1 and 1:2 lines while the dashed lines represent the mean Ya (x-axis) and Y_{TEx} (y-axis, Model I for both crops).

in Hawassa, and only few farms exhibited efficiency yield gaps greater than 50% of Y_{TEx}. Based on the four models, the magnitude of the efficiency yield gap reduced from 38% in Model I to 18% in Model IV (Table 6.3), which means that the quality of hired labour explains ca. 20% of the efficiency yield gap of wheat.

Resource yield gap

The magnitude of Y_{HF}, Y_{AF} and Y_{LF} for maize in Hawassa were 4.0, 1.6 and 0.2 t ha^{-1}, respectively (Figures 6.5A - 6.5D). Y_{HF} were obtained with significantly more mineral N applied (111 kg N ha^{-1}), compared to Y_{AF} (59 kg N ha^{-1}) and Y_{LF} (41 kg N ha^{-1}), and there were no significant differences in mineral N applied between these last two groups (Figure 6.5A). Similar results were observed for mineral P applied (Figure 6.5B). No significant differences between Y_{HF}, Y_{AF} and Y_{LF} were observed regarding animal power used for ploughing and sowing (Figure 6.5C) and regarding the labour used for the first hand-weeding (Figure 6.5D). The fitted boundary lines plateaued at slightly lower input levels than the ones observed for Y_{HF}: ca. 50 kg N ha^{-1}, 25 kg P ha^{-1}, 7 animal-days ha^{-1} and 5 person-days ha^{-1} (Figures 6.5A - 6.5D).

For wheat in Asella, Y_{HF}, Y_{AF} and Y_{LF} were 4.9, 2.6 and 1.1 t ha^{-1}, respectively (Figures 6.5E - 6.5H). Differently from maize in Hawassa, there were no significant differences in mineral N applied to wheat between Y_{HF} (57 kg N ha^{-1}), Y_{AF} (43 kg N ha^{-1}) and Y_{LF} (36 kg N ha^{-1}; Figure 6.5E). Highest farmers' yields were attained with 42 kg P ha^{-1} (Figure 6.5F), which is significantly more than the P application rates used by average (30 kg P ha^{-1}) and lowest yielding farmers (28 kg P ha^{-1}). Highest farmers' yields were also associated with more animal power and labour use. For instance, Y_{HF} was reached with 49 animal-days ha^{-1} for ploughing and sowing, and with 2 person-days ha^{-1} for herbicide application (Figures 6.5G and 6.5H). This is significantly more than the 20 - 25 animal-days ha^{-1} for ploughing and sowing, and than the 0.6 - 0.7 person-days ha^{-1} for herbicide application associated with Y_{AF} and Y_{LF}. Finally, wheat yields plateaued at ca. 20 kg N ha^{-1}, 40 kg P ha^{-1}, 20 animal-days ha^{-1} for ploughing and sowing, and 0.5 person-days ha $^{-1}$ for herbicide application (Figures 6.5E - 6.5H).

Technology yield gap

The largest share of maize and wheat yield gaps was attributed to the technology yield gap. For maize this was 3.2 t ha^{-1}, or 45% of Yw (Figures 6.5A - 6.5D), and for wheat 5.2 t ha^{-1}, which corresponds to 52% of Yw (Figures 6.5E - 6.5H). The simulated Yw for maize was on average 7.0 t ha^{-1} but highly variable (standard deviation of ± 3.7 t ha^{-1}), which indicates variable amount and/or

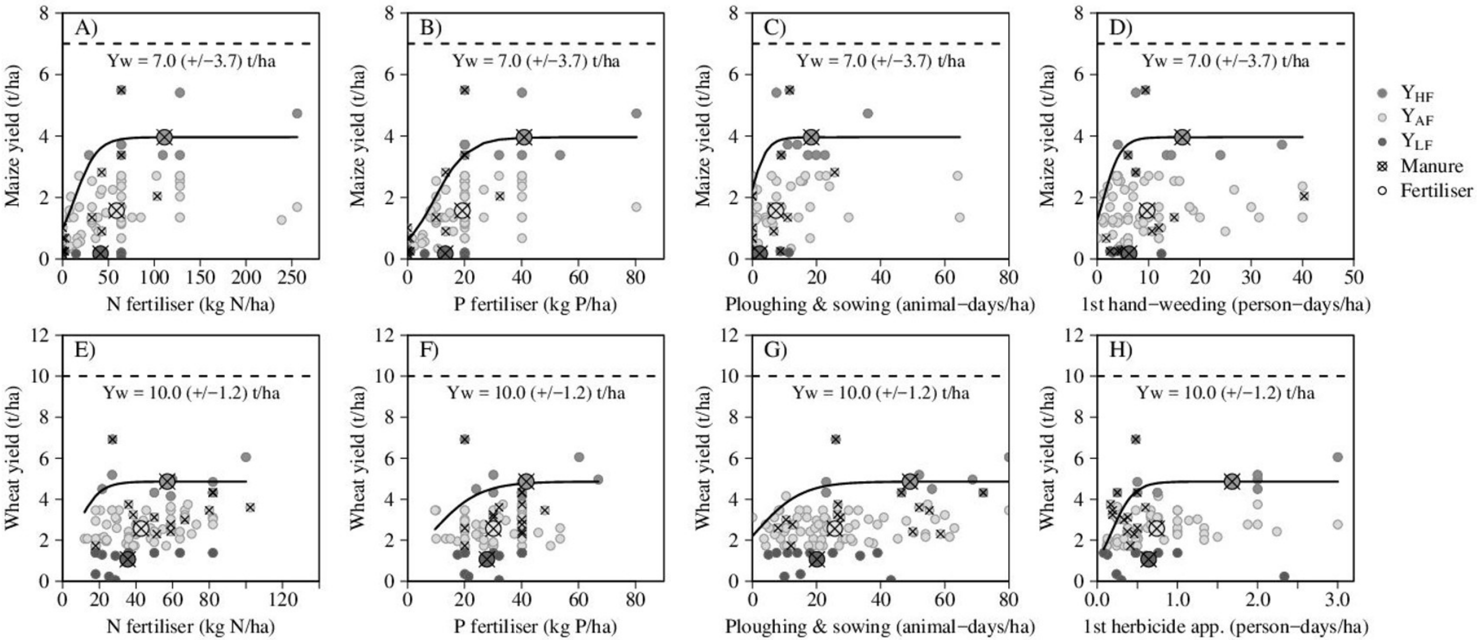

Figure 6.5: Resource (N, P and labour) and technology yield gaps for maize in Hawassa (A - D) and wheat in Asella (E - H), Southern Ethiopia. Each dot refers to a single farm in the year 2012. Mean yield and input use for Y_{HF}, Y_{AF} and Y_{LF} is represented by the largest dots with a cross. Solid lines are boundary lines fitted to the observations using mineral fertiliser only. Horizontal dashed lines show the crop-specific mean (standard deviation) Yw estimates from GYGA (1998 - 2012, www.yieldgap.org).

distribution of rainfall across different years. For wheat, the simulated Yw was 10 t ha^{-1} and its inter-annual variability was low (standard deviation of $\pm$ 1.2 t ha^{-1}), suggesting a more uniform rainfall distribution. A large difference between Yp and Yw was observed for maize (3.5 t ha^{-1}) but not for wheat (0.9 t ha^{-1}, data not shown).

6.4.3 Farm power and capital availability

In Hawassa, oxen ownership was a proxy of capital availability, but not of farm power, as farms with more pairs of oxen had significantly more assets (and/or assets of greater value) than farms with less pairs of oxen (Figure 6.6A). However, there was no difference in the cultivated area between the different groups (Figure 6.6B). Conversely, oxen ownership in Asella was a proxy for both capital availability and farm power as farms with more pairs of oxen were found to have significantly more, and/or more valuable, assets (Figure 6.6C), and significantly larger cultivated area than farms with less pairs of oxen (Figure 6.6D).

Oxen ownership in Hawassa was positively associated with greater maize yields and N application rates for maize, but not with more animal power used for ploughing and sowing or more labour used for hand-weeding. Farms owning two or more pairs of oxen produced ca. 3.2 t maize ha^{-1}, which is significantly greater than the 2.7 and 1.7 t ha^{-1} produced by respectively farms with one pair of oxen or no oxen (Figure 6.6G). N application rates were nearly double in farms with two or more pairs of oxen compared to farms with no oxen (120 vs. 53 kg N ha^{-1}; Figure 6.6H). No significant difference was observed in animal power used for ploughing and sowing (2 - 12 animal-days ha^{-1}; Figure 6.6M) or total labour used for the first hand-weeding (5 - 15 person-days ha^{-1}; Figure 6.6N) between groups differing in oxen ownership.

Increased oxen ownership and availability of animal traction in Asella resulted in more cultivated area, but not higher wheat yields. This is demonstrated by the significant differences in cultivated area between groups with one pair or two or more pairs of oxen (Figure 6.6E) and by no significant difference in wheat yields between the two groups: 2.9 vs. 3.5 t ha^{-1} for farms owning respectively one or two or more pair of oxen (Figure 6.6J). Similarly, no significant difference in N applied to wheat was observed between the two groups (ca. 50 kg N ha^{-1}, Figure 6.6K). An increase in oxen ownership did not translate into a reduction of animal power used for ploughing and sowing per unit area (Figure 6.6P), while farms with two or more pairs of oxen tended to use slightly less labour for the first hand-weeding (Figure 6.6Q) and slightly more labour for herbicide application than farms with one pair of oxen (data not shown), but the differences were not significant.

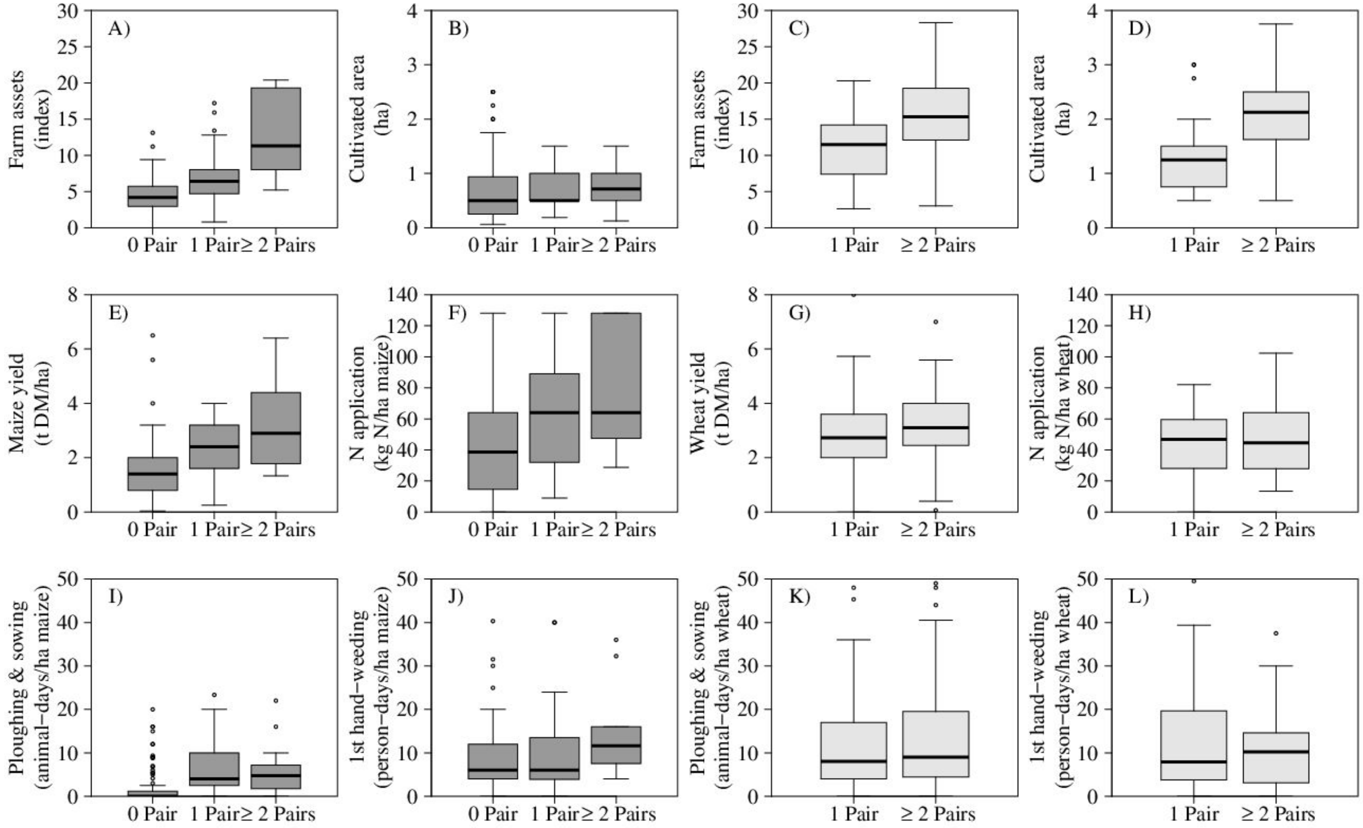

Figure 6.6: Relations between number of pair of oxen and farm assets, cultivated area, crop yield, mineral N applied and labour use for smallholder farms in Hawassa (dark grey) and Asella (light grey), Southern Ethiopia. Crop-specific effects are shown for maize in Hawassa and for wheat in Asella. Farm assets in A) and D) do not include the capital value of livestock. Significant differences between farm types are provided in Table D3 (Supplementary Material).

6.4.4 Resource allocation at farm level

We report the results of resource allocation at farm level for Asella because there is little competition for labour in Hawassa (Figure 6.3). Moreover, there seems to be little competition for capital in Hawassa as well because mineral N fertiliser is only relevant for maize (Figure 6.6F), but not for bean, an N_2-fixing crop, or enset, to which fresh manure is applied. Detailed results of farm production and resource allocation in Hawassa are thus provided in Supplementary Material (Figure D9).

Farms in Asella produced an average of 2.8 t of wheat, 1.4 t of other cereals and 0.5 t of pulses (Figures 6.7A and 6.7B). The lowest yielding wheat farms exhibited low production for all crops, while there was no clear association between the production of wheat and other crops in the highest yielding wheat farms. The concave shape of the boundary line indicates complementary use of production resources between wheat on the one hand and other cereals and pulses on the other hand. However, the relationship between wheat and pulses needs to be interpreted with caution due to the lack of observations with high pulse production and low wheat production. It is worth noting that only few farms produced more than 3.0 t of cereals or more than 1.5 t of pulses, and the ones that did so produced between 3.0 t and 3.5 t of wheat.

The area cultivated with wheat and other crops was on average 1.1 and 0.9 ha, respectively (Figure 6.7C). Few farms cultivated more than 3 ha of land and there was a nearly perfect substitution between the area cultivated with wheat and the area cultivated with other crops. The highest and lowest yielding wheat farms cultivated less than 1 ha of this crop, with the exception of two lowest yielding wheat farms which cultivated more than 2 ha of wheat.

During the months of land preparation and crop establishment, there was a nearly perfect substitution between animal power used for wheat and used for other crops for farms using an input of animal power for these operations above the 90^{th} percentile (Figure 6.7D). However, a complementary use of animal power was observed between wheat and other crops for farms using an input of animal power lower than the 90^{th} percentile. Most highest yielding wheat farms, and few lowest yielding wheat farms, used inputs of animal power close to the fitted boundary lines. During the months of hand-weeding, the boundary lines indicated nearly perfect substitution in the use of this resource for wheat and other crops across farms using an input of labour lower than the 90^{th} percentile (Figure 6.7E). However, the input of labour used for hand-weeding other crops is nearly double that used for wheat, to which herbicides are generally applied. During the months of harvesting, there was also a

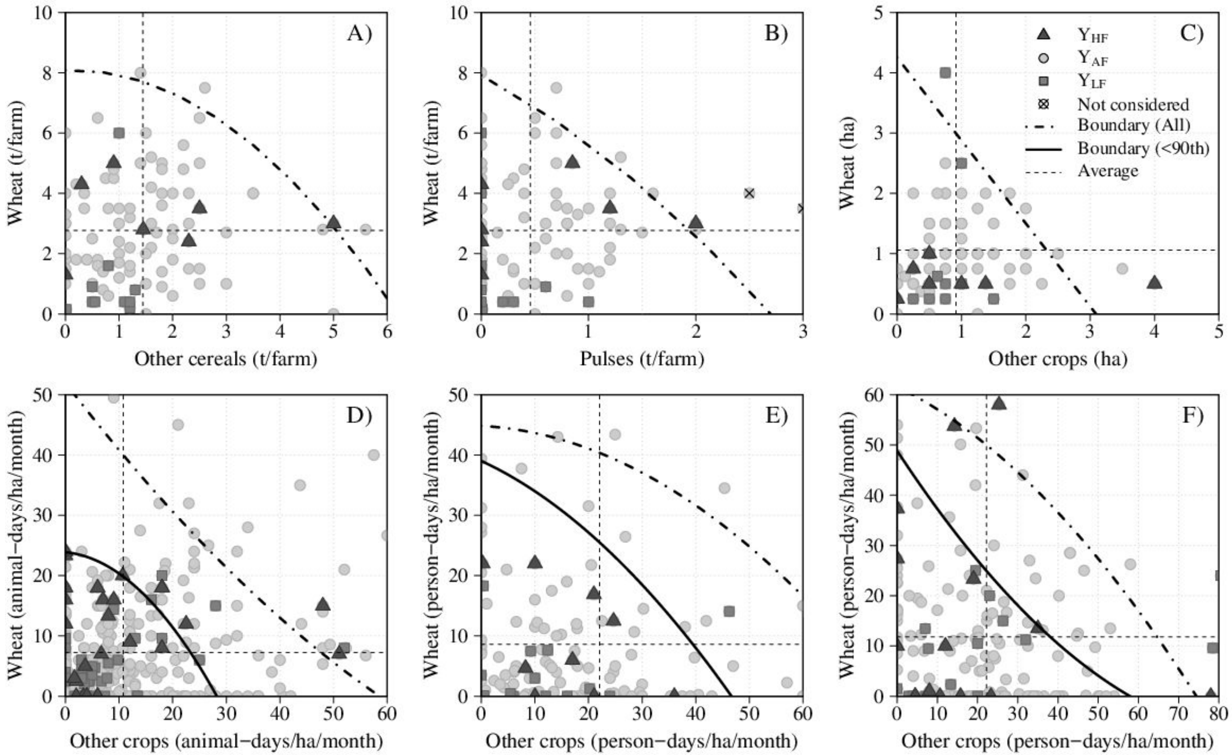

Figure 6.7: Land and labour productivity of wheat as a function of land and labour productivity of other cereals and/or pulses in Asella, Southern Ethiopia: A) relationship between production of wheat and other cereals and B) pulses, C) area share of wheat and other crops, D) animal-days used in the months for ploughing and sowing, E) person-days used in the months of hand-weeding and F) in the months of harvesting. Data for Hawassa are provided in Supplementary Material (Figure D9).

nearly perfect substitution between labour used for wheat and used for other crops for the entire sample and for farms using an input of labour lower than the 90th percentile (Figure 6.7F).

6.4.5 Crop and farm performance indicators

Farm N production in Hawassa was ca. 15 kg N ha^{-1} on average, which is much less than the average of 41 kg N ha^{-1} observed in Asella (Figure 6.8). This indicator was positively associated with gross margin per ha at farm level in both sites (Figure 6.8A). The average gross margin observed in Hawassa was ca. 4 kETB ha^{-1} and ca. 16 kETB ha^{-1} in Asella where farmers grow crops of higher value such as pulses, teff, wheat and barley. A quadratic relationship was observed between farm N production and returns to labour in both sites meaning that returns to labour increase up to 23 and 48 kg N ha^{-1} in respectively Hawassa and Asella, after which returns to labour decline as N production increases (Figure 6.8A). Returns to labour were ca. 103 ETB person-day^{-1} in Hawassa and ca. 324 ETB person-day^{-1} in Asella, which again points to better economic farm performance in Asella.

Gross margin per ha at crop level increased linearly with maize and wheat yields (Figure 6.8C) indicating little trade-off between closing the yield gap and maximising the gross margin within the yield ranges reported. The difference in gross margin observed between crops reflected the greater profitability of wheat in Asella as compared to maize in Hawassa. Similarly to the farm level analysis, there was a quadratic relationship between crop yield and returns to labour at crop level (Figure 6.8D) with optimal returns to labour observed at 3.6 t maize ha^{-1} in Hawassa and 3.2 t wheat per ha^{-1} in Asella.

6.5 Discussion

6.5.1 Scope for intensification in Southern Ethiopia

The yield gap was on average 5.4 t ha^{-1} (or 77% of Yw) for maize in Hawassa and 7.4 t ha^{-1} (or 74% of Yw) for wheat in Asella. These confirm the results of other studies, which reported large yield gaps for these crops in Ethiopia (Hoffmann et al., 2017; van Ittersum et al., 2016). For maize, the efficiency yield gap was 23% of Yw, the resource yield gap was 9% of Yw and the technology yield gap was 45% of Yw. For wheat, the efficiency and resource yield gaps were respectively 17 and 5% of Yw and the technology yield gap was ca. 52% of Yw.

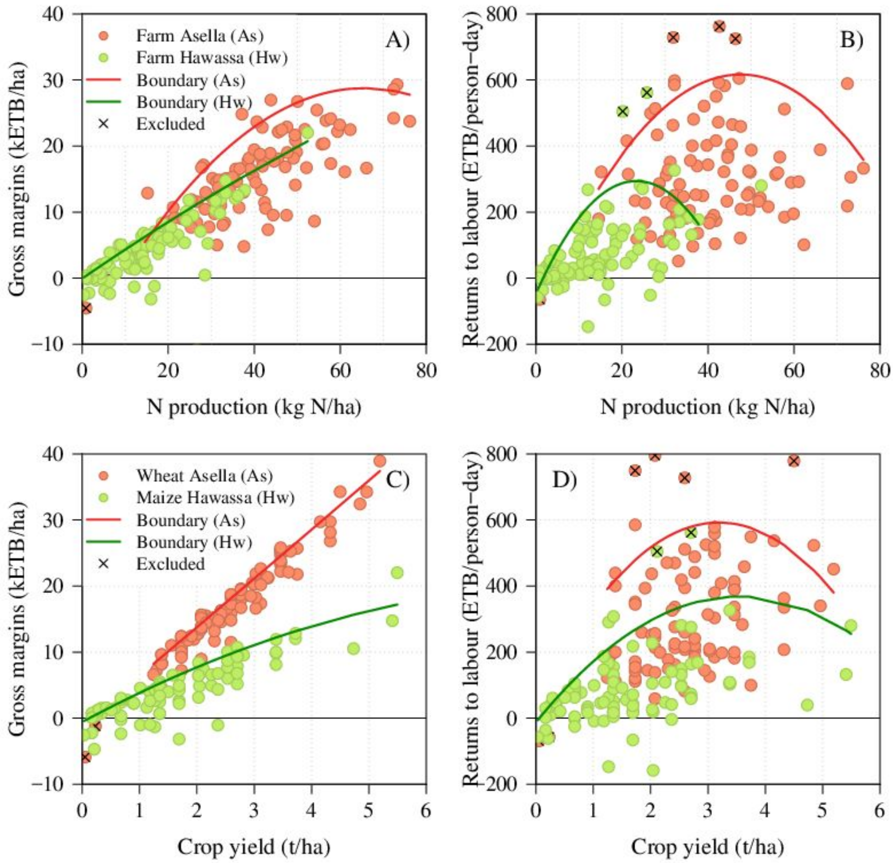

Figure 6.8: Relationship between N production and gross margin (A) and returns to labour at farm level (B), and between crop yield and gross margin (C) and returns to labour at crop level (D). Background data for each crop are provided in Figure D10 (Supplementary Material).

Stochastic frontier analysis has been increasingly applied in agronomic studies as a benchmarking technique for yield gap analysis (e.g. Chapter 3; Henderson et al., 2016; Carberry et al., 2013; Hoang, 2013). Early sowing and, to a less extent, a lower proportion of women labour involved in sowing contributed to smaller efficiency yield gaps of maize (Table 6.2). The sowing window for maize ranged from February to June (Figure 6.3) due partly to differences in agro-ecology between farms in terms of elevation. Moreover, farmers synchronized sowing of maize in e.g. Hawassa Zuria to reduce the incidence of stem borers (Kebede et al., 2015), but not in e.g. Wondo Genet where farmers focused more on cash crops (e.g. coffee and khat) instead of maize (Mellisse et al., 2017). Although maize sowing is mostly done by men (Figure D6; Supplementary Material), we have no specific explanation for the observed effects of women labour during sowing. For wheat in Asella, the efficiency yield gap increased with increasing amount of hired labour used for sowing and weed control (Table 6.3). Challenges in handling hired labour have been reported in focus group discussions conducted in this site, which support these results (van Eerdewijk and Danielsen, 2015). Moreover, it is well known that sowing and weeding are very labour intensive and tedious operations, and their impact on wheat yield in Ethiopia has been well documented (e.g. Nyssen et al., 2011; Taa et al., 2004; Tanner et al., 1993).

The resource yield gap was small for both crops due to the small difference in input use between Y_{HF} and Y_{TEx} (Figure 6.5). The amount of N applied was the key driver of maize yields (Table 6.2) while wheat yields were mostly associated with labour use for herbicide application and hand-weeding (Table 6.3). Comparisons of input and labour used by highest, average and lowest yielding farms also confirm these results (Figure 6.5). For instance, highest yielding wheat farms in Asella used nearly twice as much labour for ploughing and sowing, and for weed control, as average and lowest yielding farms. These differences are consistent with the competition for labour during the wheat growing season observed in Asella (Figure 6.3). Although not tested in the stochastic frontier, yield responses to P were also observed for both crops (Figures 6.5B and 6.5F).

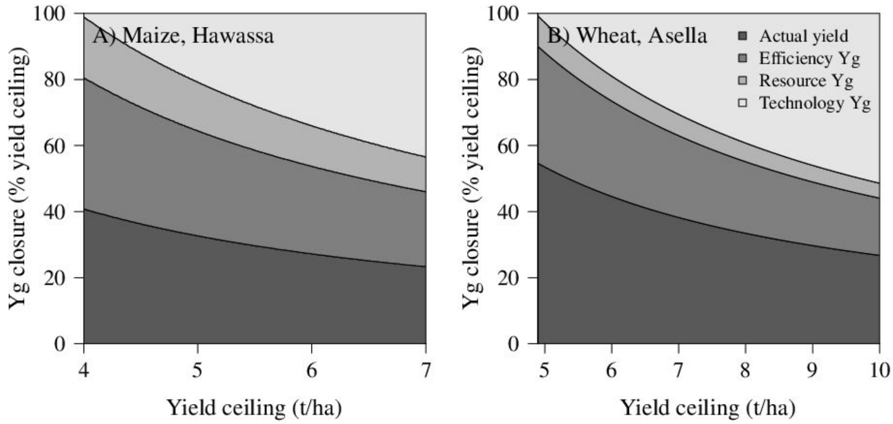

Figure 6.9: Average magnitude of the efficiency, resource and technology yield gaps (in relative terms) as a function of increasing yield ceiling for A) maize in Hawassa and B) wheat in Asella. The minimum yield ceiling presented corresponds to the crop-specific Y_{HF} and the maximum yield ceiling to the model-based simulated Yw.

Yield gaps in both sites were largely attributed to the technology yield gap (Figure 6.5), which is explained by differences in management practices, and possibly in crop varieties, used to achieve highest farmers' yields and assumed to simulate Yw. The differences in crop management are related to the lack of adoption of precision agriculture technologies as well as to resource yield gaps of specific inputs (Chapter 3). For example, broadcasting is the preferred method of fertiliser application (and sowing) for wheat in Asella while farmers in Hawassa use banding rather than more localized placement methods (Sime and Aune, 2014). Labour may be also important if more labour, or labour of better quality, is required by precision agriculture practices. The most obvious resource yield gaps refer to sub-optimal sowing densities used by highest-yielding farmers (particularly for maize), possible limitations of K (even though soils in the study sites have high values of exchangeable K) and lack of plant protection agents to control biotic factors (used by less than 10% of the farmers sampled). It is unclear whether the N application rates used and the varieties adopted by highest-yielding farmers are able to attain Yw and further research is required to clarify these aspects.

Uncertainties in the yield ceiling have a strong impact on the relative importance of the different yield gaps (Figure 6.9). For instance, the magnitude of the efficiency yield gap is nearly double if Y_{HF} is considered to be the yield ceiling instead of Yw. As expected, the relative importance of the technology yield gap increases with an increasing yield ceiling while the resource yield gap is barely affected. It is important to acknowledge this variation particularly for wheat in Asella where the maximum yield, ca. 5.5 t ha^{-1} (Habte et al., 2014), obtained in on-farm trials conducted in the area during the same year the household survey was conducted is much closer to 4.9 t ha^{-1} (Y_{HF}) than to 10.0 t ha^{-1} (Yw). To our best knowledge, the latter was only observed in advanced breeding lines of wheat (Wandera et al., 2015). The yield difference between on-farm trials (TAMASA, unpublished data) and model simulations does not seem to be that large in Hawassa: 6.0 and 7.0 t ha^{-1}, respectively.

6.5.2 Opportunities for labour-saving technologies

Oxen ownership was a proxy of farm power and/or capital availability (Baudron et al., 2015; Aune et al., 2001) and there were clear differences between the two sites regarding this metric (Figure 6.6). In Hawassa, oxen ownership was associated with intensification of crop management (Figure 6.6H) while, in Asella, it was associated with extensification as more oxen resulted in greater cultivated land (Figure 6.6E) rather than greater labour use efficiency (i.e. a decrease in the number of days worked per unit area, Figures 6.6P and 6.6Q). This has been observed in other farming systems across sub-Saharan Africa as well (Ollenburger et al., 2016; Leonardo et al., 2015; Baudron et al., 2012). This contrast between the two sites can be explained by differences in land:labour ratio: population density in Hawassa is more than 600 persons km^2 while in Asella it is ca. 200 persons km^2 (based on the 2007 census).

These 'trajectories' have important implications for agricultural development and for the interventions required to close yield gaps. Capital and land, not labour, constraints are at stake in Hawassa (Figure 6.6A), where gross margin per ha at crop and farm level are also small compared to the national poverty line (Figures 6.8A and 6.8C). Lack of capital and small farm sizes most likely triggered many farmers to expand the area of high value crops (Hazell et al., 2010), such as *khat* in this site (Mellisse et al., 2017). Conversely, labour constraints may still be problematic for some farms in Asella (Figures 6.3), particularly during the months of ploughing and sowing, hand-weeding and harvesting wheat (Figure 6.7). Despite the linear increase of gross margin per unit area with increasing wheat yield and farm N production, the gross margin per labour unit decreased after an optimal level of crop yield and farm N production (Figures 6.8B and 6.8D). This points to the importance of increasing labour productivity.

Labour-saving technologies can reduce the number of management operations (e.g. conservation agriculture; Nyssen et al., 2011) or the actual amount of labour required per management operation (e.g. machinery and herbicides). These can contribute to increase labour use efficiency (Figure 6.5) by increasing crop yield through improved timeliness and precision (e.g. sowing depth and row planting), and/or reducing labour demand for critical management operations. For example, Nyssen et al. (2011) proposed a 'bed-and-furrow' system, involving the traditional ard plough ($maresha$) and permanent seedbeds, that requires less power and shortens the number of 'passes' required for land preparation. This would be a suitable option to decrease labour demand for land preparation and sowing in Asella (Figure 6.7D). Herbicides are already widely used for wheat in this site (Figure 6.5H) which reduced the labour requirements for hand-weeding to less than 10 person-days ha^{-1} (Figures 6.7E and 6.3B). Two-wheel tractors may also be suitable for smallholder farms in Asella, where land is fragmented, to ensure labour is not limiting during land preparation (Baudron et al., 2015).

'Suitable, reliable and affordable mechanization' has been recognized as a key strategy to accelerate agricultural growth in Africa (Malabo Declaration of the African Union). At the same time, farm power is stagnating or even declining in some regions across sub-Saharan Africa (Baudron et al., 2015). This may be further exacerbated in the future by e.g. out-migration from rural to urban areas, particularly in Southern Ethiopia (Bezu and Holden, 2014), and the preference of household members for more regular and less labour intensive sources of income, such as off-farm employment (Frelat et al., 2016; Haggblade et al., 2007). Moreover, other agricultural (e.g. herding) and household activities (e.g. domestic chores) still require a substantial amount of labour (Figure D11; Supplementary Material) which further reduces labour availability for agricultural production. Reducing labour drudgery at farm level is especially important for men who are in charge of most crop management activities in the two sites (van Eerdewijk and Danielsen, 2015; Figure D11; Supplementary Material).

6.6 Conclusions

There were sharp differences between the two sites concerning the labour calendars of the main crops. In Hawassa, there was a complementary use of labour throughout the year: labour peaks for bean occurred when labour demand for maize was low while labour peaks for enset were observed when labour demand for both bean and maize was low. In Asella, the labour peaks of the main crops (wheat, barley, tef, peas and faba bean) coincided during the months of sowing and hand-weeding as well as harvesting. The different crops also competed for animal draught power in the months of land preparation. This was confirmed by the pattern of labour allocation to different crops at farm level.

Yield gaps of maize and wheat were as large as 77 and 74% of Yw, respectively. Actual yields were ca. 1.6 and 2.6 t ha^{-1} for maize and wheat, respectively. For both crops most of the yield gap was attributed to the technology yield gap, which had a magnitude of ca. 45% of Yw for maize and ca. 52% of Yw for wheat. This reflects differences in crop management observed in farmers' fields compared to the assumptions made to simulate Yw, including sub-optimal sowing practices and/or yield reduction due to pest and diseases. For maize, the efficiency yield gap explained ca. 23% of Yw and it increased with late sowing and with increased proportion of women labour used for this operation. For wheat, the efficiency yield gap was slightly smaller (ca. 17% of Yw) and it increased with increasing amounts of hired labour for sowing and weed control. The resource yield gap was less than 10% of Yw for both crops and explained mostly by N application levels for maize and herbicide use for wheat. Effects of labour used for hand-weeding were also observed for both crops.

Oxen ownership was associated with intensification of maize production in Hawassa and extensification of wheat production in Asella. For instance, farms with higher oxen ownership used higher rates of mineral N fertiliser in Hawassa and cultivated more land in Asella while no difference in labour use efficiency between farms differing in oxen ownership was observed in either site. The economic performance of farms in Hawassa was lower than in Asella. In both sites, there was little trade-off between gross margin per ha and crop yield while gross margin per unit labour declined after an optimal level of crop yield. This means that even if more labour could be used it would not be worthwhile to do so from an economic perspective. Labour and farm power were not major determinants of maize yield gaps in Hawassa but they appeared to explain wheat yield gaps in Asella. Further efforts are required to understand the magnitude and nature of labour constraints, and their impact on crop productivity, in other farming systems across sub-Saharan Africa and to target technologies that can increase labour productivity.

6.7 Acknowledgements

We thank Wageningen University for funding through the project 'Mapping for Sustainable Intensification' and the Australian Centre for International Agricultural Research (ACIAR) for funding the project 'Farm Power and Conservation Agriculture for Sustainable Intensification' (FACASI, www.facasi.act-africa.org), CRP MAIZE (www.maize.org) and CRP WHEAT (www.wheat.org). Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of Wageningen University, ACIAR, CRP MAIZE and CRP WHEAT. We thank the Ethiopian Institute of Agricultural Research (EIAR) for support with the on-farm experiments, Dr. Kindie Tesfaye for information on the yield ceilings, Mr. Ephrem Tadesse for compiling the secondary data used to compute the farm performance indicators and Mr. Tesfaye Sida, Mr. Beyene Mellisse and Mrs. Yodit Kebede for stimulating discussions.

CHAPTER 7

Whose gap counts? The role of yield gap analysis within a development-oriented agronomy

This chapter is under review as:
Silva, J.V.; Ramisch, J.J. 2017. Whose gap counts? The role of yield gap analysis within a development-oriented agronomy. *Experimental Agriculture*.

J.V. Silva contributed to the development of the framework, data analysis for the case study in rice-based farming systems in the Philippines and writing of the manuscript.

Abstract

Yield gaps have become a useful tool for guiding development-related agronomy, especially in the global South. While critics have challenged some aspects of the yield gap methodology, and the relevance of food security advocacy based on yield gaps, very few studies question the actual relevance, application and scalability of yield gaps for smallholder farmers (and researchers) in the tropics. We assess these limitations using two contrasting case studies: maize-based farming systems in Western Kenya and rice-based farming systems in Central Luzon, Philippines. From these two cases, we propose improvements in the use of yield gaps that would acknowledge both the riskiness of crop improvement options and the role that yield increases might play within local livelihoods.

Participatory research conducted in Western Kenya calls into question the actual use and up-scaling of yield measurements from on-station agronomic trials to derive estimates of actual and water-limited yields in the region. Looking at maize yield gaps as cumulative probabilities demonstrates the challenges of assessing the real magnitude of yield gaps in farmers' fields and of deciding whose yield gaps counts for agricultural development in Kenya. In the case of rice-based farming systems, we use a historical dataset (1966 - 2012) to assess changes in rice yields, labour productivity, gross margin and rice self-sufficiency in Central Luzon (Philippines). While large rice yield gaps persist here, there appear to be few incentives to close that gap once we consider the position of crop production within local livelihoods. In this context, economic returns to labour for farm work were marginal: labour productivity increased over time in both wet and dry seasons, but gross margins decreased in the wet season while no trend was observed for the dry season. Since most households were rice self-sufficient and further increases in crop production would offer minimal returns while relying increasingly on hired labour, we question who should close which yield gap.

7.1 Introduction

Crop yield gaps feature prominently in the literature not only as a framework to disentangle effects of growth-defining, -limiting and -reducing factors to actual yields (van Ittersum and Rabbinge, 1997) but also to make claims about improvements of rural livelihoods (e.g. Dzanku et al., 2015). However, very few studies question the actual relevance, application and scalability of the concept for smallholder agriculture in the context of development-oriented agronomy. Yield gaps remain a problematic concept in ways which go beyond the methodological issues raised by van Ittersum et al. (2013) and by Sumberg (2012) and which we explore in this paper.

Beyond a failure to acknowledge the reasons why smallholder farmers under-produce relative to potential yields, yield gap calculations vary widely in the reliability of the input data used to assess yield ceilings and actual yields (Grassini et al., 2015). On the one hand, the supposedly ideal, researcher-managed conditions used to calibrate crop models and to estimate yield ceilings are based on the selective erasure of social and logistical factors that determine crop performance, such as planting and harvesting dates, or crop protection measures. Some of these factors could actually reduce yields (especially when inputs and labour supply are not available at required moments), but most lead to unrealistic inflation of yield ceilings (especially when extrapolating from very small and potentially unrepresentative sample plots). On the other hand, assessments of farmers' own production are subject to similar errors, whether taken from small sample plots or national statistics. In both cases, data are especially scarce and unreliable under more marginal conditions, which is exactly where smallholders have the most need for improvement or good technology (Grace et al., 2014). These challenges undermine the ease with which either set of data can be scaled up as a 'technical' and socially-neutral artefact from small plots to represent the performance of larger agro-ecological areas, or with which crop model simulations (such as the Global Yield Gap Atlas, GYGA) can be downscaled to specific locations.

The diversity which characterizes farming systems in the tropics (e.g. Stuart et al., 2016; Giller et al., 2011) further challenges the relevance of yield gaps for improving rural livelihoods. Smallholder farmers across the globe face multiple biophysical (Tittonell and Giller, 2013; Lansigan et al., 2000) and socio-economic challenges (Ellis, 1993) and operate with scarce resources in terms of land, labour and capital, which hinder closure of yield gaps in their fields. Input-output markets and prices deserve particular attention since they determine the profitability of farming and hence both its importance for rural households and its potential for boosting rural economies. Nonetheless, staple foods must remain affordable so that they can be accessed by urban and rural poor who are net purchasers of food. Unfavourable market conditions for agricultural commodities lead to decreasing marginal returns to labour of farming activities and 'marginalization' of the agricultural sector. These conditions are more likely to force smallholder farmers to seek opportunities off-farm (e.g. Frelat et al., 2016; Takahashi and Otsuka, 2009) than to intensify their production systems in order to sustain their livelihoods. Reversing this trend is not impossible but may require the implementation of protective and strategic policies by national and regional authorities (Studwell, 2013).

The increasing misapplication and over-extension of the 'technical' yield gap concept formalized by van Ittersum and Rabbinge (1997) to justify investments in research and development as well as policy interventions in developing countries (van Oort et al., 2016; Sumberg, 2012) creates the need for a thorough analysis of the incentives available for smallholder farmers to close yield gaps. The objective of this manuscript is to understand and discuss the relevance and usefulness of yield gaps to make claims about rural development as part of a development-oriented agronomy. This manuscript builds upon two case studies analysing empirically 1) the use and up-scaling of yield measurements from on-station agronomic trials to derive estimates of actual and water-limited yields in Kenya and 2) the changes in rice yields, labour productivity, gross margin and rice self-sufficiency in Central Luzon (Philippines) during the past half-century. A map with the location of the villages where households were surveyed in both case studies is provided in Figure 7.1.

7.2 Yield gaps at local level

Yield gap analysis at local level should go beyond the traditional mean difference between a yield ceiling (e.g. potential yield, Yp, or water-limited yield, Yw) and average farmers' yields (Ya) in two different aspects. Firstly, it should acknowledge the existence of variability in yield ceilings and Ya by analysing these yield levels in terms of yield distributions instead of single point estimates. Secondly, it should be contextualized as one dimension among many within wider livelihood preoccupations that assign greater or lesser importance to narrowing any of the gaps. Differently from Snyder et al. (2016), we focused on methodology development to widen the current yield gap approach and raise awareness of the effective contribution of farming for rural households. Further, we narrow down the approach of van Oort et al. (2016) and illustrate how yield gaps can be used for R&D prioritization at farm(ing) system level.

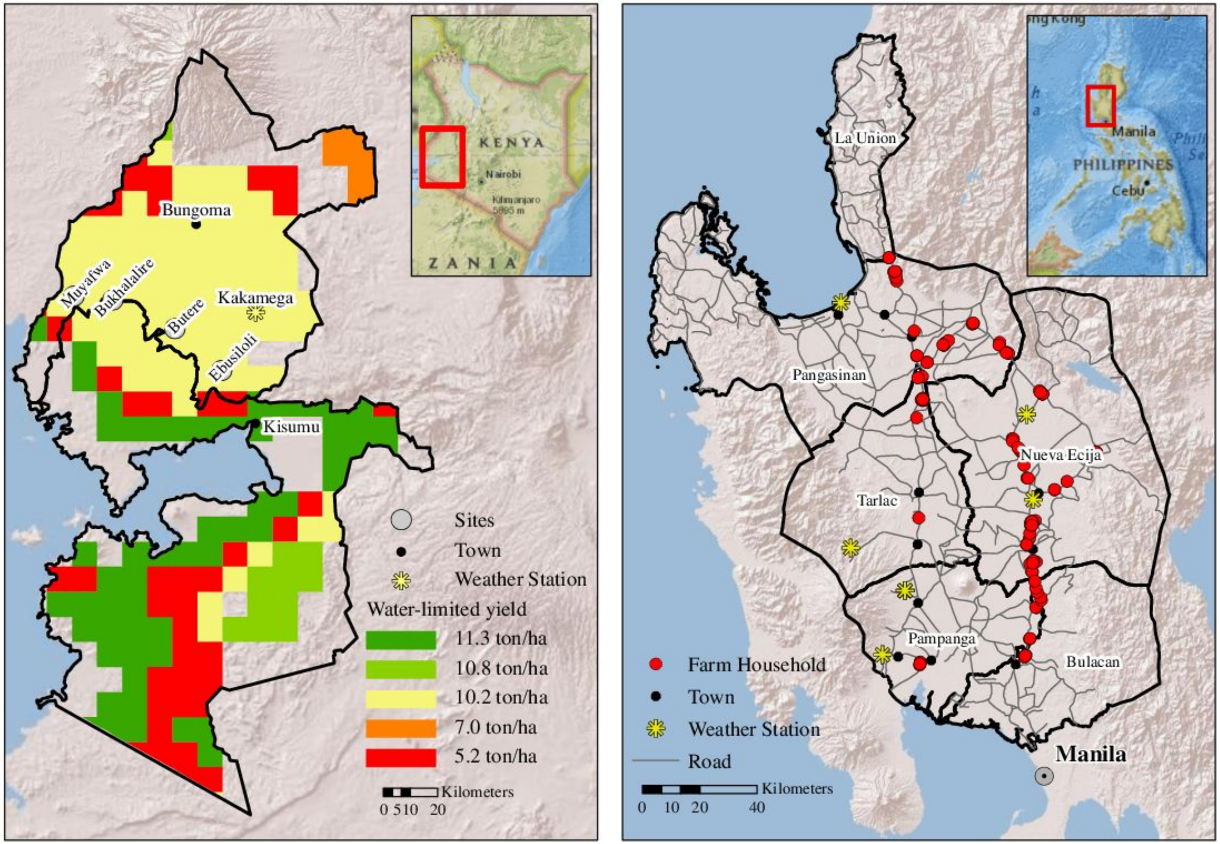

Figure 7.1: Location of the study sites in Kenya (left) and the Philippines (right). Water-limited yields in Kenya were retrieved from GYGA and correspond to the average over the period 1998 - 2014. Households in the Philippines were interviewed within the Central Luzon Loop Survey.

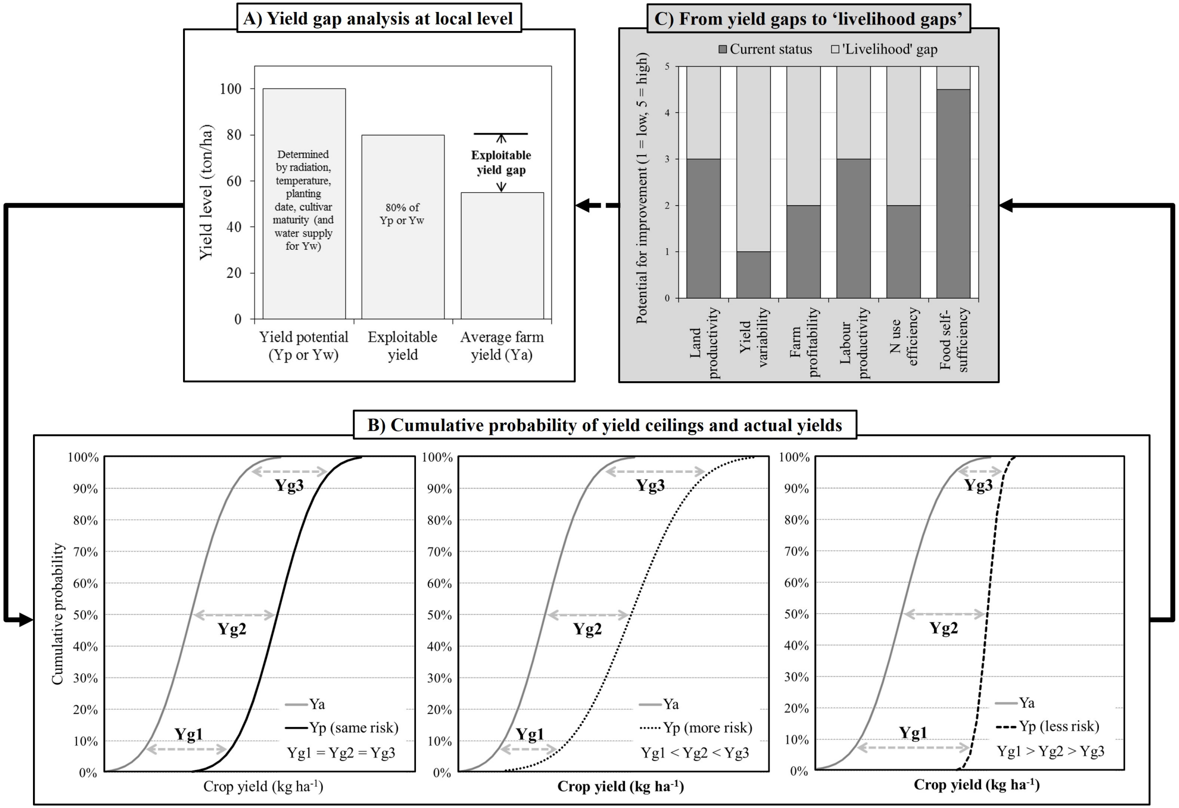

Figure 7.2: Conceptual framework to integrate yield gap analysis within development-oriented agronomy. A) Standard yield gap analysis relying on the difference between yield ceilings (Yp or Yw) and actual yields (Ya), e.g. van Ittersum et al. (2013); Lobell et al. (2009). B) Yield gap analysis based on hypothetical probability distributions of Yp and Ya with differing levels of risk (same, greater or lower than the risk of Ya). C) Integrated assessment of farm level indicators to identify 'livelihood gaps'. White boxes depict the 'farming' domain in which yield gaps are assessed and the light grey box depicts the 'livelihood' domain in which on-farm and off-farm activities are considered jointly. Arrows illustrate the flow between methodologies required to assess the importance of yield gaps for rural livelihoods.

In agronomy, yield gaps are used to understand the relative contribution of growth-defining, -limiting and -reducing factors to Ya (van Ittersum and Rabbinge, 1997; Janssen et al., 1990; French and Schultz, 1984; Herdt and Mandac, 1981). However, this original purpose has been largely over-simplified in many recent studies (e.g. Kassie et al., 2014; Angulo et al., 2012; Hochman et al., 2012; Meng et al., 2012; Boling et al., 2010; Licker et al., 2010; Lobell et al., 2009; Bhatia et al., 2008 to name just a few) which reduced it to the difference between Yp (or Yw) and Ya (Figure 7.2A). On the one hand, the simplicity of calculating yield gaps has made this a powerful framing device for justifying policy interventions (Sumberg, 2012). But on the other hand, the concept is highly problematic, not least because of its flexibility in terms of the yield ceilings considered (examples for rice can be found in Stuart et al., 2016).

A first step forward is to consider variability and acknowledge that yield ceilings and Ya are not single estimates but can be represented by probability curves of varying likelihood (cf. Vanlauwe et al., 2016; Beddow et al., 2015). In practical terms, farmers hoping to attain the yield ceiling would be interested not only in the maximum production possible but also the risks associated with it: how large is the variance associated with the reported ceiling and are the probability curves normally distributed or skewed. Figure 7.2B shows a normally distributed probability curve for a hypothetical set of Ya and Yp. The Ya curve for a given cultivar in a given season is related to factors like planting dates and densities, soil water and nutrient regimes, pest and weed pressures, as well as frequency and timing of weeding operations. The standard yield gap calculation is shown in Figure 7.2B as the difference between the median value of Yp and the median value of Ya (Yg2 in each case). This rightward shift corresponds to the "treatment effect" of eliminating all pertinent limiting factors (cf. Vanlauwe et al., 2016).

If the yield ceiling has the exact same variance as existing practices, the yield ceiling and Ya curves would be parallel to each other and the yield gap would be a constant for all farmers and conditions (Yg1=Yg2=Yg3). However, if the yield ceiling has its own variance and level of risk, the yield gap would no longer be a constant, and would not be equivalent to the difference between the mean values of the yield ceiling and Ya for all farms or conditions. Two scenarios present themselves: the yield ceiling could be associated either with more risk or with less risk than is found in the Ya probability curve's variance. In situations where the yield ceiling is associated with higher levels of risk (i.e. if the increased yield relies on management practices that, even under researcher management, increased production in some cases but with significant risks of failure in others), the yield gap would be greater than the difference between the median values of the yield ceiling and Ya for the high end of the production curves (Yg3>Yg2), and less than that difference for the low end of the production curve (Yg1<Yg2). However, if researcher-managed conditions eliminated many of the risks associated with on-farm practice (i.e. by limiting biotic and abiotic stresses) and had less risk than the Ya probability curve, we would see the reported yield gap based on the mean values of yield ceiling and Ya under-estimating the yield gap for the lower half of the Ya curve (Yg1>Yg2), while over-estimating it for the upper half of the farms and conditions (Yg3<Yg2).

The relevance of the yield gap for farmers further depends on the overall importance, and contribution, of agriculture for their livelihoods (Figure 7.2C). Farming is not only about biophysical and technical issues (e.g. land productivity and resource use efficiency) but includes other livelihood dimensions as well (e.g. food security, economic viability and social drudgery). Identifying opportunities to close yield gaps in farmers' fields require a diagnosis of a set of indicators as well as knowledge about the importance of off-farm income, the proportion of hired labour to total labour and input-output price ratios, among other issues. Moreover, agronomy can contribute to improved rural livelihoods with interventions aiming at narrowing yield gaps and increasing resource use efficiency but this may come at the expense of labour productivity and gross margin. This suggests that agronomy per se cannot eliminate the 'livelihood gap' because of possible trade-offs between different livelihood dimensions.

Finally, the identification and adoption of appropriate innovations which respond to the livelihood concerns lead to a new set of yield ceilings, and yield gaps (van Dijk et al., 2017; de Koeijer et al., 1999), which are more closely matched to farmers' personal conditions. In other words, efforts should be made to identify interventions necessary for closing the yield gap component(s) of the 'livelihood gap'.

7.3 Maize-based farming systems in western Kenya

7.3.1 Establishing yield gaps for maize in Kenya

Actual yields (Ya)

Maize has been a staple food since the early twentieth century in Kenya and at the core of national agricultural research since the 1950s. A guiding principle for this research has been closing the yield gap between on-station and on-farm conditions.

However, even from the earliest days agronomists struggled with issues of data quality for both Ya and researcher-managed conditions (e.g. Allan, 1971; Eberhart, 1971). Kenya's maize breeding programme began in Kitale, in the high potential, western highlands, but its first plant breeder, A.Y. Allan bemoaned the lack of reliable on-farm yield data (Allan, 1971). He noted measurement as costly and difficult, the widely differing husbandry practices, and the fact that farmers did not accurately know the size of their fields and or the amounts of maize produced, since it might be harvested continuously over a long period rather than on a single date. This knowledge gap was made evident when Allan's "deliberate attempt (on-station) to approximate 'poor' husbandry ... yielded almost 50 percent more (1.97 Mg ha^{-1}) than the estimated average on-farm yield (1.35 Mg ha^{-1})" (Gerhart, 1975). This attempt at duplicating on-farm practice in 26 factorial trials had set six factors sub-optimally (time of planting, plant population, type of seed, standard of weeding, and use of nitrogen and phosphate), and speaks to the challenge of understanding the full extent of the constraints on maize productivity under farmers' conditions.

Table 7.1 shows maize Ya in Kenya for 2014, disaggregated by province and county, compared to various yield ceilings. Even in the most productive parts of the country (Rift Valley, Western, and Nyanza provinces), Ya values (whether provincial averages or the Ya from the highest yielding county in each province) remain well below Yp and Yw calculated by GYGA for each province. Although these are the best and most comprehensive Ya data from the government of Kenya, many authors note large discrepancies between official yield statistics and independent yield measurements in African countries (Tittonell and Giller, 2013; Wairegi et al., 2010). And there are important errors to consider in the Kenyan case. Not all maize is marketed in Kenya as much of it is consumed by subsistence producers or traded informally, so official Ya statistics are modelled estimates based on direct measures such as the purchases by the National Cereals and Produce Board (NCPB) and sampling within selected counties (GoK, 2015). Due to resource constraints, this sampling cannot be done every year or in every jurisdiction (Kibaara et al., 2008; Hassan et al., 1998b). Errors can arise from the lack of consistent crop-cutting, area estimation on irregular fields (Jaetzold et al., 2009), and the lack of controls on moisture content at the time of weighing (Grassini et al., 2015).

Data quality both for yield ceilings and Ya is highest for the most productive parts of the country, which is logical given the economic and policy incentive to have denser data coverage in these zones, but which devotes significant national resources to gathering data in areas where farmers' best yields are already approaching Yw.

Primary data are much sparser, and remote sensing is used to model productivity in the more marginal maize growing regions, which arguably are the regions where the yield gap is greatest (e.g. semi-arid Eastern Kenya; Grace et al., 2014) or where the economic marginal rate of return for closing the yield gap would be greatest (Hassan et al., 1998a). A similar knowledge gap appears in the absence of Yp or Yw for Coast or North Eastern provinces, which are absent from GYGA even though national research stations operate in Kwale and Mtwapa, and national seed companies have further developed hybrid and composite varieties suited to coastal soils (as reflected in the "optimal yield" targets column).

Yield ceilings (Yw and Yp)

Table 7.2 illustrates the most commonly used yield ceilings in yield gap estimations in Kenya. Many are variants of Yw, i.e. from researcher-managed conditions either on-station or on-farm, that reduce all biophysical limitations except moisture, since irrigation is not widely available in Kenya, especially in smallholder conditions. As a result, models rely heavily on the quality of input rainfall and soil moisture data (van Wart et al., 2013; Jaetzold and Schmidt, 1982), while Yw based on empirical data (the best yields under managed conditions) show considerable inter-annual and inter-seasonal variation (Rojas, 2007; Hassan et al., 1998a; Smaling and Janssen, 1993; KARI, 1993).

Field measures

Since Kenya liberalized its maize research in the 1990s, both private and public research are testing maize performance and Yw in similar ways. While initial on-station research in the 1960s and 1970s established the importance of early planting, high plant densities, and high standards of weeding (Allan, 1971), subsequent research has explored the crop response to inputs especially N and P (Njoroge et al., 1995; KARI, 1990). As a result, all treatments in more recent trials share many elements of crop husbandry that, for reasons discussed below, are not necessarily desirable or suitable for smallholder farmers. Although most farmers intercrop their maize with legumes, the vast majority of research trials simplify by planting sole stands of maize. For example, the National Maize Productivity Trials (NPTs) use comparatively few factorial treatments (usually only N or P combinations), with 1 - 4 replicates on 5m x 5m plots, meaning typically 24 - 48 maize plants harvested from the inner 4 out of 6 rows planted (Njoroge et al., 1995). Trial sites are chosen to be "representative of the target area" yet, unlike many of the farms that they are supposed to represent, should also be "as uniform as possible in terms of slopes,

Table 7.1: Regional distribution of county-level 2014 maize actual yields (Ya) in Kenya versus yield ceilings (Yw and Yp). Provincial and county Ya and areas (GoK, 2015); "Optimal yields" are the yield ceilings reported for the varieties best suited to each province from Kang'ethe, 2011; Yw and Yp are the reported averages based on 14 years of data 1998-2011 from GYGA. '*' Provinces were abolished as administrative units in 2010 but remain a useful way to categorize the 47 counties that were created to take their place.

Province *	**Maize Area (ha)**	**Maize Production (Mg)**	**Counties (#)**	**Country level Ya ($Mg\ ha^{-1}$)**		**County (name)**	**"Optimal yields" ($Mg\ ha^{-1}$)**	**GYGA Yw ($Mg\ ha^{-1}$)**	**GYGA Yp ($Mg\ ha^{-1}$)**
Rift Valley	670,847	1,816,386	14	Mean	**2.2**		**8.7 - 14.9**	**7.5**	**14.3**
				Max.	4.2	Trans Nzoia		8.1	13.2
				Min.	1.0	Baringo		n.d.	n.d.
Eastern	575,023	336,778	8	Mean	**0.7**		**5.0 - 8.0**	**4.2**	**10.2**
				Max.	1.0	Meru		3.5	12.1
				Min.	0.4	Kitui		3.2	9.7
Nyanza	350,193	547,199	6	Mean	**1.6**		**7.5 - 9.0**	**6.0**	**13.6**
				Max.	2.0	Kisumu		3.9	15.3
				Min.	1.2	Nyamira		n.d.	n.d.
Central	190,894	154,217	5	Mean	**0.9**		**7.0 - 8.0**	**4.6**	**6.2**
				Max.	1.3	Nyandarua		7.0	14.5
				Min.	0.3	Kiambu		n.d.	n.d.
Western	243,239	558,966	4	Mean	**2.1**		**7.0 - 11.0**	**8.9**	**12.0**
				Max.	2.7	Bungoma		8.9	12.0
				Min.	1.4	Busia		8.9	12.0
Coast	81,446	79,873	6	Mean	**0.9**		**6.0 - 9.0**	**n.d.**	**n.d.**
				Max.	1.3	Tana River		n.d.	n.d.
				Min.	0.5	Kilifi		n.d.	n.d.
North Eastern	3,587	1,919	3	Mean	**0.5**		**3.0 - 5.0**	**n.d.**	**n.d.**
				Max.	0.8	Garissa		n.d.	n.d.
				Min.	0.2	Wajir		n.d.	n.d.
Country Total	**2,115,229**	**3,495,339**	**47**	**Mean**	**1.7**			**7.1**	**14.7**

soil types and previous cropping history" and "preferably not planted with maize the previous season" (Njoroge et al., 1995).

Although this configuration is in accordance with international agronomic norms, and allows for robust statistical testing while maximising the use of scarce research station land, this design erases or renders invisible a number of important production factors that would not be found in farmers' fields, or design elements that would themselves be considered treatments. Participatory, on-farm research with farmers in six communities of western Kenya between 2001 and 2008 under the "Strengthening Folk Ecology" project (Ramisch et al., 2006) worked to build a shared understanding of soil fertility and crop husbandry under smallholder conditions. Focus group discussions and individual interviews with smallholder farmers and scientists in the early stages of the project helped identify discrepancies in the knowledge and attitudes of the different groups.

Table 7.3 summarizes findings from 2002 and 2003 about how researchers and smallholder farmers each defined "good husbandry". These are not trivial differences. Many of the high yielding practices defined as good husbandry by researchers, and which are used as the basis for determining yield ceilings under researcher-managed trials, were seen as too risky or simply unwise by smallholder farmers. For example, planting on a single date, planting relatively few seeds per hole, and planting maize without an intercrop (especially a legume) each appeared to many respondents as risky gambles that did not offer alternatives if the promised yield improvements failed to materialize. Farmers also discounted the value of selection criteria such as minimizing in-field soil variability, competition with woody plants, and avoiding histories of continuous maize cultivation, since technologies were deemed more impressive if they could succeed in more typical and challenging contexts. Finally, standard measures taken by researchers to protect trial sites and treatments from interference or loss (fencing, herbicides, pesticides, and clean weeding) were each considered treatments in their own right and therefore as interesting as potential innovations as the varietal testing or soil fertility management treatments ostensibly being tested (Ramisch, 2011).

Early planting was especially noted as a problematic practice even while acknowledging its favourable impact on yields if rains fell reliably. The key concern is a social one: maize that matures earlier than the maize of one's neighbours would attract considerable attention in the hungry season between weeding and harvest when food is scarce. Farmers with maize ripening in this period face considerable social pressure to share this abundance with less fortunate kin or neighbours (Ramisch, 2016), or

Table 7.2: Examples of yield ceilings against which actual on-farm yields in Kenya have been compared. Terms are reported as in the original reference.

Source	**Term**	**Type**	**Description**
GYGA	Potential yield	Yp	Modeled maximum yield assuming no limitations (nutrients, soil, competition, biotic stresses)
Jaetzold and Schmidt (1982)	Potential yield	Yp	Based on agro-ecological zone, soil maps, top farmers in competitions, on-station research
GYGA	Water-limited potential yield	Yw	Similar to Yp above, but limited by water supply, and hence influenced by soil type and field topography
Jaetzold and Schmidt (1982)	Climatic yield potential	Yw	Modeled maximum yield based on soil-crop-water data (6 million rainfall points, crop data from best on-farm and on-station trials)
KARI (1993)	Water-limited potential yield	Yw	Best yields in researcher managed fertilizer trials on-station
FURP (1987)	'Good' potential yield	Yw	80% of the best yields in researcher managed fertilizer trials on-farm
Kenya Seed Co. (unpublished)	Potential yield	Yw	Best yields in researcher managed varietal trials on-farm
Hassan et al. (1998a)	Economic optimum	Y_{AE}	Target based on profit maximization (i.e. where the marginal product fertilizer is equal to the nutrient price ratio)
Hassan et al. (1998b)	Feasible yield	50% of (Yw Y_{HF})	Target based on averaging the best present yields of farmers (Y_{HF}) and of researcher-managed trials (Yw)

Table 7.3: Husbandry on researcher-managed sites that are taken as standard ”best practices” before testing treatment efforts, versus equivalent on-farm practices in western Kenya. Source: Focus group discussions conducted with farmers and scientists in the six study communities of western Kenya (2002 - 2003; cf. Ramisch, 2014).

Researcher-managed treatments	Farmer practices (Western Kenya)
• Early planting (at or within a few days of the onset of rains)	• Delayed planting (waiting to confirm rains, for soil to soften, or when labour is available).
• Single planting date	• Staggered planting is common (especially if labour is scarce or rains fall intermittently, or if seed remains from initial planting).
• 2 (or 3) plants per hole, thinned when plants reach 20 cm	• 2-6 plants per hole, 1-3 allowed to mature, others thinned for fodder at or before cobbing.
• Gaps (double-planted holes can be left unthinned to compensate for poor germination in adjacent spots)	• Gaps may be filled by replanting if rains continue.
• Sole stands of maize	• Maize is normally intercropped with common beans.
• Avoiding shading or root competition with woody plants	• Boundary hedges (to mark tenure) and timber or fuel trees are common on farms; technologies must be able to do well in shaded contexts.
• Row planting (with tape)	• Row planting is done for maize, but spacing is based on experience or energy levels; Beans or intercrop may be broadcast depending on time.
• Clean weeding (within two weeks of seedling appearance, again as needed)	• Weeding labour is scarce (especially for small households) and often prioritises only problem areas or high value crops.
• Herbicides (pre- and post-emergence)	• Any herbicide use would be considered a treatment.
• Previous crop residues and stubble removed	• Previous crop residues may be burnt in field (or removed for fodder).
• Top dressing to follow up fertilizer application at planting	• Top dressing is considered an additional treatment.
• Minimizing in-field soil fertility variability (hotspots, waterlogging, etc.)	• Variability is the norm; areas that underperform one season (e.g. waterlogged) might be the only areas that yield in a different season.
• Avoiding previous cultivation history (e.g. long history of continuous maize cropping)	• Areas with long cultivation history are the norm; a technology that works in such sites it will be seen much more favourably than one that only works on virgin or privileged sites.
• Pesticides (e.g. stemborers)	• Any pesticide use would be considered a treatment.
• Fencing against wildlife incursion	• Fencing (or crop guarding) would be considered a treatment.
• Single harvest date	• Maize is harvested over extended periods, i.e. green (for roasting), before or once it has dried (usually determined by labour availability or threats of theft or wildlife loss), or dried grain is allowed to stay unharvested in the field until labour is available.

worse may have this maize stolen and roasted if they are unwilling to share (Verma, 2001). In other words, the ability to valorize on practices that support a supposed yield ceiling may be compromised by a farmers' unwillingness to incur the social or moral consequences. Bunei et al. (2013) confirm that fear of crop theft indeed drives some farmers in the Rift Valley to avoid planting early ripening varieties of maize or beans, or to not invest in soil fertility management that would otherwise increase their maize yields.

7.3.2 Maize yield distributions in western Kenya

One way to accommodate these differences in perception is to acknowledge the range of probabilities of outcomes for farmers' yields and yield ceilings. Researcher- and farmer-managed maize trials from four of the 'Strengthening Folk Ecology' field sites are presented in Figure 7.3 to illustrate the challenge of moving from yield gaps (as conceptualized in Figures 7.2A and 7.2B) to 'livelihood gaps' (Figure 7.2C). As discussed above, the choice of the yield ceiling (Table 7.2) in the Kenyan context is often related to Yw.

Figure 7.3 illustrates the challenges of a yield gap analysis in western Kenya once we acknowledge probability curves for both Ya and Yw. In the lower panel (Figure 7.3B), four cumulative probability curves represent the range of farmer-managed Ya harvested in the long rains of 2007 in four communities (Muyafwa, Bukhalalire, Ebusiloli and Butere; see Figure 7.1). Households differed substantially in their socio-economic and soil fertility status in ways which were unequally distributed through the sample, making it important to consider the cumulative distribution of Ya and not just a measure of central tendency such as the mean. The upper panel (Figure 7.3A) shows four different but potentially plausible yield ceilings:

1. The cumulative probability curve for the twelve top-yielding researcher-managed plots in the same communities in the same season (Yw_researcher), which averaged 5.1 Mg ha^{-1}. Each study site had three researcher-managed experiments that were used for demonstrations. The highest yielding plots in each of these experiments (typically a "best practice" management of nutrients and crop husbandry), could therefore be considered by farmers as the upper benchmark of possible production under local conditions.

2. The single highest yield recorded on a farmer-managed plot (Yw_farmer) in any of the communities over the project's lifetime (6.0 Mg ha^{-1}). Many farmers (and the research team) took note of the fact that the best recorded yield by a

farmer was higher than that seen on the researchers' own "best practice" plots (at least in 2007). Other participating farmers found it easy to accept this yield as the most logical target as the "maximum" production possible under local conditions.

3. The Yw curve for all 154 farmer-managed plots simulated for non-limiting nutrient conditions using QUEFTS (Smaling and Janssen, 1993, as applied by Tittonell et al., 2008b; Yw_QUEFTS), which averaged 5.4 Mg ha^{-1}. Because soil fertility is heterogeneous in these sites (Vanlauwe et al., 2007), the QUEFTS model was used to predict crop yields based on actual soil fertility measures and assumptions about the recovery fractions of applied nutrients. Although the model does not consider other husbandry factors, it does provide another possible measure of the yields possible if nutrients were non-limiting.

4. The Yw simulated by the Global Yield Gap Atlas for these sites and which is shown in Figure 7.1 for the climate zone to which all four sites belong (Yw_GYGA, 10.2 Mg ha^{-1}). The model simulates the performance of recently-released, high yield cultivars grown in single stands, and is calibrated on the basis of local soil properties, long-term (>10 years) daily weather data, and crop management data (sowing date or sowing rule and plant population density) as per http://www.yieldgap.org/web/guest/methods-model-calibration.

5. The Yw curve generated from the 14 years of experimental data collected by the Global Yield Gap Atlas for the Kakamega weather station, which averaged 8.9 Mg ha^{-1} (Yw_GYGA14). Unlike the modelled data (Yw_GYGA), this value might be easier for farmers to relate to, since it was derived from actual yields obtained under optimal, researcher-managed conditions.

Yield gaps for any of the four site's Ya curves represented in Figure 7.3B could legitimately be calculated against each of the five yield ceilings shown in Figure 7.3A. The dotted lines bridging the upper and lower panels of Figure 7.3 link to either the point estimate or the mean value of the probability curve for each of the five Yw curves or estimates, leading to five different sets of possible ceilings. The highest yield ceiling (the GYGA point estimate of 10.2 Mg ha^{-1}) is twice the value of the lowest ceiling, the average of the top-yielding researcher-managed plots (5.1 Mg ha^{-1}). Including all fourteen years of available data moderates the GYGA target somewhat, to 8.9 Mg ha^{-1} but precipitation at the Kakamega weather station on which Yw_GYGA14 is based is 1971 mm yr^{-1}, which is at the more favourable end of the precipitation range recorded in the study sites (1270 - 2000 mm yr^{-1}, Ramisch et al., 2006). Beyond the diversity of possible Yw targets, yield gap calculations based on the averages or point estimates of Yw would tell us little about the probabilities of attaining the maximal yields for any of the given farms within the sample. Even the QUEFTS calculations, which drew upon the farmers' own local conditions, generated yield gaps between 0.7 - 6.5 Mg ha^{-1} but do not convey the probabilities related to closing those gaps. Yet this is a crucial consideration for farmers deciding whether/how to close any gap.

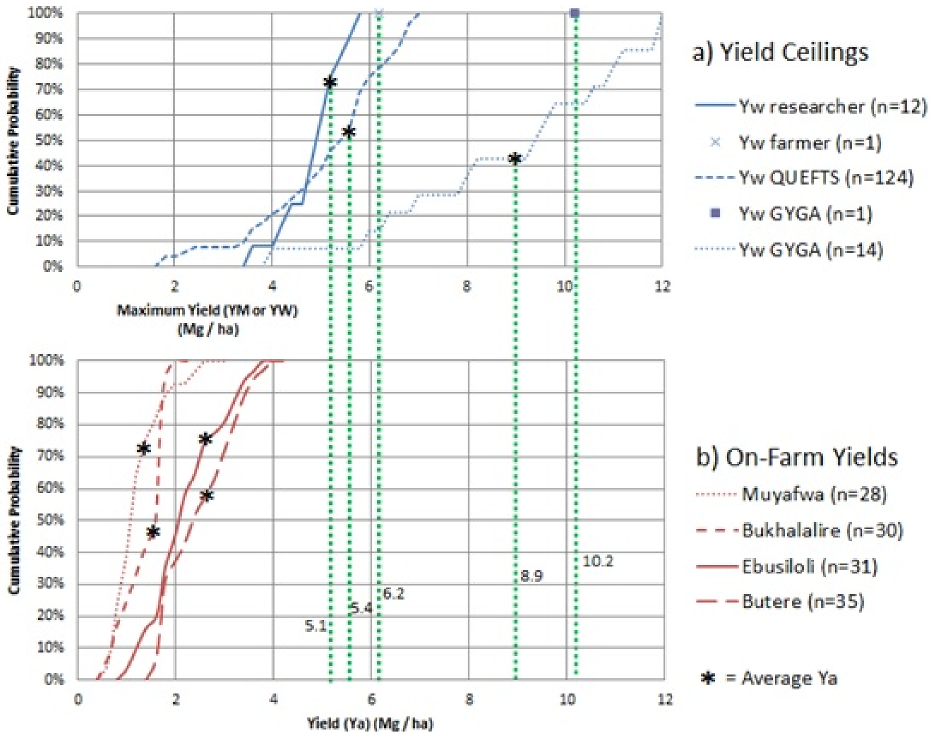

Figure 7.3: Yield gaps based on (a) four different variants of Yw and (b) cumulative probability curves of Ya of four communities (Muyafwa, Bukhalalire, Ebusiloli and Butere) of western Kenya. Data refer to the 2007 long rains. The asterisks (*) represent the mean Ya for each community: not only is there a wide range of Ya in each site but (as the means' position above or below the median show) the yields are also not normally distributed. Ya as well as Yw_farmer, Yw_researcher and Yw_QUEFTS are all unpublished data from the 'Strengthening Folk Ecology' project. Ya curves are single farm harvests (n=124) in the four listed study communities. Yw_QUEFTS was calculated for all 124 farms using the same methodology as the concurrent study reported by Tittonell et al. (2008b). In a), the point estimate (Yw_GYGA) is the mean value of water-limited yield for rainfed maize, filtered by national scale climatic zone reported for western Kenya (see Figure 7.1). The 14 year data set (Yw_GYGA14) refers to the Kakamega weather station (1998-2011).

The Kenyan case study suggests that neither Ya nor Yw data can easily be taken at face value, since the calculation of both is based on sets of socially-determined decisions about what to accept and what to exclude. By the same token, the significance of the presence (and scale) of a gap between Yw and Ya needs to be fit within a context of the probabilities of attaining a given yield as well as the livelihood significance of what closing (or failing to close) that gap would mean.

7.4 Rice farming systems in Central Luzon, Philippines

7.4.1 Context and household survey

Rice is the staple food in the Philippines. The average per capita rice consumption in the country has increased over time from ca. 100 kg capita^{-1} year^{-1} in the 1980s up to ca. 130 kg capita^{-1} year^{-1} in the 2000s (USDA and FAOSTAT databases). Despite the rapid development of the off-farm sector (e.g. construction, industry, transport, services and remittances) over the past decades, rice farming remains an important activity in Central Luzon contributing up to 25% of the total household income (Takahashi and Otsuka, 2009) and being an important source of employment to many (landless) peasants (Kerkvliet, 1990). Double rice cropping systems are common in this region, with a wet season (WS) crop between June/July and September/October and a dry season (DS) crop between December/January and March/April. Historically, the traditional season for rice farming was the WS but the increasing investments in irrigation facilities and the release of short cycle varieties made possible the cultivation of a subsequent rice crop in the DS (Cassman and Pingali, 1995).

The Central Luzon Loop Survey is a historical household survey which has been collected by the International Rice Research Institute (IRRI) every 4 to 5 years since 1966 up to now. It has been conducted since then to monitor changes over time in crop management and household characteristics in the rice-based farming systems of Central Luzon (Moya et al., 2015). On average, 103 rice farming households were interviewed in the WS and 59 in the DS. Most households were interviewed in the WS and DS of the same crop year but the sample size is lower in the DS because of water-related constraints or cultivation of other crops.

7.4.2 Rice yields and yield gaps

Rice yield gaps (Yp - Ya) in Central Luzon were on average 3.2 Mg ha^{-1} in the WS and 4.8 Mg ha^{-1} in the DS during the period 1979 - 2012 (Chapter 3; Laborte et al., 2012). There was no significant increase in Ya during the WS, which remained ca. 3.8. Mg ha^{-1} over the period analysed (Figure 7.4A). Stagnation of Ya in the WS may be attributed to greater risks of lodging due to typhoons at high N application levels (Lampayan et al., 2010; Lansigan et al., 2000; Loevinsohn et al., 1993). Conversely, there was a significant increase in Ya during the period 1979 - 2012 from ca. 4.0 Mg

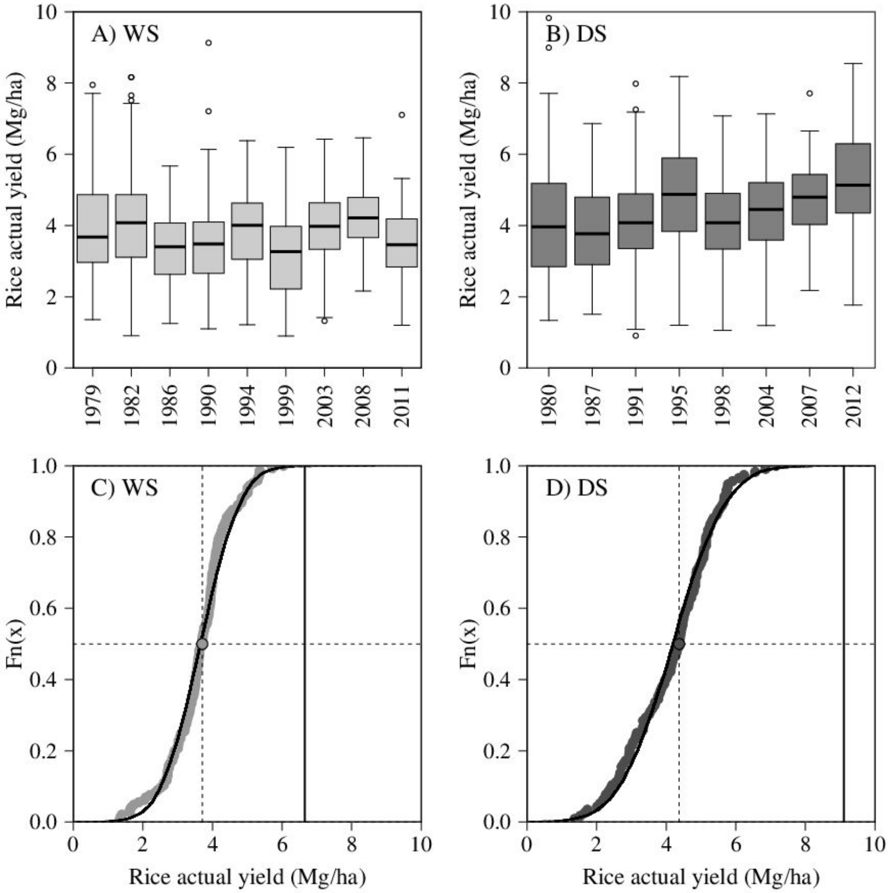

Figure 7.4: Rice yields and yield gaps based on cumulative probability curves of Ya in Central Luzon, Philippines. Data for the WS and DS are presented on the left (A and C) and right (B and D), respectively. In C) and D), Ya refer to averages across years for each farm and the horizontal dashed line shows the 50% probability, the vertical dashed line the median of Ya and the vertical solid line the simulated Yp averaged over the WS and DS periods (6.7 and 9.2 Mg ha^{-1}, respectively; see Chapter 3 for further details). The normal distribution is shown in the solid line next to the distribution observed in the data.

ha^{-1} in 1980 up to 5.2 Mg ha^{-1} in 2012 (Figure 7.4B). Typhoons do not occur during this season, which reduces climatic risks considerably, and provides more favourable growing conditions for rice, as indicated by the higher Yp compared to the WS. In addition to increases in Ya, there was also a significant increase of N application during this season. Modern varieties were readily and widely adopted over the past half-century in Central Luzon (Chapter 4), which also contributed to increases in Ya.

The distribution observed for rice yields approximated the normal distribution in both seasons (Figures 7.4C and 7.4D). The median of Ya was 3.7 Mg ha^{-1} in WS and 4.4 Mg ha^{-1} in DS. However, there was a large variability around this value with a minimum of 1.3 and 1.4 Mg ha^{-1} and a maximum of 6.1 and 7.2 Mg ha^{-1} in the WS and DS, respectively. The highest Ya values were close to Yp in the WS (minimum difference of 0.6 Mg ha^{-1}) but not in the DS (minimum difference of 2.0 Mg ha^{-1}). Yield gap variability can be directly observed in Figures 7.4C and 7.4D as the horizontal distance between Yp and Ya. From these, it can be concluded that rice yield gaps in Central Luzon were as large as 5.3 Mg ha^{-1} and as small as 0.6 Mg ha^{-1} in WS and as large as 7.7 Mg ha^{-1} and as small as 2.0 Mg ha^{-1} in DS.

7.4.3 Moving towards ’livelihood gaps’

Yield gaps per se are not very informative about the possibility to improve livelihoods of rural households (Figure 7.2C). For this purpose, they need to be analysed in relation to other indicators of farm performance capturing the broader livelihood aspects within which farming takes place. The farm level indicators analysed were labour productivity (kg ld^{-1}), gross margin (PhP ha^{-1}) and rice self-sufficiency at household level (%). Labour productivity and gross margin were quantified as described in Chapter 4 and the calculation of rice self-sufficiency is detailed in Appendix (Section 7.8).

Labour productivity

Labour productivity increased over time in both WS and DS (Figures 7.5A and 7.5B). The increase in this indicator was particularly evident in the DS, from about 51 kg ld^{-1} in 1980 up to about 120 kg ld^{-1} in 2012, which can be explained by a combination of increases in Ya (Figure 7.4B) and adoption of labour-saving technologies including direct-seeding, small tractors, threshers and herbicides (Moya et al., 2015). The median labour productivity observed during the WS and DS was about 55 and 79 kg ld^{-1} (Figures 7.5C and 7.5D), respectively, and the variation observed in this indicator approximated a normal distribution in the WS, but not in the DS.

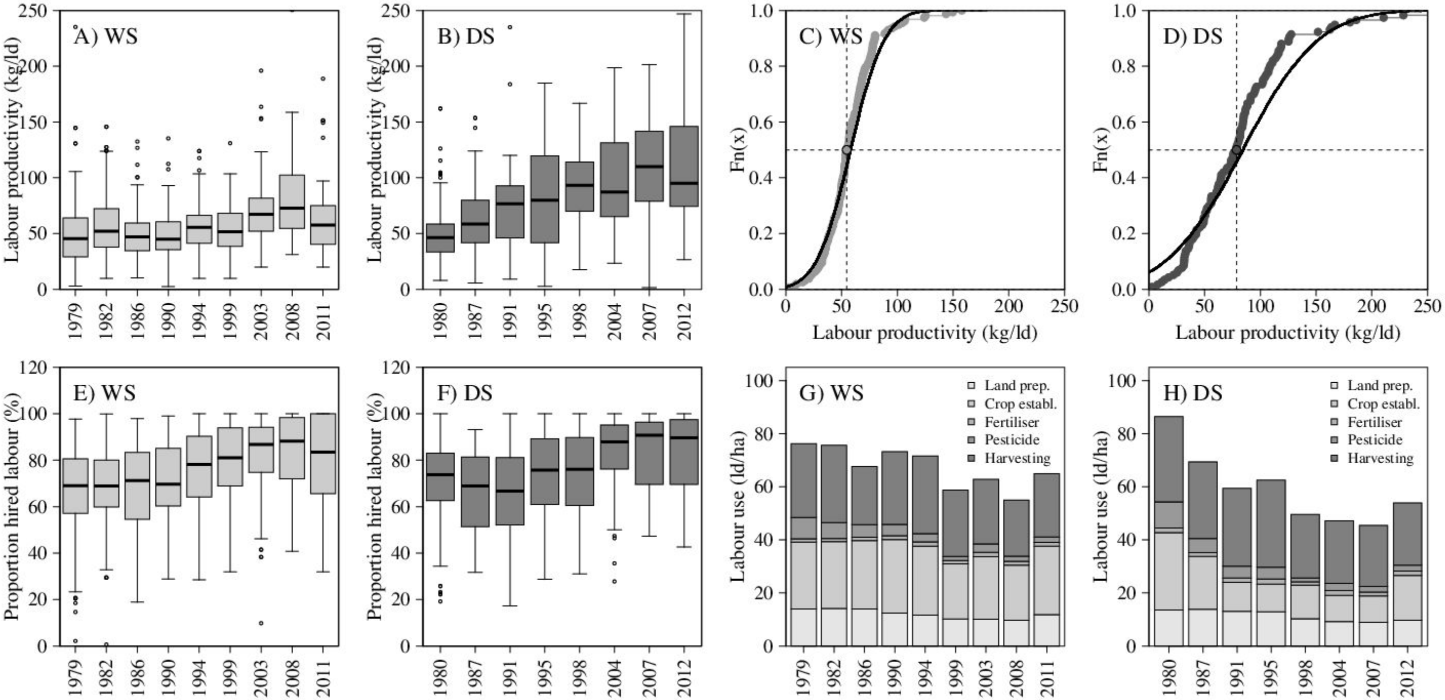

Figure 7.5: Labour productivity (A - D) and use (E- H) in rice-based farming systems in Central Luzon, Philippines. In C) and D), labour productivity refers to averages across years for each farm. In these panels, the horizontal dashed line shows the 50% probability, the vertical dashed line the median of labour productivity while the normal distribution is shown as a solid line next to the distribution observed in the data.

Four labour arrangements predominate in the rice-based farming systems of Central Luzon. Family labour refers to household members directly involved in rice farming and it is mostly used for operations which are less labour intensive (e.g. fertiliser application). Exchange labour is a peculiar labour arrangement in which members of one household help members of another household in exchange for a similar amount of labour of the latter household in the fields of the former household. This type of labour was categorized as family labour. Hired labour can take two forms: 1) hiring of temporary workers to perform labour demanding activities such as transplanting and harvesting or 2) hiring of permanent labourers who are responsible for all the crop management activities in exchange of ca. 10% of the total production. Hiring labour has had a positive effect in the rural economy by providing employment to many landless peasants (Estudillo and Otsuka, 1999; Kerkvliet, 1990).

There was a sharp increase in the proportion of hired labour over time up to an average of ca. 82% of the total labour used for rice farming in 2011/2012 (Figures 7.5E and 7.5F). The replacement of family by hired labour was particularly evident after the year 2000. This trend suggests that the importance of rice farming as a landowning household's primary occupation declined over time. The decline in hired labour between 1979 and 2000, and its increase afterwards, was associated with land preparation and crop establishment practices used by farmers. For example, labour use for crop establishment declined in the 1980s and 1990s and slightly increased again in the 2000s (particularly in the DS) because of the adoption of direct seeding during the 1990s and the re-adoption of transplanting after the 2000s (Figures 7.5G and 7.5H), respectively. Labour use for land preparation and crop management slightly decreased over time mostly due to the adoption of small tractors and herbicides, respectively.

Gross margin

There was a significant decline in rice gross margin over time during the WS (Figure 7.6A) while no trend was observed in the DS (Figure 7.6B). In the WS this can be explained by a slight decline in revenues, attributed to yield stagnation and a slight decline in paddy prices (Chapter 4), and by an increasing of production costs due to greater use of hired labour and material inputs (data not shown). The negative gross margin of WS rice observed for many households shows rice farming is not economically rewarding during this season especially given current amounts of hired labour and high labour wages in the region (Moya et al., 2004). It is worth noting that the median gross margin in the 2008 WS was even negative. By contrast, revenues and production costs during the DS increased over time. The increase in revenues was explained by increases in Ya (Figure 7.4B), since paddy prices slightly declined (Chapter 4). Similarly to the WS, increasing production costs were explained by greater use of hired labour and, to a less extent, material inputs (e.g. real prices of N declined between the late 1960s and early 2000s, after which there was a sharp increase).

The median gross margin from WS rice and DS rice was 5305 and 9525 PhP ha^{-1}, respectively, and the distribution observed in this indicator approximated a normal

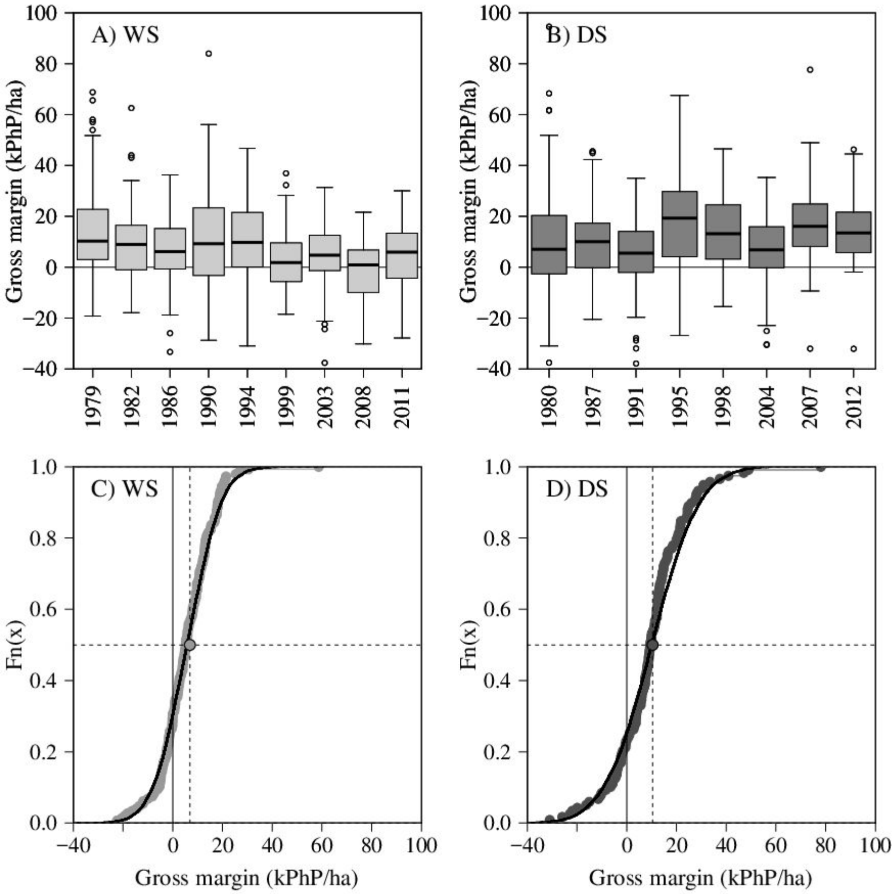

Figure 7.6: Gross margin of rice-based farming systems in Central Luzon, Philippines: A) and B) trends over time, C) and D) cumulative probabilities. In C) and D), gross margin refers to averages across years for each farm. In these panels, the horizontal dashed line shows the 50% probability, the vertical dashed line the median of gross margin while the normal distribution is shown as a solid line next to the distribution observed in the data.

distribution in both seasons (Figures 7.6C and 7.6D). The threshold probability for positive gross margins was ca. 30% in the WS and ca. 20% in the DS. The maximum gross margin observed was far greater than the minimum gross margin in absolute terms. However, the low median and high probability of negative gross margin in the WS question again the importance of rice farming as a source of income for many households in this season. For instance, in 1979 WS and 1980 DS ca. 25% of the households had negative gross margin from rice farming as compared to 2011 WS and 2012 DS in which 14% and 5% of the households had negative gross margin, respectively. The largest number of households with negative gross margin (ca. 48%) was recorded in 2008 WS. The importance of positive gross margin for each household is likely to increase with decreasing importance of off-farm income. Negative gross margin over consecutive seasons for households not depending on off-farm income may limit investments in e.g. education in the short-term and lead to migration to urban areas in search of non-agricultural employment in the long-term.

Rice self-sufficiency

During the period 1979 - 2012, less than 15% of the households did not meet their yearly rice requirements (Figure 7.7). This indicates that the majority of households in Central Luzon achieved rice self-sufficiency given current rice yields, area available and consumption requirements (i.e. household size and *per capita* rice consumption) and that most households have a considerable amount of land surplus to produce rice for the market. Indeed, Central Luzon is known to be the rice bowl of the Philippines, and particularly of Metro Manila.

The land required for rice self-sufficiency at household level remained constant over time with an average of ca. 0.5 ha per household, and ranging between ca. 0.4 and 0.6 ha per household in 2011/2012 and 1986/1987, respectively. The additional amount of land required to achieve rice self-sufficiency at household level (i.e. land deficit) was also rather constant over time and its negligible magnitude confirms that most households were able to meet their domestic consumption needs. The average land surplus, i.e. the actual amount of land cultivated not needed for rice self-sufficiency at household level, declined significantly over time from ca. 2.7 ha in 1979/1980 to 2.1 ha in 2011/2012. The significant decline in land surplus can be attributed to increased rice demand and decreased rice supply. The latter occurred due to greater *per capita* consumption and adult:child ratio, while the former was due to a slight, but significant, decline in rice cropped area explained by the cultivation of other crops and/or land conversion to non-agricultural uses (Chapter 4).

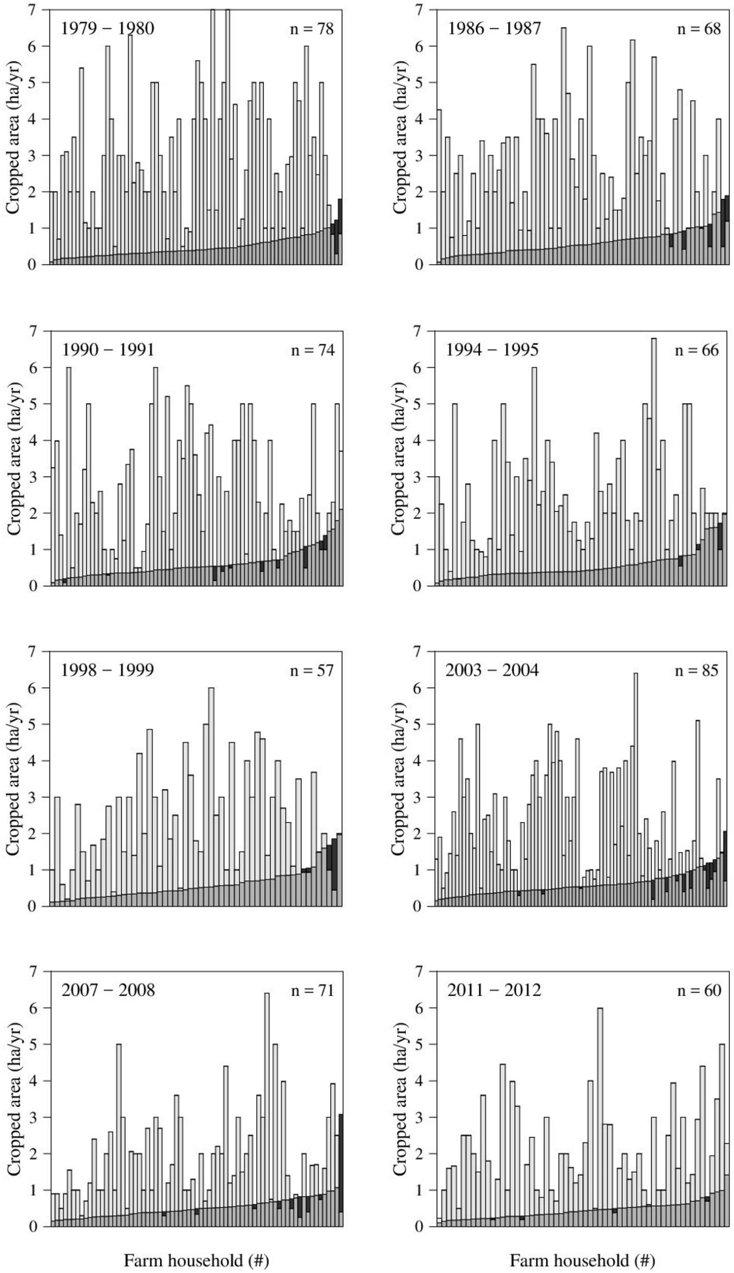

Figure 7.7: Yearly land requirements for rice self-sufficiency (ha, medium grey), land surplus (ha, light grey) and land deficit (ha, dark grey) between 1979 - 2012 in Central Luzon, Philippines. Cropped area refers to the sum of the area of each field cultivated by an individual household in both WS and DS. Each bar represents the total cropped area for an individual household.

Rice and rural livelihoods

The data from our sample shows large rice yield gaps persist in Central Luzon, particularly during the DS (Figure 7.4), indicating there is considerable scope to increase rice production in this farming system (Chapter 3; Laborte et al., 2012). However, as with the Kenyan examples, the benefits from narrowing such yield gaps depend on whether or not increased production 1) is associated with similar or reduced risks compared to the current situation, 2) translates into significantly greater returns to land and/or labour and 3) is required to meet self-sufficiency needs at household level.

Climatic risk is a very important aspect in Central Luzon as intensification of N use increases the probability of lodging during the WS due to frequent incidence of typhoons during this time of the year (Lampayan et al., 2010; Lansigan et al., 2000; Loevinsohn et al., 1993). An example of the impact of typhoon damage on rice stands under field conditions is depicted in Figure 7.8. In this on-station trial, the treatment replicating N application rates used by farmers in the region (i.e. 60 kg N ha^{-1}) did not suffer from lodging after typhoon while the treatment aiming to achieve Yp (i.e. 180 kg N ha^{-1}) resulted in complete crop failure. In these extreme situations, farmers would most likely find little incentive to increase N application rates beyond levels currently used.

In addition to the biophysical risks associated with higher yields and more intensive management practices, it is important to also consider farmers' financial and food security incentives to narrow the yield gap. Figure 7.7 demonstrates that actual rice yields and current farm sizes have been able to satisfy the caloric need of most households between 1979 - 2012. Rice self-sufficiency at household level indicates that the main incentive for narrowing rice yield gaps would then be for commercial, rather than domestic consumption, purposes. Formal research might indeed encourage such a shift in production orientation, but in a context where narrowing the yield gap goes in tandem with high input costs (especially hired labour), marginal additional income from selling rice (Figure 7.6) and increased risk of more intensive input use in the WS (e.g. Figure 7.8).

The low gross margin of rice farming has been compensated by an increase in non-agricultural sources of income such as off-farm employment and remittances (Chapter 4, Moya et al., 2015; Takahashi and Otsuka, 2009; Estudillo and Otsuka, 1999). The high labour wages for on-farm work observed in Central Luzon compared to other rice bowls in Southeast Asia (Moya et al., 2004) did not discourage most households from replacing family by hired labour over the past half-century (Figures 7.5E and 7.5F). This indicates that opportunity costs for family labour in the non-agricultural sector are also high in the region as otherwise it would not be economically rational to depend so much on hired labour for rice production, which currently accounts for more than 50% of the total production costs. In this context, any efforts to narrow the yield gap would most likely require even heavier reliance on hiring permanent and/or casual labourers for operations like crop establishment and harvesting (Figure 7.5).

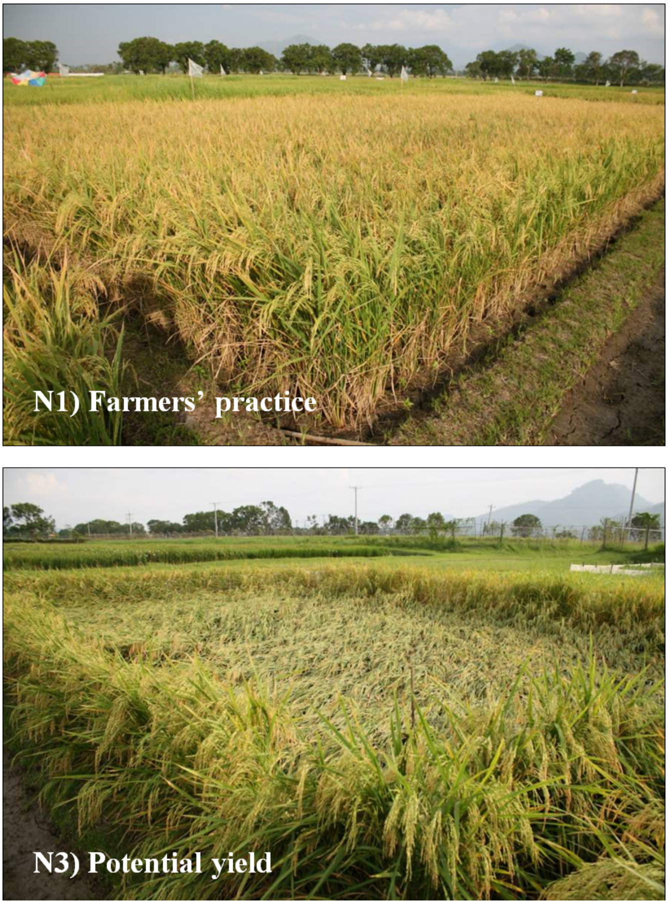

Figure 7.8: Lodging of rice in an on-station trial conducted in IRRI during 2014 WS. The photo on top refers to a treatment replicating the farmers' practices in the region (N application of 60 kg N ha^{-1}) while the photo on the bottom refers to a treatment aiming for climatic potential yield (N application of 180 kg N ha^{-1}). The photos were taken in plots adjacent to each other a few days after the region was hit by a typhoon. *Source*: João Vasco Silva, IRRI, 23rd September 2014

The analysis of rice yield gaps vis-à-vis gross margin, labour productivity and rice self-sufficiency provides insights into the relative importance of rice farming for rural livelihoods in Central Luzon. In addition to the trade-offs between yield gap closure on the one hand and labour productivity and N use efficiency on the other hand (Chapter 4), it is questionable whether rice production will be intensified in Central Luzon under prevailing conditions given that 1) greater yields are associated with greater risk of crop failure in some periods of the year (Figure 7.8), 2) economic returns to land are marginal (Figure 7.6) and 3) rice self-sufficiency at household level is not at stake (Figure 7.7). Observed trends in off-farm income and labour dynamics further confirm the clear preference of most households for more regular, and less labour intensive, sources of income which leaves us with the question of who should close which yield gap.

7.5 Discussion

The standard yield gap is conceptually simple and therefore easy to explain to non-specialist policy-makers (e.g. 'how much more food could our farmers grow?'). This makes it all the more challenging to contest or replace. The two case studies illustrate how moving from yield gaps to 'livelihood gaps', and acknowledging the probabilities and risks inherent in bridging such gaps, are important points to communicate in contexts where agronomy has much to offer but where farming is only one element within a suite of livelihood options. Understanding how farmers (and researchers) perceive risk and variability is crucial to help making better choices about new technologies.

7.5.1 Incorporating variability into yield gap analysis

Farming systems research can point to multiple successes in both the Kenyan and the Philippines context. The dramatic increase of maize yields observed in Kenya between 1965 and 1980 is widely hailed as a success story for African maize agronomy, which saw the release and widespread adoption of varieties suited to a range of agro-climatic conditions (Smale and Jayne, 2003). This raised maize Yp in Kenya, not only by improving the grain production potential of individual plants but also by allowing farmers to grow maize in both the long and short rains season in many parts of the country. A similar productivity improvement was observed in the Philippines with the development of irrigation facilities and short cycle varieties which allowed the cultivation of rice in both WS and DS (Cassman and Pingali, 1995).

Despite these successes, current yields (Ya) both of maize in Kenya and rice in the Philippines lag behind their full potential. The two case studies demonstrate that by failing to consider variability and risk in both actual, on-farm conditions (Ya) and yield ceilings (Yw or Yp), conventional yield gap analyses only partially explain where and how Ya for these staple crops could be improved. For example, in both contexts experimental evidence shows that increased fertilizer use is associated with greater risk of crop failure. In Kenya, the nationwide Fertilizer Use and Recommendations Project (FURP) trials provided experimental and modelled Yw far above Ya, but the yield increases due to fertilizer use were also often associated with increased yield variability (Rötter and van Keulen, 1997). In the Philippines, greater N use was also associated with greater yield variability due to increased risk of lodging during the WS (e.g. Figure 7.8; Lampayan et al., 2010). Extending the yield frontier should not only offer the promise of higher Ya but should also seek, wherever possible, to reduce the risk associated with those improvements, i.e. tightening the distribution of the cumulative probability curves rather than stretching or skewing them.

Field and farm variability is recognized in different site-specific approaches developed in agronomy (Vanlauwe et al., 2015; Dobermann et al., 2002). As data from Central Luzon shows, where a farm lies on the Ya probability curve (Figure 7.4) is not independent of its placement on labour productivity or gross margin probability curves (Figures 7.5 and 7.6). In western Kenya, the FURP trials not only included a range of cropping conditions for maize (sole stands and intercropped), but reported its findings using three different probability thresholds (0.33, 0.50 and 0.66) for purposes of calculating economic benefits. Many of the yield ceilings in Table 7.2 claim to be based on purely biophysical criteria which are not as objective as they might first appear: their selection and application help shape crop productivity outcomes. For example, Herdt (1979) found a significant decrease in on-station rice yields when the production objective was switched from yield maximization to gross margin maximization and recent efforts have shown the same for maize yields in Tanzania (van Dijk et al., 2017). However, even these optimal yields need to be contextualized at farm(ing) systems level to understand how exactly households may benefit from possible yield gains.

7.5.2 Embedding yield gaps into rural livelihoods

Yield gaps that identify where (and by how much) crops and farm(ing) systems can be improved are important guides for technology development, but must also help guide decision-making about which technologies are worthwhile investments. In western

Kenya, many farms are smaller than 0.5 ha and can provide 3 - 7 months of food security (Ramisch, 2014). The most food insecure households in this site would need to increase Ya by a factor of four while avoiding e.g. theft or post-harvest losses.

Ideally, in an on-farm setting, it would be possible to identify the degree to which a given variable could be improved given specific objectives and resource constraints. Optimisation techniques can be used for the more easily quantified data, which tend to be the biophysical and those economic aspects that engage with the formal sector. However, these methods are more conjectural and ill-suited for qualitative data, or those quantitative data that lack precision or accuracy. Farmers make decisions about how much energy or resources they want to invest in agriculture (let alone in closing yield gaps). However, while decisions to opt out of farming (or to rely on hired rather than family labour) may be rational for many households, better information about how farmers perceive and respond to variability and risk could guide better technology development for those households that decide to persist in the agricultural sector. Such qualitative information would also be helpful to guide those households at the lower end of the Ya or yield ceiling cumulative probability curves to understand the risks, opportunities and trade-offs.

This makes it essential that the real costs and risks associated with attaining yield ceilings are made more explicit. The differing perspectives on what constitutes a ”treatment” (Table 7.3) reveal the frustrations that farmers have with researchers as well as the ways in which researchers perceive and understand the constraints farmers face. The fact that a crop will perform better under non-limiting conditions is no surprise, but there is generally a strong incentive on the part of a researcher to demonstrate a convincingly large yield gap between researcher-managed and farmer practice. As Allan (1971) noted ”demonstrating differences of 10% in yield is a waste of time in developing countries”, which led him to advocate for a ”package approach” that would maximize all possible input factors so that even if one factor had no effect, the overall differences of the new maize-growing system would be ”very much higher than the average yields [Ya] round about”. While such strategy might persuade some farmers that agronomists can be ”good farmers”, it is arguable that many other farmers might see the complex lengths required to bring about this new system as impossible, uneconomic, or irrelevant.

7.5.3 Looking up: Dynamics at national level

Even if individual households are making the appropriate choice (for them) to get out of staple food production, the national food security picture is deteriorating in both

Kenya and the Philippines. From a policy standpoint, the yield gap is paired with the costs of imports to meet consumption needs not covered by domestic production (e.g. van Ittersum et al., 2016). Between 1970 and 1991, maize imports represented 2.9% of total annual maize consumption in Kenya, growing to an average of 12% for the period 2000-2010 (Kimani and Gruere, 2010). Population increase over this period accounts for some of the increased demand, but the liberalization of the maize marketing system in the 1990s played a role in making low cost maize available from the neighbouring countries. This discouraged some smallholders from investing resources in higher cost domestic maize production (Nyoro et al., 2004). Most of the maize is grown in the former Rift Valley, Western, and Nyanza Provinces (84%; Table 7.1) and, other things being equal, the country could meet its consumption needs by closing part of the yield gap in these regions.

National policies in the Philippines have focused on promoting rice self-sufficiency and providing high income to farmers (e.g. PDA, 2012) while ensuring affordable prices for consumers (PSA, 2017). However, rice imports in the country accounted for ca. 10% of total rice consumption in 2015 and there has been an increasing import dependency over time (PSA, 2017), particularly from Vietnam and Thailand. The Philippines has been the largest rice importing country in the world (Dawe et al., 2006) due to e.g. unfavourable geography (Dawe, 2006), population growth (FAOSTAT, 2017) and poor irrigation infrastructure (Barker and Levine, 2012). Future perspectives are not very promising given that rice production needs to double, due to population growth, if self-sufficiency is to be achieved by 2050. This seems very improbable given ”current trends in yield growth, existing production technologies, and prevailing conditions” (Laborte et al., 2012). The later include the fact that rice farming is one in many livelihood activities and getting close to ’hobby farming’ for many households in Central Luzon, the prime rice producing area of the country.

In both countries, it is important that consumers have confidence that the national food system delivers the commodities in sufficient quantity and at affordable prices. Regular maize shortages due to fertilizer price instability, or to extreme events such as the Kenya’s 2008 post-election violence indicate that, even if the Rift Valley could grow all of Kenya’s maize, it is logic for smallholders countrywide to keep growing their own (higher cost) maize ”just in case” (Brooks et al., 2009). Similar concerns are observed in the Philippines where rice farm-gate prices (Cabling and Dawe, 2006) and consumer rice prices (FAO/FPMA, 2017) are among the highest in Southeast Asia. Rice prices are established by the National Food Authority (NFA), a governmental agency which controls the national rice market in terms of rice imports and acquisition of rice from farmers at support price.

7.6 Conclusion

The two case studies explored how yield gap analysis can be expanded to consider the variability of yield ceilings and actual yield and to situate the yield gap within a wider livelihood context. Using the 'livelihood gap' concept to develop pathways to improved food security needs to acknowledge that smallholder farming in many contexts is one within multiple livelihood activities and it often entails biophysical and/or socio-economic risk as yield gaps are narrowed. Future efforts to improve the crop performance in smallholder farming systems must not entail greater risks than are found in the current situation, nor should they encourage out-migration at the expense of the national (or global) food security discourse. Developing a shared understanding of risk and opportunity is crucial, since farmers and researchers often differ in their perceptions and understanding of risk and even of what constitutes a "treatment".

Finally, moving from yield gaps to 'livelihood gaps' will require more careful consideration of data needs. Data intensive modelling approaches such as the Global Yield Gap Atlas could incorporate more layers devoted to modelling uncertainty and variability at the sub-national scale. For instance, the communities studied in Kenya correspond to four different production contexts, as seen in their distinct probability curves (Figure 7.3), but are represented by a single climatic zone. Finally, while acknowledging that such efforts are useful to sketch the operational space within which food production operates, it is not clear that smallholders' decisions are well-served by a model whose coarse resolution of agro-meteorology impedes the provision of context-specific estimates of yield ceilings or the probabilities of attaining them.

7.7 Acknowledgements

JVS would like to acknowledge Dr. Pytrik Reidsma (PPS-WUR) and Dr. Alice Laborte (SSD-IRRI) for their comments in previous versions of this manuscript. JJR would like to acknowledge Dr. Michael Misiko (SEP-CIMMYT) for his insights and discussions in the field in Kenya, which inspired an earlier version of this manuscript. Both authors are indebted to the editors of the Special Issue, Dr. Jens Andersson (SIP-CIMMYT) and Prof. Dr. Ken Giller (PPS-WUR), for the stimulating discussions and for the invitation to draft this manuscript.

7.8 Appendix

The land required for rice self-sufficiency at household level was quantified for each unique household × year combination in the Central Luzon Loop Survey using the methodology of Hengsdijk et al. (2014).

Annual rice supply was calculated based on the quantity of rice kept by each household on a yearly basis, i.e. in both WS and DS. Total rice production was derived from the cultivated area and rice actual yields adjusted to 14% moisture content (expressed in kcal; Quilty et al., 2014). Areas and yields reported for each growing season were summed and expressed on a yearly basis. The quantity of rice kept by a household was obtained by subtracting in-kind payments of permanent workers as well as harvesting and threshing activities (ca. 25% of total rice production) from the total rice consumption. The quantity of rice kept was further corrected to a constant milling rate of 65% and post harvest losses were assumed to be as high as 37% (IRRI Rice Knowledge Bank).

Annual rice demand was estimated based on the number of household members and their energy requirements per year. There was a slight decline in the number of household members during the study period from an average of ten individuals in 1979 - 1980 to seven individuals in 2011 - 2012. There was also a sharp increase in annual per capita rice consumption from about 105 kg capita^{-1} year^{-1} in 1980s and 1990s up to 129.4 kg capita^{-1} year^{-1} in 2000s (USDA and FAOSTAT databases)[†]. These were converted to energy units by assuming an energy content for rice of 3630 kcal kg^{-1} (Quilty et al., 2014). Energy requirements of household members under 18 years old were assumed to be on average 50% of those of an adult.

A 'land deficit' was identified for the households unable to produce enough rice to meet the energy needs of the household members, i.e. for households in which the demand for rice was greater than the supply from their fields. The land deficit thus indicates how much additional land is required for a household to achieve rice self-sufficiency given observed actual yields. In contrast, a 'land surplus' was identified for the households which were able to produce rice beyond household energy needs; hence it indicates the area which is not required for rice self-sufficiency.

[†]It is important to mention that our results may under-estimate the rice area required for rice self-sufficiency as *per capita* rice consumption may be considerably higher in Central Luzon. As an example, the average *per capita* rice consumption recorded in a recent household survey conducted by IRRI in 2014 (Chapter 4) was about 170 kg capita^{-1} year^{-1} (data not shown).

CHAPTER 8

General discussion

8.1 Introduction

This thesis provides insights into the magnitude and causes of yield gaps in rice-based farming systems of Central Luzon (Philippines, Chapters 3 and 4), arable farming systems in the Netherlands (Chapter 5) and mixed-crop livestock systems in Southern Ethiopia (Chapter 6). Methods of frontier analysis were combined with concepts of production ecology so that the richness of individual farm level data, in terms of variation in yields and management practices, could be analysed using agronomic theory. Yield gaps were further embedded within the broader farm level constraints and opportunities using different statistical methods. The analysis was complemented as much as possible with field visits and interactions with farmers, local researchers and other key informants in the three countries. The combination of statistical, model-based and empirical approaches made it possible to achieve a high level of detail and comprehensiveness, which complements initiatives at higher scales such as the Global Yield Gap Atlas (www.yieldgap.org).

This final chapter summarizes the main findings obtained regarding the magnitude of yields gaps as well as their crop and farm level determinants. Moreover, it provides a comparative analysis of crop and farm performance across the three case studies with the objective of contextualizing the dynamics observed in each farming system and of identifying options for sustainable intensification. The three overarching hypotheses introduced in Chapter 1 and revisited in this chapter are: 1) the efficiency yield gap is most important in the Netherlands, the technology yield gap is most important in Ethiopia and the efficiency, resource and technology yield gaps (abbreviated as intermediate yield gaps) are of similar importance in the Philippines; 2) yield gaps are determined by farm level conditions related to farm(er) characteristics and cropping frequencies as well as farmers' objectives and resource constraints and; 3) agriculture needs to be intensified in the 'developing South' while improving sustainability is more applicable in the 'developed North'. The chapter ends with a reflection on the sustainable intensification of agriculture and on the methodological contributions and limitations of this thesis, and opportunities for further research.

8.2 Decomposing yield gaps at crop level

8.2.1 Magnitude across case studies

Crop yields were smallest in Southern Ethiopia, intermediate in Central Luzon (Philippines) and highest in the Netherlands (Figure 8.1, year 2012). For instance, average maize and wheat yields in Hawassa (Southern Ethiopia) were 1.6 and 2.7 t ha^{-1}, respectively, which corresponds to ca. 23 and 27% of the simulated water-limited yield (Yw). Rice yields in Central Luzon varied between 3.2 t ha^{-1} in the wet season and 4.8 t ha^{-1} in the dry season, which were ca. 55 and 49% of the simulated climatic potential yield (Yp), respectively. In the Netherlands, actual yields of spring barley were 5.5 t ha^{-1} (64 % of Yp) and actual yields of winter wheat were nearly

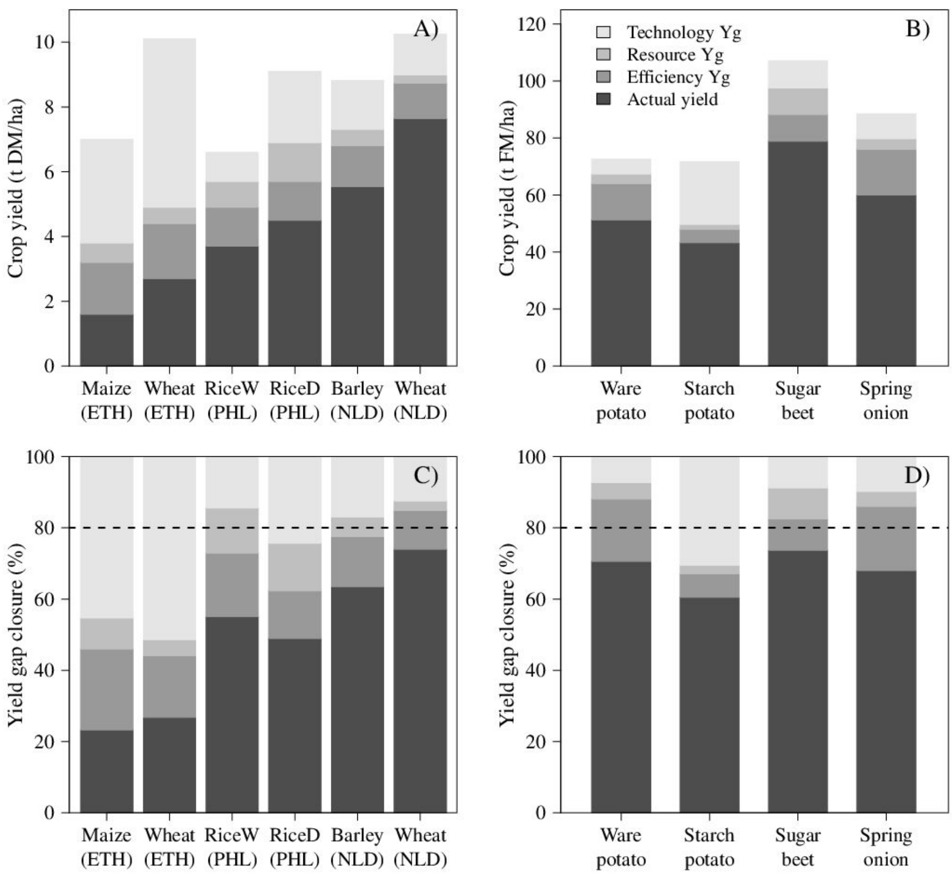

Figure 8.1: Magnitude of actual yields and yield gaps for cereals and arable crops in the three case studies analysed in this thesis. All data refers to the year 2012. Panels A) and B) show the yield gap in absolute terms, while panels C) and D) show the yield gap in relative terms (as % of Yw in Ethiopia and Yp in other countries). Yield data for cereals is provided on a dry matter (DM) basis and for other crops in the Netherlands on a fresh matter (FM) basis. Background data are provided in Chapter 3, Chapter 5 and Chapter 6. Country codes: ETH = Ethiopia, PHL = Philippines, NLD = Netherlands.

three times as high (i.e. 7.6 t ha^{-1} or 74% of Yp) as the wheat yields observed in Asella (Southern Ethiopia). Yield gaps smaller than 30% of Yp were also observed for ware potato, sugar beet and spring onion in the Netherlands. These figures are in line with conventional wisdom that yield gaps are small (20 - 30% of Yp) in North-west Europe, intermediate (30 - 50% of Yp) in Southeast Asia and large (>50% of Yw) in East Africa (e.g. Chapter 2; van Oort et al., 2016; van Ittersum et al., 2013; Mueller et al., 2012; Neumann et al., 2010).

Yield gaps were mostly attributed to the technology yield gap in Southern Ethiopia, to the efficiency (and technology) yield gaps in the Netherlands and to three intermediate yield gaps in Central Luzon, Philippines (Figure 8.1). The technology yield gap of maize and wheat in Southern Ethiopia was as high as 45 and 52% of Yw, respectively. The efficiency yield gap was also high in these farming systems as its closure would nearly double actual yields. The yield gap of rice in Central Luzon during the wet season was equally explained by the efficiency, resource and technology yield gaps (18, 13 and 15% of Yp) while during the dry season the relative contribution of the technology yield gap (24% of Yp) was slightly higher than that of the other yield gaps (each ca. 13% of Yp). For cereals in the Netherlands, the yield gap was equally explained by the efficiency and technology yield gaps: 11 and 12% of Yp for wheat and 14 and 17% of Yp for barley, respectively. The results for other arable crops indicated a large relative importance of the efficiency yield gap for ware potato (18% of Yp) and spring onion (18% of Yp), of the technology yield gap for starch potato (31% of Yp) and an equal importance of the three intermediate yield gaps for sugar beet (each ca. 9% of Yp). These results confirm the first overarching hypothesis related to the relative importance of each intermediate yield gap in the different case studies.

8.2.2 Crop management determinants

The determinants of the efficiency yield gap differed per farming system (Table 8.1). For maize in Hawassa, the efficiency yield gap was greater for late sown crops and increased with increased proportion of women labour used for sowing (Chapter 6). For wheat in Asella, the efficiency yield gap increased with the increasing proportion of hired labour for sowing and weeding (Chapter 6). Indeed, farmers reported difficulties in handling hired labourers during focus group discussions conducted in this site (van Eerdewijk and Danielsen, 2015). For rice in Central Luzon, the efficiency yield gap declined with later first fertiliser application date, with earlier second fertiliser application date and with earlier first and second pesticide application dates

Table 8.1: Summary of the most important management factors explaining the efficiency, resource and technology yield gaps. Background information can be found in Chapter 3 (Central Luzon), Chapter 5 (the Netherlands) and Chapter 6 (Southern Ethiopia). Arable crops include ware, seed and starch potato, sugar beet, spring onion, winter wheat and spring barley.

	Efficiency yield gap	**Resource yield gap**	**Technology yield gap**
Southern Ethiopia			
Maize, Hawassa	Delayed sowing time Women labour for sowing	N application rate Labour for hand-weeding	Sub-optimal fertiliser use Pest and disease incidence
Wheat, Asella	Hired labour for sowing Hired labour for weeding	Herbicide application Labour for hand-weeding	Crop establishment method Sub-optimal fertiliser use Pest and disease incidence
The Philippines			
Rice, Central Luzon	Fertiliser application date Pesticide application date	N & K application rate Seed rate	Sub-optimal fertiliser use Water limitations
The Netherlands			
Arable crops	Unfavourable weather Machinery constraints	Management of growth-reducing factors	Narrow crop rotations Water limitations

(Chapter 3). This partially contradicts the recommendation of performing the first N application 10 days after transplanting (DAT) since the efficiency yield gap was smaller in fields where the first fertiliser application was done later than 19 DAT. It was not possible to explain the efficiency yield gap in the Netherlands due to lack of detailed management data (Chapter 5). However, the timeliness of management operations is likely to be affected by e.g. machinery constraints on large farms (Reidsma et al., 2015b) and/or unfavourable weather and soil conditions (van Oort et al., 2012; Droogers et al., 1996).

The resource yield gap was associated mostly with the management of growth-limiting factors in Hawassa and Central Luzon and with the management of growth-reducing factors in Asella and the Netherlands (Table 8.1). In Hawassa, maize yield responses to N and labour used for the first hand-weeding were observed (Chapter 6). Similar results were obtained in on-farm trials conducted in this site (for N: TAMASA, unpublished data; for weeding frequency: Workayehu and Wortmann, 2011). The determinants of the resource yield gap for wheat in Asella captured well the importance of weed management for this production system (Amare, 2014; Taa et al., 2004; Tanner et al., 1993). Rice yields in Central Luzon were greater in transplanted than direct-seeded fields, which influences seed use, and were positively associated with N application (especially in the dry season) and K application (Chapter 3). Matching N application levels with indigenous soil N supply is crucial to obtain high yields in this production system (Cassman et al., 1998) and yield responses to K have also been reported for rice in Central Luzon elsewhere (Witt et al., 1999; Dobermann et al., 1996). Management of growth-reducing factors was important in the Netherlands and a positive association was observed between e.g. fungicide use and the yields of ware potato and winter wheat (Chapter 5).

The causes behind the technology yield gap could only be assessed based on agronomic trials (and crop model simulations), literature and/or expert knowledge (Table 8.1). The large technology yield gap observed for cereals in Southern Ethiopia was attributed to sub-optimal management practices used in highest yielding farms compared to best-performing technologies assumed to simulated Yw (Chapter 6). These included lack of precision agriculture technologies (e.g. sowing and fertiliser application are broadcast in wheat), resource yield gaps of particular inputs (e.g. low N rates and sowing densities to achieve Yw and no use of K or plant protection agents) and, possibly, lack of adoption of modern varieties. Sub-optimal N rates explained the technology yield gap for (dry season) rice in Central Luzon as N rates in highest yielding fields were generally lower than the ca. 200 kg N ha^{-1} required to achieve

Yp (Chapter 3; Kropff et al., 1993). Water limitation during the growing season is another possible constraint in this region (Barker and Levine, 2012). This was also true for arable crops, especially cereals, in the Netherlands, in addition to narrow crop rotations for starch potato (Chapter 5), which increase the pressure of soil borne diseases (Scholte and s'Jacob, 1990). Further research is needed to gain deeper insights into the causes of the technology yield gap, particularly in Southern Ethiopia where it is largest.

8.2.3 Yield responses to N and P

Pooled data from the three case studies showed that cereal yield responses to N and P follow the law of diminishing returns (Figure 8.2; Nijland et al., 2008; Janssen et al., 1990), which was not evident when data were analysed separately per farming system (Chapters 3, 5 and 6). Maize and wheat farms in Southern Ethiopia, rice farms in Central Luzon and Dutch farms cultivating barley were in the steep part of the response curve while Dutch wheat farms were in the plateau of the response curve (ca. 8 t ha^{-1} or 75% of Yp). This means that yield responses to N and P were smaller at low (i.e. Southern Ethiopia and wet season in Central Luzon) and at high yield and nutrient levels (i.e. wheat in the Netherlands) than at 'intermediate' yield and nutrient levels (i.e. dry season in Central Luzon and barley in the Netherlands).

Most farms in Southern Ethiopia used less than 75 kg N ha^{-1} and 50 kg P ha^{-1} for cereals and achieved cereal yields lower than 4 t ha^{-1} (ca. 40% of Yw). Maize and wheat yield responses to N and P in Hawassa and Asella, respectively, were not as clear when the data was analysed separately compared to the pooled analysis, which suggests that yield responses to nutrients are small at low yield levels. Other studies conducted in the Ethiopia Rift Valley and the Southern Highlands reported similar ranges for crop yields and input use as well as a lack of clear yield responses to N and P at low yield and input levels (Getnet et al., 2016; Baudron et al., 2014).

In Central Luzon, N and P application rates for rice ranged between 20 - 200 kg N ha^{-1} and 0 - 80 kg P ha^{-1} and rice yields ranged between 2 - 6 t ha^{-1} (ca. 20 - 80% of Yp). Rice yield responses to N and P were observed during the dry season but during the wet season in this site. Clear yield responses to nutrients during the dry season can be explained by the more favourable climatic conditions to grow rice during this period of the year (when water is available) and by the higher nutrient application rates used by farmers compared to the wet season (Chapter 3). These observations were corroborated by a more representative household survey of rice farms in the region (Figure B6; Chapter 4, Supplementary Material) and by a number of independent studies (e.g. Gines et al., 2004; Witt et al., 1999; Cassman et al., 1998).

Arable farms in the Netherlands used 50 - 300 kg N ha^{-1} and 0 - 140 kg P ha^{-1} for cereals and obtained cereal yields between 4 - 9 t ha^{-1} (ca. 60 - 80% of Yp). Rates of P application to cereals were (close to) zero for many farms because management of P is done at crop rotation level and most P is applied to the potato crop. No yield

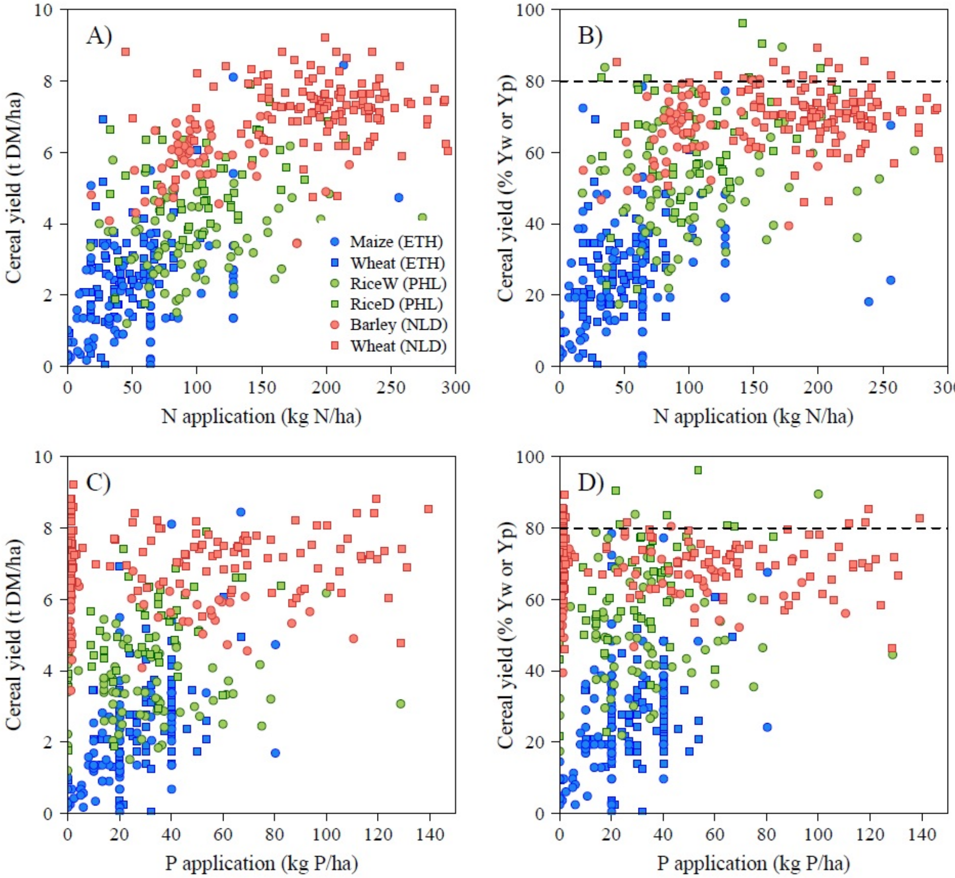

Figure 8.2: Cereal yield response to N and P applied in the year 2012 across the three case studies analysed in this thesis. Panels A) and C) present cereal yield in absolute terms while panels B) and D) present cereal yield in relative terms (as % of Yw in Ethiopia and Yp in other countries). Background data to this figure is provided in Chapter 3, Chapter 5 and Chapter 6. Each observation corresponds to one individual farm in each of the three case studies. Country codes: ETH = Ethiopia, PHL = Philippines, NLD = Netherlands.

responses to N were observed for wheat, a crop which received ca. 200 kg N ha^{-1} on average and yielded ca. 9 t ha^{-1}. Conversely, yield responses to N were observed for barley, a crop which received less than 150 kg N ha^{-1} on average and yielded ca. 7 t ha^{-1}. These data are currently used to inform official statistics (www.agrimatie.nl) and policy makers as to whether or not farmers comply with the N application standards in place (J.J. Schröder, personal communication).

Closing the yield gap of maize and wheat in Southern Ethiopia, of (dry season) rice in Central Luzon and, possibly, of barley in the Netherlands requires greater N and P application rates. Conversely, the lack of yield response to N and P for wheat in the Netherlands suggests there is scope to decrease nutrient application rates for this crop without compromising yield. This type of comparative analysis of farm(ing) systems at local level is valuable to gain deeper insights into crop performance in farmers' fields and to complement narratives at higher levels aiming to understand opportunities to increase food production and/or reduce the environmental footprint of agriculture (Hoffmann et al., 2017; Rockström et al., 2017; Mueller et al., 2012; Tilman et al., 2011; Neumann et al., 2010; Foley et al., 2005).

8.3 Crop yield gaps and farm performance

8.3.1 Intensification and resource availability

Farms in the Netherlands exhibited smaller yield gaps, had greater farm sizes and used much less labour than farms in Central Luzon and Southern Ethiopia (Figure 8.3). The average farm size was 0.8 ha in Hawassa, 2.0 ha in Asella and 1.7 ha in Central Luzon, which is way lower than the 54.1 ha in the Netherlands (Figure 8.3A; logged data). Moreover, the maximum farm size recorded in Southern Ethiopia and Central Luzon (ca. 10 ha) corresponded roughly to the minimum farm size observed in the Netherlands. Land was distributed unequally between farms in the three case studies with 45% of the farms owning only ca. 20% of the land. Similar distribution of farm sizes has been reported for the same regions by other studies (Lowder et al., 2016; van Vliet et al., 2015; Masters et al., 2013). At first sight, these case studies provide no clear evidence that small farms are more productive than large farms (even in Central Luzon a negative relationship was only observed in the wet season and when data from multiple years was pooled, Chapter 4). However, further research is needed to confirm this preliminary observation.

Farms in Southern Ethiopia and Central Luzon used up to 10 times more labour to cultivate one ha of land than farms in the Netherlands and a threshold of 15 ld[†] ha^{-1} was identified as the minimum and maximum labour use observed in the tropics and in the Netherlands, respectively (Figure 8.3B). There was no association between yield gap closure and labour use at farm level in the Netherlands while there was a positive association between these two variables in Southern Ethiopia and Central Luzon. This is not surprising given that most management operations in arable farms in the Netherlands are mechanized. By contrast, farming in Southern Ethiopia remains largely a manual activity and the use of the traditional oxen ploughing system is a good example of the little change observed since 'biblical' times (Gebregziabher et al., 2006; Aune et al., 2001; McCann, 1995). More recently, the little capital available was directed to purchase improved seeds, mineral fertilisers and, to a less extent, herbicides (Sheahan and Barrett, 2017; Abate et al., 2015).

The situation of rice farming in Central Luzon in the late 1960s was not much different. Back then, most farms used traditional varieties, applied ca. 40 kg N ha^{-1} and relied on hand-weeding, not herbicides, for weed control (Chapter 4; Moya et al., 2015; Launio et al., 2008). Rice yields of ca. 2 t ha^{-1} were common up to the mid-1970s (Chapter 4), yield levels comparable to the ones observed nowadays for cereals in Southern Ethiopia (Chapter 6; Getnet et al., 2016). As time passed, farmers were able to access credit and to replace labour by capital (Takahashi and Otsuka, 2009). These might have facilitated the adoption of improved varieties and of labour-saving technologies (Moya et al., 2015; Launio et al., 2008; Pandey and Velasco, 2002) which, together with investments in irrigation, contributed to increase rice yields from the 1970s onwards (Estudillo and Otsuka, 2006). A similar, but more dramatic, transformation occurred in the Netherlands where labour has been gradually replaced by capital to a point in which the level of debt and investment capacity became a major determinant of farm performance (Zhengfei and Lansink, 2006). The economic pressure associated with farming drove many non-profitable holdings out of business and triggered a steady increase in farm size over the past decades (CBS, 2015).

Farms in the Netherlands used considerably more energy, and capital, than farms in Central Luzon and Southern Ethiopia (Figure 8.3C). This was captured by the area

[†] Labour used was expressed in labour-days (ld) per ha for farms in the Philippines (Chapter 4) and in person-days per ha for farms in Ethiopia (Chapter 6). The two units are identical and express the total number of days worked by family and hired labourers per ha on an 8 hour working day basis. The unit hour per ha was used for farms in the Netherlands (Chapter 5) and this was converted to labour-days per ha also assuming an 8 hour working day.

cultivated per unit labour (in m^2 ld^{-1}), a proxy for the amount of energy used at the expense of capital under the assumptions that 1) the same area of land can be cultivated with less labour if more energy and labour-saving technologies are used and 2) labour and energy can be mutually substituted (de Wit, 1979). Dutch arable

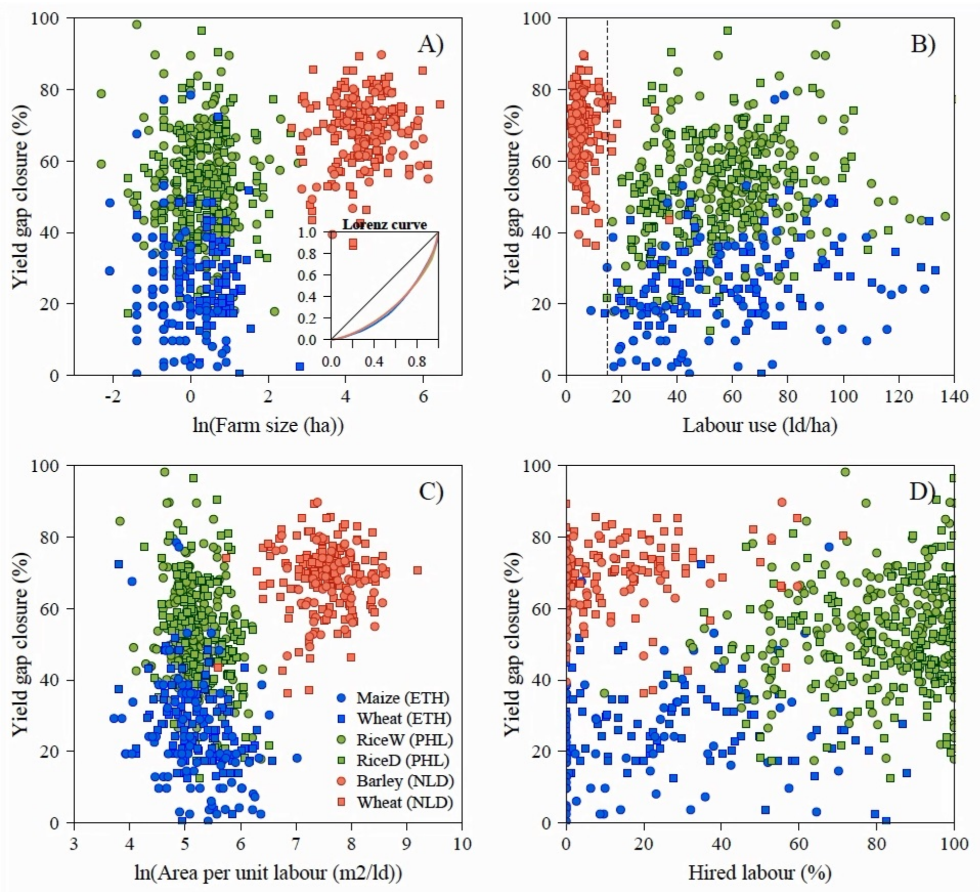

Figure 8.3: Yield gap closure for cereals and farm resources: A) farm size, B) labour use at farm level, C) area cultivated per labour-day at farm level and D) proportion of hired labour at farm level. The inset in panel A) shows the Lorenz curve for the variable farm size for the three case studies (y-axis = cumulative share of farm size; x-axis = cumulative share of households). Data for Ethiopia and the Netherlands refer to the year 2012 and data for the Philippines refer to years included in the Central Luzon Loop Survey after 2000. Background data to this figure is provided in Chapter 3, Chapter 5 and Chapter 6. Each observation corresponds to one individual farm in each of the three case studies. Country codes: ETH = Ethiopia, PHL = Philippines, NLD = Netherlands.

farms cultivated up to 2100 m^2 ld^{-1} while that value was about 10 times lower for farms in the tropics. Historical data for Central Luzon showed sharp increases in land productivity during the 1970s followed by sharp increases in labour productivity from the late 1980s onwards (Chapter 4). The latter was associated with an increase of the area cultivated per unit labour from ca. 160 m^2 ld^{-1} in the late 1980s up to ca. 225 m^2 ld^{-1} in 2012 during the dry season, as a result of the adoption of direct-seeding and small machinery. The increase in the wet season was not as sharp because transplanting remained the preferred crop establishment method during this season.

It is not unlikely that a greater use of external energy may favour yield gap closure in farmers' fields as this is associated with capital intensive technologies (e.g. combine harvester). The rationale behind this is that capital-intensive technologies require long-term strategic investments and force farmers to maximise the returns to that technology as it can only be used for specific purposes. This is very different from a situation where farmers strive to maximise returns to labour, which encompasses both on- and/or off-farm employment, and have limited access to e.g. information and/or markets. In these cases, *ad-hoc* investments, which are often not 'locked in', may be more suitable to meet household needs and preferences in the short term and to match farmers' perspectives and aspirations in the long-term (Dorward, 2009).

The argument above may explain why the yield gap is greater in Southern Ethiopia, and to a less extent in Central Luzon (where rice farming is one within many household activities), than in the Netherlands. However, it cannot explain why yield gaps are smaller in Central Luzon than in Southern Ethiopia. The difference in yield gap closure between these two regions may be attributed to a 'comparative advantage' of farming in Central Luzon compared to Southern Ethiopia due to three main reasons:

1. Rice farming in Central Luzon is mostly irrigated while cereal farming in Southern Ethiopia is rainfed;

2. Farms in Central Luzon have had better access to e.g. innovation, input markets and infrastructure than farms in Southern Ethiopia, where timely fertiliser availability is still problematic, and;

3. Farms in Central Luzon are less diverse (mostly dominated by rice and one additional maize or vegetable crop) than farms in Southern Ethiopia.

The last point deserved some recent attention in relation to the stagnation and decline of farm power in sub-Saharan Africa (Baudron et al., 2015). As an example, Chapter 6 showed temporal overlap in labour peaks during the growing season for small grains and pulses in Asella, but not in Hawassa where capital constraints and small farm sizes are widespread. Labour constraints to farm performance occur in other farming systems across sub-Saharan Africa as well (Ollenburger et al., 2016; Leonardo et al., 2015; Baudron, 2011; Stone et al., 1990). These are unlikely to have posed serious limitations to intensification of rice farm(ing) systems in the irrigated lowland plains of Southeast Asia, which may further explain the success of the Green Revolution in this region.

The proportion of hired labour used for farming was much greater in Central Luzon than in Southern Ethiopia or in the Netherlands (Figure 8.3D). Indeed, the contribution of family labour to rice farming in Central Luzon has gradually declined over the past half-century as a result of adoption of labour-saving technologies and increasing importance of hired labour (Chapters 4 and 7). This happened despite the high labour wages in Central Luzon relative to other rice bowls of Southeast Asia (Moya et al., 2004). Economic development in Central Luzon allowed family members to engage in multiple off-farm activities (Takahashi and Otsuka, 2009; Estudillo and Otsuka, 1999) and to hire landless peasants and/or permanent agricultural labourers to perform farm work (Otsuka, 2000; Kerkvliet, 1990), even if at relatively high labour wages and under peculiar contractual arrangements (Moya et al., 2004). This contrasts with the situation in Southern Ethiopia, where hiring labour is less common possibly due to a lack of capital to hire external workers, and in the Netherlands, where labour wages are much higher and most farm work is mechanised.

8.3.2 Intensification and sustainability

More transparency and better analysis of trade-offs between the 'intensification' and 'sustainability' aspects of agriculture is required in the sustainable intensification debate (Struik et al., 2014). The level of intensification in each case study (as measured by yield gap closure) was compared to sustainability indicators capturing socio-economic (gross margin and labour productivity) and environmental performance (N use efficiency). These indicators are far from extensive (e.g. Smith et al., 2017) but provide some interesting first insights. Other biophysical indicators which could be considered to improve the current trade-off analysis include water use efficiency (Grassini et al., 2011b), energy use efficiency (Quilty et al., 2014; Meul et al., 2007), nutrient balances (Cobo et al., 2010; Ponsioen et al., 2006) or greenhouse gas emissions (An et al., 2015; Grassini and Cassman, 2012).

There were sharp differences in revenues and costs[†]A. at farm level across case studies (Figure 8.4A). Farms in Southern Ethiopia obtained at most 1500 € ha^{-1} and had production costs lower than 300 € ha^{-1}. The average value was much lower, ca. 800 and 100 € ha^{-1} for gross returns and production costs, respectively. Gross returns derived from rice farming in Central Luzon were not much different from the ones observed in Southern Ethiopia (mean and maximum of ca. 900 and 2000 € ha^{-1}). However, production costs were considerably higher (average and a maximum of ca. 650 and 1250 € ha^{-1}) with labour costs accounting for more than 50% of the total production costs (Moya et al., 2015, 2004). For Dutch arable farms, gross returns per ha were way above the maximum of 2000 € ha^{-1} observed in Central Luzon and the same is true regarding production costs. In all farming systems, some farms had gross returns nearly double the production costs incurred while some other farms incurred greater production costs than the gross returns obtained. This was true for many rice farms in the Philippines, particularly in the wet season, and for some arable farms in the Netherlands and shows that farming is not always a profitable economic activity.

Gross margin per ha at crop level was greater for wheat and barley in the Netherlands and for wheat in Asella on the one hand than for maize in Hawassa and rice in Central Luzon on the other hand (Figure 8.4B). This indicator was estimated for each crop (wet and dry season rice were treated separately) as the difference between revenues per ha (crop yield $\times$ price) and input costs per ha (materials and hired labour in Southern Ethiopia and Central Luzon and only materials in the Netherlands). The maximum gross margin per ha in Asella and in the Netherlands was close to 0.33 € kg^{-1} grain and differences between these systems were largely attributed to lower cereal yields in Asella than in Netherlands. Conversely, gross margin per ha lower than 500 € ha^{-1} (and even negative) were observed for maize in Hawassa and rice in Central Luzon independently of the intensification level. This comparison indicates that gross margins per ha for wheat and barley are nearly three times greater than for maize and rice, but it is unclear whether this holds in other farming systems as well.

There was no trade-off between cereal yields and gross margins per ha at crop level in all farming systems analysed (Figure 8.4B). Although intensification goes in tandem with economic performance per unit land, there were trade-offs between yield gap closure and labour productivity in Central Luzon (Chapter 4) and in Southern

[†]Revenues were calculated at the farm level as the sum of the product between individual crop yields and market prices for each crop (other sources of agricultural-related income were also included in the Netherlands). Production costs were computed at farm level considering variable costs (material inputs and hired labour) in the three case studies and additional fixed costs in the Netherlands (e.g. rented land, energy costs, depreciation costs, other financial costs). Their variation is presented in Figure 8.4

Ethiopia (Chapter 6). Conversely, no significant relationships between labour use at farm level and crop yields were observed in the Netherlands (Chapter 5). The quadratic relationship between cereal yield and labour productivity observed in the tropics indicates that, even if it would be possible to use more labour, it would not be

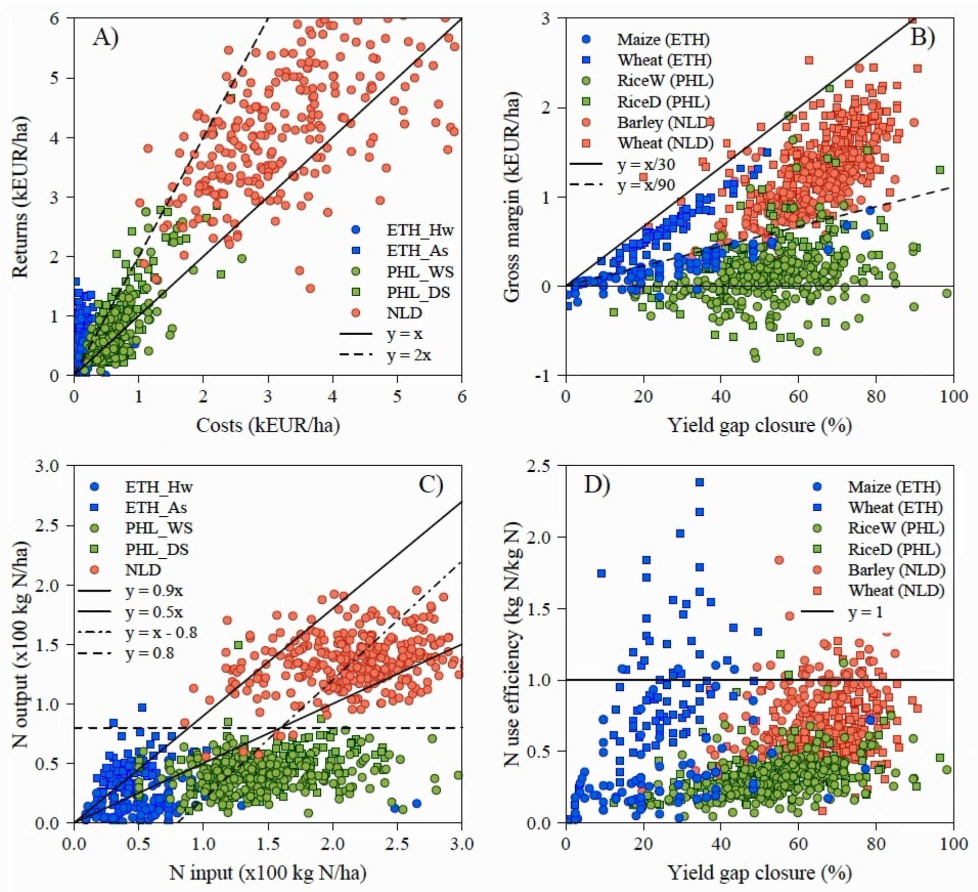

Figure 8.4: Economic and environmental performance of individual farms across the three case studies: A) returns and costs at the farm level, B) yield gap closure for cereals and gross margin at crop level, C) N outputs and N inputs at farm level and D) yield gap closure for cereals and N use efficiency at crop level. N use efficiency did not take into account soil mineralization (input) and crop residues (output) under the assumption both N flows correspond to internal recycling. Data for Ethiopia refer to the year 2012, for the Netherlands to the years 2010 - 2012 and for the Philippines to years after 2000 included in the Central Luzon Loop Survey. Background data to this figure is provided in Chapter 4, Chapter 5 and Chapter 6. Country codes: ETH = Ethiopia, PHL = Philippines, NLD = Netherlands.

worth for most farmers to do so because of lower returns to labour at high yield levels. Land productivity has been the focus of agricultural development in the past decades but these results indicate that labour productivity requires further attention in the years to come (Woodhouse, 2010), especially for smallholder farming systems in the tropics.

The indicator of N use efficiency (NUE) proposed by the EU Nitrogen Expert Panel (2015) was applied to each farming system. This framework states that NUE greater 0.9 kg N kg^{-1} N indicate the risk of soil mining while levels lower than 0.5 kg N kg^{-1} N point to inefficient use of N. This optimal range (50 - 90%) should be obtained with N output (i.e. farm productivity in kg N ha^{-1}) greater than 80 kg N ha^{-1} and N surplus (i.e. N input − N output) of less than 80 kg N ha^{-1}. Based on this framework, the following conclusions can be derived (Figure 8.4C):

1. There is evidence of mining soil N in Asella (Southern Ethiopa) as NUE greater than 0.9 kg N kg^{-1} N were observed for some farms. N output and N surplus are well below 80 kg N ha^{-1} for most farms in this site.

2. There is scope to increase N use efficiency in Hawassa (Southern Ethiopia) and Central Luzon (Philippines), where NUE is lower than 0.5 kg N kg^{-1} N for most farms. N output is lower than 80 kg N ha^{-1} for nearly all farms in these sites. N surplus is lower than 80 kg N ha^{-1} for most farms in Hawassa, while the opposite is true in Central Luzon.

3. Most arable farms in the Netherlands are N use efficient ($0.5 \leq$ NUE ≤ 0.9 kg N kg^{-1} N) and have a N output greater than 80 kg N ha^{-1}. However, N surplus is still well above 80 kg N ha^{-1} for many farms pointing at high N losses.

Observations at crop level support the statements above (Figure 8.4D). N use efficiency for wheat in Asella ranged between 0.5 - 1.5 kg N kg^{-1} N, which indicates soil N mining. N use efficiency was low for maize in Hawassa and rice in Central Luzon, typically between 0.3 - 0.5 kg N kg^{-1} N. The low NUE values are common in irrigated lowland rice systems due to the challenge of synchronizing N application with, a highly variable, indigenous soil N supply in flooded soils (Olk et al., 1998; Cassman et al., 1996). Finally, N use efficiency of arable farms in the Netherlands was highest with values between 0.4 - 1.0 kg N kg^{-1} N). Trade-offs between yield gap closure and optimal NUE at crop level were observed in Central Luzon where rice yields maximising NUE were ca. 68% of Yp (Chapter 4). Wheat yields of ca. 75% Yp maximised NUE in the Netherlands while this trade-off was less relevant in Southern Ethiopia, where mining of soil N prevails (Figure 8.4D)

Nutrient imbalances in contrasting farming systems were reported in other studies at local level (e.g. Carberry et al., 2013; Vitousek et al., 2009) and at global level (e.g. Zhang et al., 2015; Lassaletta et al., 2014). Studies conducted at global level are often not representative of farming systems in specific countries. For instance, Lassaletta et al. (2014) estimated for the year 2009 an N input of 22, 77 and 360 kg N ha^{-1} and NUE of 1.2, 0.6 and 0.4 kg N kg^{-1} N for Ethiopia, the Philippines and the Netherlands, respectively. These are similar to the results obtained for Southern Ethiopia, but not for Central Luzon where NUE was overestimated and N inputs underestimated. The opposite is true for Dutch arable farming where NUE above 50%, and N inputs lower 300 kg N ha^{-1}, were observed for most farms. Similar discrepancies were observed with the study of Zhang et al. (2015) conducted for the year 2011.

The uncertainty and resolution of the data used by global studies seem to provide unreliable estimates of NUE and N input at farming systems level. Similar problems were reported for the estimation of yield gaps in global studies (van Ittersum et al., 2013). In addition to the lack of accuracy, global studies hardly accommodate the variability and diversity observed at local level (Figures 8.2 and 8.3) and barely consider the socio-economic context where farming systems operate (Figure 8.4; Chapter 7). The type of analysis and data documented in this thesis are thus needed to gain deeper insights in the performance of specific farm(ing) systems and hence, to give sustainable intensification local meaning.

8.4 The sustainable intensification of agriculture

8.4.1 Where?

Food security is a pressing reality in Ethiopia, a country where increases in cereal yields will be required if cereal self-sufficiency is to be achieved by 2050 (van Ittersum et al., 2016). In the Philippines, (sustainable) intensification of rice production must be considered for similar reasons: the country has historically been the largest rice importer in the world (Dawe et al., 2006) and it is expected to remain that way if rice production does not increase (Laborte et al., 2012). Moreover, there is no strong evidence of e.g. groundwater pollution due to past intensification (Bouman et al., 2002). The situation is different in the Netherlands where reducing fertiliser use without compromising yields is possible for most crops (Chapter 5).

There is considerable scope to improve NUE in the different farming systems, but the way that can be best achieved depends on e.g. the yield gap closure currently realized in each farming system (Figures 8.1, 8.2 and 5.6C). Three main 'directions of change' pose themselves for the future: sustainable intensification ('more output with less input'), extensification ('same output with less input') and intensification ('more output with more input'). Sustainable intensification as well as intensification are appropriate in Hawassa and Central Luzon where NUE is low and the yield gap large. Intensification is needed in Asella where values of NUE greater than 1 kg N kg^{-1} N, and thus mining of soil N, and large yield gaps are observed for many farms. This contrasts with the extensification pathway currently observed for this site (Chapter 6). Finally, sustainable intensification, or even extensification, is desirable in the Netherlands where NUE is within the optimal range and yield gaps small, but N surplus is still high in many farms.

The insights gained in Southern Ethiopia and in the Netherlands can be extrapolated to other regions in East Africa (Tittonell and Giller, 2013) and Northwestern Europe (Stoate et al., 2009), respectively. However, that is not true in Southeast Asia due to large differences in NUE and yield gaps and across the major rice bowls (Stuart et al., 2016; Laborte et al., 2012). The socio-economic setting is also quite peculiar compared to other rice bowls in Southeast Asia (Moya et al., 2004; Kerkvliet, 1990).

8.4.2 How?

Intensification in Ethiopia will require increases in N and P application rates, as well as in their use efficiency, to narrow the resource and technology yield gaps (Chapter 6, Figure 8.2). This shall be accompanied by the use of other inputs (e.g. herbicides: Tanner et al., 1993), some of which are not yet readily available throughout the country (e.g. composite fertilisers with K), and knowledge intensive ecological principles to control pests, diseases and weeds where possible (e.g. Kebede et al., 2015; Taa et al., 2004). Moreover, releasing labour and capital constraints will be required to ensure more timely and precise crop management (to narrow the efficiency yield gap) and to intensify input use (Chapter 6), while reducing labour drudgery (Baudron et al., 2015).

In Central Luzon, (sustainable) intensification of rice farming requires adoption of precision agriculture practices to improve the timeliness of fertiliser and pesticide applications and narrow the efficiency yield gap (Chapter 3; Savary et al., 2012; Dobermann et al., 2002; Bouman et al., 2001). Narrowing the resource and technology yield gaps by increasing N application rates is not desirable in the wet season due to increased risk of lodging as a result of typhoons (Chapter 7; Lampayan et al., 2010) but should occur in the dry season when the climatic potential yield is higher

(Chapters 3 and 4). Attention should also be paid to K limitations (Chapter 3; Dobermann et al., 1996) and to the interaction between crop establishment method and weed pressure (Chapter 4; Lantican et al., 1999).

Sustainable intensification, or even extensification, of arable farming in the Netherlands requires increased precision of crop management in time and space to narrow the efficiency yield gap (Chapter 5). This may still be constrained by weather extremes influencing the date of sowing and harvesting (van Oort et al., 2012), problems with trafficability when soils are wet (Droogers et al., 1996), machinery constraints on large farms meaning that management operations are performed when machines are available and not necessarily when crops most need (Reidsma et al., 2015b) and/or maximisation of economic performance in the short term which favour as less management interventions as possible (Mandryk et al., 2014). In addition, it is also important to reduce the pressure of over-fertilisation by seeking alternative ways for disposing excessive organic manures.

8.4.3 Whom?

In Southern Ethiopia and Central Luzon, (sustainable) intensification is expected to occur in farming systems where the farm size is small, capital scarce and labour 'plenty' (Figure 8.3; Hazell et al., 2010) and where returns to land are as important as returns to labour (Chapters 4 and 6; Woodhouse, 2010). While large yield gaps, and low resource use efficiencies, suggest considerable room for improvement in these farming systems (Figures 8.1 and 8.4C), it is unclear whether that potential can be realized where farming is one in many livelihood activities and where lack of capital is still the rule rather than the exception.

The growing importance of *khat*, a highly profitable narcotic crop, in the farming systems around Hawassa (Mellisse et al., 2017) may be seen as a response of farmers' to the lack of capital available and the small farm sizes (Chapter 6). In a region where the population density is as high as 600 person km^2 and farm sizes are small, intensifying agriculture further requires solving liquidity problems through e.g. off-farm employment and, even then, it is unclear whether the additional income earned will be re-invested on-farm. However, the prospects for Asella are different given the lower population density (ca. 200 person km^2) and good market opportunities (e.g. breweries and increasing demand for small grains). Current returns to land and labour also indicate that farming can be a profitable activity in this site.

A historical diagnosis of rural livelihoods in Central Luzon indicates that: 1) labour productivity, hired labour and off-farm income increased sharply, 2) profitability stagnated or declined and, 3) current areas and crop yields are sufficient to meet rice demands at household level (Chapters 4 and 7; Moya et al., 2015; Takahashi and Otsuka, 2009; Estudillo and Otsuka, 1999). Despite being the main rice supplier to Metro Manila, this indicates that rice farming in Central Luzon is 'an additional activity' to rural households, possibly getting close to hobby farming, rather than a 'professional business'. In a country where rice production needs to double if rice self-sufficiency is to be achieved by 2050 (Laborte et al., 2012), it is unclear to which extent smallholder farmers in Central Luzon will be the ones ensuring such future rice demand with current technologies and under prevailing conditions.

The large 'cash flows' and 'economic size' of Dutch arable farms require professional managerial skills, especially in a country where land and labour costs are very high. In this context, it is no surprise that some farms struggle to be profitable (even when considering other sources of agricultural related income, Figure 8.4A), which in turn encompasses prioritization of short-term economic performance over long-term 'sustainability' (Mandryk et al., 2014). This 'quest for survival' can be observed in e.g. the diversification of traditional crop rotations with high value root and tuber crops. The cropping frequency of root and tuber crops had no effect on crop yields (which contrasts with previous experimental evidence, Scholte and s'Jacob, 1990), suggesting it is possible to reconcile high yields with frequent cultivation of high value crops (Chapter 5). However, further research is needed to verify these results at field scale.

8.5 Methodological considerations

8.5.1 Main contributions

The literature review conducted in Chapter 2 highlighted that the factors considered for yield gap analysis are subjected to e.g. the methods applied and that most studies have focused on the contribution of input quantity rather than on crop management in its broader sense. Moreover, labour, farm characteristics and socio-economic factors have been hardly considered in the agronomic yield gap literature (Chapter 2), but widely considered in agricultural economics technical efficiency literature (Thiam et al., 2001).

Overcoming the caveats identified in Chapter 2 required new concepts, multiple methods and detailed datasets. Integrating methods of frontier analysis with concepts of production ecology proved useful to decompose and explain crop yield gaps (Chapters 3, 5 and 6) as this allowed merging the flexibility of the frontier approach with process knowledge on crop growth. Moreover, this approach could be linked to crop models which are useful to identify biophysical yield ceilings and to benchmark the most productive technology observed in a given farming system against technologies with improved resource use efficiency (van Ittersum et al., 2013). Additional statistical methods had to be deployed to further integrate yield gaps within farm level conditions but this was highly dependent on the type, and detail, of data available. The frameworks and methods used in this thesis are generic and provide a comprehensive ’toolbox’ for future studies on yield gaps and resource use efficiency.

The datasets compiled were the most comprehensive, in terms of detail or temporal and spatial span, available for each case study and their quality is certified up to a certain extent. First, the instrument used in Southern Ethiopia was designed by anthropologists and tested prior to its administration to farmers. Second, the Central Luzon Loop Survey is a unique long-term effort in which the same farms were visited multiple times by the same researchers. In addition, there was a recent strategic investment towards the harmonization of the dataset (Moya et al., 2015). Third, the Dutch farms included in the database gave voluntarily consent to the use of their private data for official monitoring and evaluation purposes. Certainly, the datasets contain errors and inaccuracies but these may be reduced due to the peculiarities highlighted.

8.5.2 Main limitations

The use of frontier analysis to explain yield gaps has three main limitations. First, benchmark yields (e.g. Y_{TEx}) depend largely on the data available in the household surveys related to e.g. soil properties and crop management (Chapter 2). The lack of detailed soil information in most household surveys may confound the estimates of the efficiency and resource yield gaps and it is important to understand how large the errors incurred may be. For instance, the effect of soil properties on the efficiency yield gap of rice in Central Luzon was minor (Chapter 3) but that is unlikely to be the case in regions where soil heterogeneity is more pronounced (Tittonell and Giller, 2013). Other example refers to regions where rainfall is highly erratic and unequally distributed within small spatial units (e.g. Kassie et al., 2014; Hochman et al., 2012). Thus, knowledge about the local conditions is needed to judge whether the data available in these regions is appropriate for this type of analysis so that technical efficient farms are classified as such because they are more efficient in converting inputs to outputs and not because they experienced more favourable growing conditions.

Second, it is difficult to derive policy recommendations using household surveys in farming systems where the technology yield gap is large (which can of course still be done with e.g. expert knowledge). This may well be the case for most farming systems in sub-Saharan Africa (Chapter 6; van Dijk et al., 2017; Tittonell and Giller, 2013). Detailed information on the input combinations (e.g. water and nutrients) and management practices (e.g. timing and method of input application) required to attain the yield ceilings would be useful to explain the technology yield gap further. Moreover, variability in yield ceilings, particularly Yw, needs to be considered more often and in addition to variability in actual yields (Chapter 7). For example, actual yields in a large potato farm in the Netherlands ranged between 20 and 90 t ha^{-1}, following a normal distribution, and the same might hold for Yw due to differences in soil properties, depth of groundwater table and past history of different fields (Reidsma et al., 2015b).

The last limitation is the accuracy of farmers reported estimates of land areas, and crop yield, in household surveys. This is a well-known problem, particularly in sub-Saharan Africa (Carletto et al., 2015), and has to do with farmers' lack of precise knowledge or unwillingness to report the right information. Improving the accuracy of this type of data in 'developing countries' requires focused questionnaires and on-farm measurements over multiple visits to the same farm, but this would reduce sample sizes due to increased costs of data collection. Data quality is somewhat less problematic in 'developed countries' where local authorities compile and harmonize farm databases to inform official statistics (Paustian, 2013). The problem there refers to data ownership and accessibility due to privacy reasons. Moreover, the degree of detail is often not ideal for in-depth agronomic analysis (e.g. crop rotations effects in Chapter 5). Data collection approaches based on e.g. crowdsourcing will be useful to address part of the 'data problem' in the future (Chapter 2; Beza, 2017).

8.5.3 Future applications

This thesis focused on benchmarking different farms within specific regions (Chapters 3, 5 and 6, Figure 8.5). This allowed to decompose yield gaps at lower systems levels and to use those as entry points for further explanations at higher systems levels. However, a different approach could have been followed. For instance, stochastic frontier analysis could have been used to estimate 'system-wide yield gaps' at the farm level first (based on distance functions; Henderson et al., 2016) followed by analysis of an in-depth analysis at crop level to explain the variation observed in farm performance. Crop models could have also been used to assess the variability of Yw for different farms within a region (especially in Chapters 5 and 6), similarly to previous approaches benchmarking different fields within a farm (Oliver and Robertson, 2013). The application of these methods may be further extended in the future to sensor data collected for precision farming with the objective of benchmarking different 'soil management units' within single fields.

Further efforts are required to explain yield gaps vis-à-vis resource use efficiency gaps (van Noordwijk and Brussaard, 2014). The yield responses to nutrients analysed in this thesis fall in 'quadrant II' (Figure 8.2) of the 'three-quadrant diagram' used to decompose nutrient use efficiency (de Wit, 1992). However, the concepts and methods developed can be used to benchmark uptake and conversion efficiencies as well, and to identify their drivers, if detailed on-farm data on e.g. nutrient uptake by crops are available. Existing datasets from past research can be used to develop a

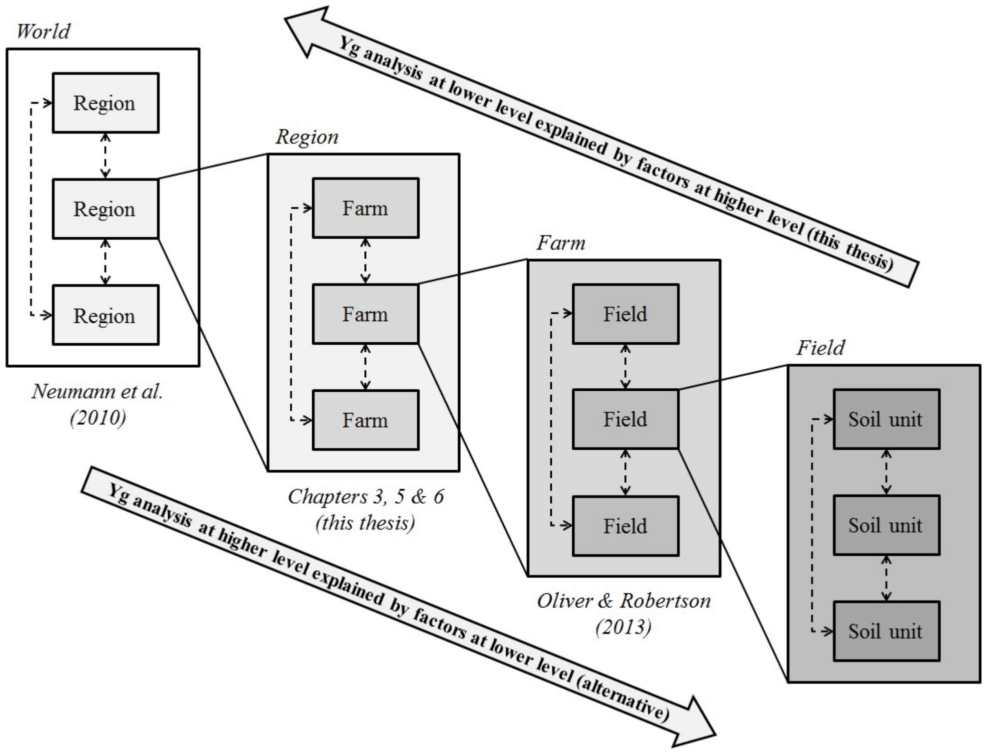

Figure 8.5: Opportunities for yield gap (Yg) analysis at different scales. As different scales are nested, it is possible to explain yield gaps at lower systems levels and seek for explanations at higher systems levels (this thesis) or *vice versa*.

prototype for this type of analysis (e.g. Sattari et al., 2014).

Optimisation techniques can complement the current analyses in at least two different ways. First, they can be used to identify input-output combinations in the production frontier which optimise a specific farmer objective given a set of constraints (e.g. Kanellopoulos et al., 2014). This would complement the framework introduced in Chapter 3 by extending the quantification of the resource yield gap to different farmers' objectives. Second, they can also be used to identify the optimal yield level that can be realized in a specific farm under prevailing conditions of input prices, climatic and price variability, labour requirements, opportunity costs for labour off-farm and resource availability. This acknowledges the fact that for some households it may be more interesting to target yield levels lower than the modelled Yp or Yw, which in some cases are even unrepresentative of current farmers' practices (Chapters 6 and 7). This would complement the current yield gap analysis by shifting the focus from what is 'biophysically possible' towards what is 'socio-economically feasible'.

8.6 Conclusions

Yield gaps have been traditionally estimated and explained using crop models. Although these incorporate detailed process knowledge about crop growth, they are ill suited to embed yield gaps within the broader farm level conditions (Chapter 2). To overcome these caveats, a method combining frontier analysis and concepts of production ecology was developed to decompose crop yield gaps into efficiency, resource and technology yield gaps (Chapter 3). This was successfully tested in rice-based farming systems in Central Luzon, Philippines (Chapter 3), arable farming systems in the Netherlands (Chapter 5) and mixed-crop livestock systems in Southern Ethiopia (Chapter 6). In short, the efficiency yield gap was most important in the Netherlands, the technology yield gap was most important in Southern Ethiopia and the three intermediate yield gaps were of similar importance in Central Luzon.

The farm level determinants of yield gaps were clearer for farming systems in Southern Ethiopia than for rice farming in Central Luzon and arable farming in the Netherlands. For example, there is evidence of intensification in the farming systems around Hawassa, where farm sizes are very small, as capital constraints are alleviated and evidence of extensification in Asella as oxen ownership increases (Chapter 6). The labour peaks for management operations of cereals and pulses and limited supply of human labour and animal draught power are still possible constraints to yield gap closure in Asella. As to Central Luzon, it is possible to conclude that farm and regional factors did not lead to different levels of intensification within the variation investigated (Chapter 4). The most striking effect observed was that direct-seeding (and thus slightly lower rice yields) was mostly adopted in larger farms, and used lower amounts of hired labour than transplanting. The analysis of rotational effects in Dutch arable farms yielded inconclusive results but confounding effects with e.g. rented land do not allow to conclude that these are not at stake in this farming system (Chapter 5).

Sustainable intensification requires new sources of knowledge and methods as well as more transparency and better analysis of trade-offs (Struik et al., 2014). This thesis shows that indeed a thorough understanding of the causes of yield gaps at farm(ing) systems level is a knowledge, and data intensive exercise, and it requires empirical and modelling approaches. For the farming systems analysed (Chapters 4, 5 and 6), yield gap closure does not seem to entail trade-offs with gross margin per unit land. However, it does so with N use efficiency and labour productivity in Southern Ethiopia and Central Luzon, and to a less extent in the Netherlands. This means that (sustainable) intensification needs to go hand-in-hand with agronomic interventions that increase land productivity while ensuring high resource use efficiency and with labour-saving technologies that can reduce the drudgery of farming without compromising crop yields. Future yield gap studies should not only assess the contribution of growth-defining, -limiting and -reducing factors to crop yields but also integrate those within the diversity of responses observed at local level and the broader livelihood context within which farming takes place.

This thesis broadens the discussion on yield gaps by moving from the technical aspects underlying their estimation towards the broader farm level opportunities and constraints undermining their closure. In doing so, it shows that agriculture is a source of staple foods, income and employment rather than a mere source of emissions and environmental degradation. Insights from contrasting farming systems indicate that further intensification of input use and crop management is required in the 'developing South', where yield gaps are large and resource use efficiency low, while sustainable intensification (or even extensification) needs to be targeted in the 'developed North', where yield gaps are small and resource use efficiency high. In either case, this will need to be achieved by individual farms who are far too often required to prioritise short-term needs over long-term aspirations.

References

Abate, T., Shiferaw, B., Menkir, A., Wegary, D., Kebede, Y., Tesfaye, K., Kassie, M., Bogale, G., Tadesse, B., and Keno, T. (2015). Factors that transformed maize productivity in Ethiopia. *Food Security*, 7:965 – 981.

Abeledo, L. G., Savin, R., and Slafer, G. A. (2008). Wheat productivity in the Mediterranean Ebro Valley: Analyzing the gap between attainable and potential yield with a simulation model. *European Journal of Agronomy*, 28:541 – 550.

Affholder, F., Poeydebat, C., Corbeels, M., Scopel, E., and Tittonell, P. (2012). The yield gap of major food crops in family agriculture in the tropics: Assessment and analysis through field surveys and modelling. *Field Crops Research*, 143:106 – 118.

Aggarwal, P. K. and Kalra, N. (1994). Analyzing the limitations set by climatic factors, genotype, and water and nitrogen availability on productivity of wheat II. Climatically potential yields and management strategies. *Field Crops Research*, 38:93 – 103.

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

Aker, J. C. (2011). Dial ”A” for agriculture: a review of information and communication technologies for agricultural extension in developing countries. *Agricultural Economics*, 42:631 – 647.

Ali, M. (1995). Institutional and socioeconomic constraints on the second-generation Green Revolution: A case study of basmati rice production in Pakistan’s Punjab. *Economic Development and Cultural Change*, 43:835 – 861.

Allan, A. Y. (1971). *The influence of agronomic factors on maize yields in western Kenya with special reference to time of planting*. PhD thesis, Nairobi: University of East Africa.

Amare, T. (2014). Effect of weed management methods on weeds and wheat (Triticum aestivum L.) yield. *African Journal of Agricultural Research*, 9:1914 – 1920.

An, N., Fan, M., Zhang, F., Christie, P., Yang, J., Huang, J., Guo, S., Shi, X., Tang, Q., Peng, J., Zhong, X., Sun, Y., Lv, S., Jiang, R., and Dobermann, A. (2015). Exploiting co-benefits of increased rice production and reduced greenhouse gas emission through optimized crop and soil management. *PLOS ONE*, 10:1 – 17.

Anderson, W. K., Hamza, M. A., Sharma, D. L., DÁntuono, M. F., Hoyle, F. C., Hill, N., Shackley, B. J., E, M. A., and Zaicou-Kunesch, C. (2005). The role of management in yield improvement of the wheat crop - a review with special emphasis on Western Australia. *Australian Journal of Agricultural Research*, 56:1137 – 1149.

Andert, S., Burger, J., Stein, S., and Gerowitt, B. (2016). The influence of crop sequence on fungicide and herbicide use intensities in North German arable farming. *European Journal of Agronomy*, 77:81 – 89.

Angulo, C., Becker, M., and Wassmann, R. (2012). Yield gap analysis and assessment of climate-induced yield trends of irrigated rice in selected provinces of the Philippines. *Journal of Agriculture and Rural Development in the Tropics and Subtropics*, 113:61 – 68.

Audebert, A. and Fofana, M. (2009). Rice yield gap due to iron toxicity in West Africa. *Journal of Agronomy and Crop Science*, 195:66 – 76.

Aune, J. B., Bussa, M. T., Asfaw, F. G., and Ayele, A. A. (2001). The ox-ploughing system in Ethiopia: Can it be sustained? *Outlook on Agriculture*, 30:275 – 280.

Bacigaluppo, S., Bodrero, M. L., Balzarini, M., Gerster, G. R., Andriani, J. M., Enrico, J. M., and Dardanelli, J. L. (2011). Main edaphic and climatic variables explaining soybean yield in Argiudolls under no-tilled systems. *European Journal of Agronomy*, 35:247 – 254.

Bai, J., Li, J., and Li, S. (2011). *Monitoring the Plant Density of Cotton with Remotely Sensed Data*, pages 90 – 101. Springer Berlin Heidelberg, Berlin, Heidelberg.

Barker, R. and Levine, G. (2012). Water productivity in context: The experiences of Taiwan and the Philippines over the past half-century. Technical report, International Water Management Institute (IWMI).

Battese, G. E. and Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20:325 – 332.

Baudron, F. (2011). *Agricultural Intensification - Saving space for wildlife?* PhD thesis, Wageningen University.

Baudron, F., Andersson, J. A., Corbeels, M., and Giller, K. E. (2012). Failing to yield? Ploughs, conservation agriculture and the problem of agricultural intensification: An example from the Zambezi Valley, Zimbabwe. *Journal of Development Studies*, 48:393 – 412.

Baudron, F., Jaleta, M., Okitoi, O., and Tegegn, A. (2014). Conservation agriculture in African mixed crop-livestock systems: Expanding the niche. *Agriculture, Ecosystems & Environment*, 187:171 – 182.

Baudron, F., Sims, B., Justice, S., Kahan, D. G., Rose, R., Mkomwa, S., Kaumbutho, P., Sariah, J., Nazare, R., Moges, G., and Gérard, B. (2015). Re-examining appropriate mechanization in Eastern and Southern Africa: two-wheel tractors, conservation agriculture, and private sector involvement. *Food Security*, 7:889 – 904.

Becker, M. and Johnson, D. E. (1999). Rice yield and productivity gaps in irrigated systems of the forest zone of Côte d'Ivoire. *Field Crops Research*, 60:201 – 208.

Becker, M., Johnson, D. E., Wopereis, M. C. S., and Sow, A. (2003). Rice yield gaps in irrigated systems along an agro-ecological gradient in West Africa. *Journal of Plant Nutrition and Soil Science*, 166:61 – 67.

Beddow, J. M., Hurley, T. M., Pardey, P. G., and Alston, J. (2015). Rethinking yield gaps. Technical report, University of Minnesota, Department of Applied Economics.

Beddow, J. M., Hurley, T. M., Pardey, P. G., and Alston, J. M. (2014). Food security: Yield gap. In *Encyclopedia of Agriculture and Food Systems*, pages 352 – 365. Academic Press.

Belden, C., Bothwell, C., Etulain, T. M., Figueres, C. M. D. N., Goyal, A., Kelly, T. J. C., Pidatala, K., Pruuden, P., and Surya, P. (2013). ICT for data collection and monitoring and evaluation: opportunities and guidance on mobile applications for forest and agricultural sectors. Technical report, The World Bank.

Beza, E. (2017). *Citizen science and remote sensing for crop yield gap analysis*. PhD thesis, Wageningen University.

Beza, E., Silva, J. V., Kooistra, L., and Reidsma, P. (2017). Review of yield gap explaining factors and opportunities for alternative data collection approaches. *European Journal of Agronomy*, 82, Part B:206 – 222.

Bezu, S. and Holden, S. (2014). Are rural youth in Ethiopia abandoning agriculture? *World Development*, 64:259 – 272.

Bhatia, V. S., Singh, P., Wani, S. P., Chauhan, G. S., Rao, A. V. R. K., Mishra, A. K., and Srinivas, K. (2008). Analysis of potential yields and yield gaps of rainfed soybean in India using CROPGRO-Soybean model. *Agricultural and Forest Meteorology*, 148:1252 – 1265.

BMGF (2010). Agricultural development outcome indicators: Initiative and sub-initiative progress indicators and pyramid of outcome indicators. Technical report, Bill and Melinda Gates Foundation, Seattle.

Boatman, N., Stoate, C., Gooch, R., Carvalho, C. R., Borralho, R., de Snoo, G., and Eden, P. (1999). The environmental impact of arable crop production in the European Union: Practical options for improvement. Technical report, European Commission.

Boling, A. A., Bouman, B. A. M., Tuong, T. P., Konboon, Y., and Harnpichitvitaya, D. (2011). Yield gap analysis and the effect of nitrogen and water on photoperiod-sensitive Jasmine rice in north-east Thailand. *NJAS - Wageningen Journal of Life Sciences*, 58:11 – 19.

Boling, A. A., Tuong, T. P., van Keulen, H., Bouman, B. A. M., Suganda, H., and Spiertz, J. H. J. (2010). Yield gap of rainfed rice in farmers' fields in Central Java, Indonesia. *Agricultural Systems*, 103:307 – 315.

Bommarco, R., Kleijn, D., and Potts, S. G. (2013). Ecological intensification: Harnessing ecosystem services for food security. *Trends in Ecology & Evolution*, 28:230 – 238.

Boogaard, H., Wolf, J., Supit, I., Niemeyer, S., and van Ittersum, M. K. (2013). A regional implementation of WOFOST for calculating yield gaps of autumn-sown wheat across the European Union. *Field Crops Research*, 143:130 – 142.

Bos, J. F. F. P., Smit, A. B. L., and Schröder, J. J. (2013). Is agricultural intensification in the Netherlands running up to its limits? *NJAS - Wageningen Journal of Life Sciences*, 66:65 – 73.

Bos, J. F. F. P., ten Berge, H. F. M., Verhagen, J., and van Ittersum, M. K. (2016). Trade-offs in soil fertility management on arable farms. *Agricultural Systems*, 157:292 – 302.

Bouman, B. A. M., Castaneda, A. R., and Bhuiyan, S. I. (2002). Nitrate and pesticide contamination of groundwater under rice-based cropping systems: past and current evidence from the Philippines. *Agriculture, Ecosystems & Environment*, 92:185 – 199.

Bouman, B. A. M., Kropff, M. J., Tuong, T. P., Wopereis, M. C. S., ten Berge, H. F. M., and van Laar, H. H., editors (2001). *ORYZA2000: Modelling lowland rice*. Los Baños (Philippines): International Rice Research Institute, and Wageningen: Wageningen University and Research Centre.

Bouman, B. A. M. and van Laar, H. H. (2006). Description and evaluation of the rice growth model ORYZA2000 under nitrogen-limited conditions. *Agricultural Systems*, 87:249 – 273.

Bouwman, L., Goldewijk, K. K., van Der Hoek, K. W., Beusen, A. H. W., van Vuuren, D. P., Willems, J., Rufino, M. C., and Stehfest, E. (2011). Exploring global changes in nitrogen and phosphorus cycles in agriculture induced by livestock production over the 1900-2050 period. *Proceedings of the National Academy of Sciences*, 110:20882 – 20887.

Bravo-Ureta, B. E. and Pinheiro, A. (1993). Efficiency analysis of developing country agriculture: A review of the frontier function literature. *Agricultural and Resource Economics Review*, 22:88 – 101.

Brooks, S., Thompson, J., Odame, H., Kibaara, B., Nderitu, S., Karin, F., and Millstone, E. (2009). Environmental change and maize innovation in Kenya: Exploring pathways in and out of maize. Technical report, STEPS Working Paper, 36. Brighton: STEPS Centre.

Bunei, E., Rono, J., and Chessa, S. (2013). Factors influencing farm crimes in Kenya: Opinions and experiences from farmers. *International Journal of Rural Criminology*, 2:75 – 100.

Cabling, J. M. and Dawe, D. C. (2006). *Why does the Philippines Import rice? Meeting the challenge of trade liberalization*, chapter Filipino farmers receive high palay prices, pages 9 – 11. International Rice Research Institute (IRRI).

Calvino, P. and Sadras, V. (2002). On-farm assessment of constraints to wheat yield in the south-eastern Pampas. *Field Crops Research*, 74:1 – 11.

Cao, X., Luo, Y., Zhou, Y., Duan, X., and Cheng, D. (2013). Detection of powdery mildew in two winter wheat cultivars using canopy hyperspectral reflectance. *Crop Protection*, 45:124 – 131.

Carberry, P. S., Liang, W.-l., Twomlow, S., Holzworth, D. P., Dimes, J. P., McClelland, T., Huth, N. I., Chen, F., Hochman, Z., and Keating, B. A. (2013). Scope for improved eco-efficiency varies among diverse cropping systems. *Proceedings of the National Academy of Sciences*, 110:8381 – 8386.

Carletto, C., Gourlay, S., and Winters, P. (2015). From Guesstimates to GPStimates: Land area measurement and implications for agricultural analysis. *Journal of African Economies*, 24:593.

Casanova, D., Goudriaan, J., Bouma, J., and Epema, G. F. (1999). Yield gap analysis in relation to soil properties in direct-seeded flooded rice. *Geoderma*, 91:191 – 216.

Cassman, K. G. (1999). Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture. *Proceedings of the National Academy of Sciences*, 96:5952 – 5959.

Cassman, K. G., De Datta, S. K., Olk, D. C., Alcantara, J., Samson, M., Descalsota, J., and Dizon, M. (1995). Yield decline and the nitrogen economy of long-term experiments on continuous, irrigated rice systems in the tropics. *Soil management: Experimental basis for sustainability and environmental quality*, pages 181 – 222.

Cassman, K. G., Dobermann, A., Walters, D. T., and Yang, H. (2003). Meeting cereal demand while protecting natural resources and improving environmental quality. *Agronomy & Horticulture – Faculty Publications*, 318.

Cassman, K. G., Gines, G. C., Dizon, M. A., Samson, M. I., and Alcantara, J. M. (1996). Nitrogen-use efficiency in tropical lowland rice systems: contributions from indigenous and applied nitrogen. *Field Crops Research*, 47:1 – 12.

Cassman, K. G., Peng, S., Olk, D. C., Ladha, J. K., Reichardt, W., Dobermann, A., and Singh, U. (1998). Opportunities for increased nitrogen-use efficiency from improved resource management in irrigated rice systems. *Field Crops Research*, 56:7 – 39.

Cassman, K. G. and Pingali, P. L. (1995). Intensification of irrigated rice systems: Learning from the past to meet future challenges. *GeoJournal*, 35:299 – 305.

CBS (2015). Centraal bureau voor de statistiek. Available at: http://www.cbs.nl/en-GB/menu/home/default.htm (accessed 05.11.2015).

Chadwick, D., Wei, J., Yan'an, T., Guanghui, Y., Qirong, S., and Qing, C. (2015). Improving manure nutrient management towards sustainable agricultural intensification in China. *Agriculture, Ecosystems & Environment*, 209:34 – 46.

Cobo, J. G., Dercon, G., and Cadisch, G. (2010). Nutrient balances in African land use systems across different spatial scales: A review of approaches, challenges and progress. *Agriculture, Ecosystems & Environment*, 136(1):1 – 15.

Coelli, T. J. and Henningsen, A. (2013). *frontier: Stochastic Frontier Analysis*. R package version 1.1-0.

Confalonieri, R., Foi, M., Casa, R., Aquaro, S., Tona, E., Peterle, M., Boldini, A., Carli, G. D., Ferrari, A., Finotto, G., Guarneri, T., Manzoni, V., Movedi, E., Nisoli, A., Paleari, L., Radici, I., Suardi, M., Veronesi, D., Bregaglio, S., Cappelli, G., Chiodini, M. E., Dominoni, P., Francone, C., Frasso, N., Stella, T., and Acutis, M. (2013). Development of an app for estimating leaf area index using a smartphone. Trueness and precision determination and comparison with other indirect methods. *Computers and Electronics in Agriculture*, 96:67 – 74.

Connor, D. J., Loomis, R. S., and Cassman, K. G. (2011). *Crop Ecology. Productivity and management in agricultural systems*. Cambridge University Press.

Cordell, D., Drangert, J.-O., and White, S. (2009). The story of phosphorus: Global food security and food for thought. *Global Environmental Change*, 19:292 – 305.

CSA and WB (2013). Ethiopia Rural Socioeconomic Survey (ERSS) - Survey report. Technical report, Central Statistical Agency and The World Bank.

Dawe, D. (2006). *Why does the Philippines Import rice? Meeting the challenge of trade liberalization*, chapter The Philippines imports rice because it is an island nation. International Rice Research Institute (IRRI).

Dawe, D., Moya, P. F., and Casiwan, C. B. (2006). *Why does the Philippines Import rice? Meeting the challenge of trade liberalization*. International Rice Research Institute (IRRI).

de Bie, C. A. J. M. (2004). The yield gap of mango in Phrao, Thailand, as investigated through comparative performance evaluation. *Scientia Horticulturae*, 102:37 – 52.

de Datta, S. K. (1986). Technology development and the spread of direct-seeded flooded rice in Southeast Asia. *Experimental Agriculture*, 22:417 – 426.

de Haan, J. J. and van Geel, W. (2013). Adviesbasis voor de bemesting van akkerbouw- en vollegrondsgroentengewassen (in dutch). Technical report, Plant Research International, Wageningen University.

de Koeijer, T. J., Wossink, G. A. A., Struik, P. C., and Renkema, J. A. (2002). Measuring agricultural sustainability in terms of efficiency: the case of Dutch sugar beet growers. *Journal of Environmental Management*, 66:9 – 17.

de Koeijer, T. J., Wossink, G. A. A., van Ittersum, M. K., Struik, P. C., and Renkema, J. A. (1999). A conceptual model for analysing input - output coefficients in arable farming systems: From diagnosis towards design. *Agricultural Systems*, 61:33 – 44.

de Mendiburu, F. (2015). *agricolae: Statistical Procedures for Agricultural Research*. R package version 1.2-3.

de Ponti, T., Rijk, B., and van Ittersum, M. K. (2012). The crop yield gap between organic and conventional agriculture. *Agricultural Systems*, 108:1 – 9.

de Wit, C. T. (1975). Substitution of labour and energy in agriculture and options for growth. *Netherlands Journal of Agricultural Sciences*, 23:145 – 162.

de Wit, C. T. (1979). The efficient use of labour, land and energy in agriculture. *Agricultural Systems*, 4:279 – 287.

de Wit, C. T. (1992). Resource use efficiency in agriculture. *Agricultural Systems*, 40:125 – 151.

Delmotte, S., Tittonell, P., Mouret, J.-C., Hammond, R., and Lopez-Ridaura, S. (2011). On farm assessment of rice yield variability and productivity gaps between organic and conventional cropping systems under Mediterranean climate. *European Journal of Agronomy*, 35:223 – 236.

Desbiez, A., Matthews, R., Tripathi, B., and Ellis-Jones, J. (2004). Perceptions and assessment of soil fertility by farmers in the mid-hills of Nepal. *Agriculture, Ecosystems & Environment*, 103:191 – 206.

DeTar, W. R. and Penner, J. V. (2007). Airborne remote sensing used to estimate percent canopy cover and to extract canopy temperature from scene temperature in cotton. *Transactions of the ASABE*, 50:495 – 506.

Dewitte, O., Jones, A., Spaargaren, O., Breuning-Madsen, H., Brossard, M., Dampha, A., Deckers, J., Gallali, T., Hallett, S., Jones, R., Kilasara, M., Roux, P. L., Micheli, E., Montanarella, L., Thiombiano, L., Ranst, E. V., Yemefack, M., and Zougmore, R. (2013). Harmonisation of the soil map of Africa at the continental scale. *Geoderma*, 211 - 212:138 – 153.

Dobermann, A. (2000). Future intensification of irrigated rice systems. In J. E. Sheehy, P. L. M. and Hardy, B., editors, *Redesigning Rice Photosynthesis to Increase Yield, Proceedings of the Workshop on the Quest to Reduce Hunger: Redesigning Rice Photosynthesis*, volume 7 of *Studies in Plant Science*, pages 229 – 247. Elsevier.

Dobermann, A. (2005). Nitrogen use efficiency - State of the art. *University of Nebraska - Lincoln*.

Dobermann, A., Cruz, P. C. S., and Cassman, K. G. (1996). Fertilizer inputs, nutrient balance, and soil nutrient-supplying power in intensive, irrigated rice systems. I. Potassium uptake and K balance. *Nutrient Cycling in Agroecosystems*, 46:1 – 10.

Dobermann, A., Dawe, D., Rötter, R. P., and Cassman, K. G. (2000). Reversal of rice yield decline in a long-term continuous cropping experiment. *Agronomy Journal*, 92:633 – 643.

Dobermann, A. and Witt, C. (2004). The evolution of site-specific nutrient management in irrigated rice systems of Asia. In Dobermann, A. and Witt, C., editors, *Increasing productivity of intensive rice systems through site-specific nutrient management*. Enfield, N.H. (USA) and Los Baños (Philippines): Science Publishers, Inc., and International Rice Research Institute (IRRI).

Dobermann, A., Witt, C., Dawe, D., Abdulrachman, S., Gines, H. C., Nagarajan, R., Satawathananont, S., Son, T. T., Tan, P. S., Wang, G. H., Chien, N. V., Thoa, V. T. K., Phung, C. V., Stalin, P., Muthukrishnan, P., Ravi, V., Babu, M., Chatuporn, S., Sookthongsa, J., Sun, Q., Fu, R., Simbahan, G. C., and Adviento, M. A. A. (2002). Site-specific nutrient management for intensive rice cropping systems in Asia. *Field Crops Research*, 74:37 – 66.

Dogliotti, S., Rossing, W. A. H., and van Ittersum, M. K. (2003). ROTAT, a tool for systematically generating crop rotations. *European Journal of Agronomy*, 19:239 – 250.

Donald, P. F., Green, R. E., and Heath, M. F. (2011). Agricultural intensification and the collapse of Europe's farmfarm bird populations. *Proceedings of the Royal Society of London*, 268:25 – 29.

Dorward, A. (2009). Integrating contested aspirations, processes and policy: Development as hanging in, stepping up and stepping out. *Development Policy Review*, 27:131 – 146.

Droogers, P., Fermont, A., and Bouma, J. (1996). Effects of ecological soil management on workability and trafficability of a loamy soil in the Netherlands. *Geoderma*, 73:131 – 145.

Droogers, P., Immerzeel, W. W., and Lorite, I. J. (2010). Estimating actual irrigation application by remotely sensed evapotranspiration observations. *Agricultural Water Management*, 97:1351 – 1359.

Dzanku, F. M., Jirstrom, M., and Marstorp, H. (2015). Yield gap-based poverty gaps in rural sub-Saharan Africa. *World Development*, 67:336 – 362.

Eberhart, S. A. (1971). Developing improved maize hybrids with a comprehensive breeding system. In *Proceedings of the 20th Annual National Breeders Roundtable (USA)*.

Ellis, F. (1993). *Peasant Economics: Farm households and agrarian development*. Cambridge University Press.

Emvalomatis, G., Stefanou, S. E., and Lansink, A. O. (2011). A reduced-form model for dynamic efficiency measurement: Application to dairy farms in Germany and the Netherlands. *American Journal of Agricultural Economics*, 93:161 – 174.

Erenstein, O. (2006). Intensification or extensification? Factors affecting technology use in peri-urban lowlands along an agro-ecological gradient in West Africa. *Agricultural Systems*, 90:132 – 158.

Erguiza, A., Duff, B., and Khan, C. (1990). Choice of rice crop establishment technique: Transplating vs wet seeding. *IRRI Research Paper Series*, 139:3 – 10.

Estudillo, J. P. and Otsuka, K. (1999). Green Revolution, human capital, and off-farm employment: Changing sources of income among farm households in Central Luzon, 1966 - 1994. *Economic Development and Cultural Change*, 47:497 – 523.

Estudillo, J. P. and Otsuka, K. (2006). Lessons from three decades of Green Revolution in the Philippines. *The Developing Economies*, 44:123 – 148.

EU Nitrogen Expert Panel (2015). Nitrogen Use Efficiency (NUE) - an indicator for the utilization of nitrogen in agriculture and food systems. Technical report, Alterra, Wageningen University.

Ewert, F., van Ittersum, M. K., Heckelei, T., Therond, O., Bezlepkina, I., and Andersen, E. (2011). Scale changes and model linking methods for integrated assessment of agri-environmental systems. *Agriculture, Ecosystems & Environment*, 142:6 – 17.

Falconnier, G. N., Descheemaeker, K., van Mourik, T. A., Sanogo, O. M., and Giller, K. E. (2015). Understanding farm trajectories and development pathways: Two decades of change in southern Mali. *Agricultural Systems*, 139:210 – 222.

FAO (2014). Fao Statistical Yearbook. Africa food and agriculture. Technical report, Food and Agriculture Organization of the United Nations.

FAO/FPMA (2017). Food Price Monitoring and Analysis (FPMA): http://www.fao.org/giews/pricetool/.

FAOSTAT (2017). Food and Agriculture Data: http://www.fao.org/faostat/en/.

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, 120:253 – 290.

Féon, V. L., Schermann-Legionnet, A., Delettre, Y., Aviron, S., Billeter, R., Bugter, R., Hendrickx, F., and Burel, F. (2010). Intensification of agriculture, landscape composition and wild bee communities: A large scale study in four European countries. *Agriculture, Ecosystems & Environment*, 137:143 – 150.

Fermont, A. M., van Asten, P. J. A., Tittonell, P., van Wijk, M. T., and Giller, K. E. (2009). Closing the cassava yield gap: An analysis from smallholder farms in East Africa. *Field Crops Research*, 112:24 – 36.

Ferster, C. J. and Coops, N. C. (2013). A review of earth observation using mobile personal communication devices. *Computers & Geosciences*, 51:339 – 349.

Fletcher, A. L., Brown, H. E., Johnstone, P. R., de Ruiter, J. M., and Zyskowski, R. F. (2011). Making sense of yield trade-offs in a crop sequence: A New Zealand case study. *Field Crops Research*, 124:149 – 156.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., Chapin, F. S., Coe, M. T., Daily, G. C., Gibbs, H. K., Helkowski, J. H., Holloway, T., Howard, E. A., Kucharik, C. J., Monfreda, C., Patz, J. A., Prentice, I. C., Ramankutty, N., and Snyder, P. K. (2005). Global consequences of land use. *Science*, 309:570 – 574.

Frelat, R., Lopez-Ridaura, S., Giller, K. E., Herrero, M., Douxchamps, S., Djurfeldt, A. A., Erenstein, O., Henderson, B., Kassie, M., Paul, B. K., Rigolot, C., Ritzema,

R. S., Rodriguez, D., van Asten, P. J. A., and van Wijk, M. T. (2016). Drivers of household food availability in sub-Saharan Africa based on big data from small farms. *Proceedings of the National Academy of Sciences*, 113:458 – 463.

French, R. J. and Schultz, J. E. (1984). Water use efficiency of wheat in a Mediterranean-type environment. I. The relation between yield, water use and climate. *Australian Journal of Agricultural Research*, 35:743 – 764.

Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., Duerauer, M., Albrecht, F., Schill, C., Perger, C., Havlik, P., Mosnier, A., Thornton, P., Wood-Sichra, U., Herrero, M., Becker-Reshef, I., Justice, C., Hansen, M., Gong, P., Abdel Aziz, S., Cipriani, A., Cumani, R., Cecchi, G., Conchedda, G., Ferreira, S., Gomez, A., Haffani, M., Kayitakire, F., Malanding, J., Mueller, R., Newby, T., Nonguierma, A., Olusegun, A., Ortner, S., Rajak, D. R., Rocha, J., Schepaschenko, D., Schepaschenko, M., Terekhov, A., Tiangwa, A., Vancutsem, C., Vintrou, E., Wenbin, W., van der Velde, M., Dunwoody, A., Kraxner, F., and Obersteiner, M. (2015). Mapping global cropland and field size. *Global Change Biology*, 21(5):1980 – 1992.

FURP (1987). Final report: Annex III - Volume 5. Busia district. Technical report, Government of Kenya, Ministry of Agriculture, Livestock and Fisheries.

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., and Godfray, H. C. J. (2013). Sustainable intensification in agriculture: Premises and policies. *Science*, 341:33 – 34.

Gebregziabher, S., Mouazen, A. M., van Brussel, H., Ramon, H., Nyssen, J., Verplancke, H., Behailu, M., Deckers, J., and de Baerdemaeker, J. (2006). Animal drawn tillage, the Ethiopian ard plough, maresha: A review. *Soil and Tillage Research*, 89:129 – 143.

Geerts, S., Raes, D., Garcia, M., Taboada, C., Miranda, R., Cusicanqui, J., Mhizha, T., and Vacher, J. (2009). Modeling the potential for closing quinoa yield gaps under varying water availability in the Bolivian Altiplano. *Agricultural Water Management*, 96:1652 – 1658.

Gerhart, J. (1975). The diffusion of hybrid maize in western Kenya. Technical report, CIMMYT.

Getnet, M., van Ittersum, M. K., Hengsdijk, H., and Descheemaeker, K. (2016). Yield gaps and resource use across farming zones in the Central Rift Valley of Ethiopia. *Experimental Agriculture*, 52:493 – 517.

Gianessi, L. P. (2013). The increasing importance of herbicides in worldwide crop production. *Pest Management Science*, 69:1099 – 1105.

Giannakas, K., Schoney, R., and Tzouvelekas, V. (2001). Technical efficiency, technological change and output growth of wheat farms in Saskatchewan. *Canadian Journal of Agricultural Economics*, 49:135 – 152.

Giller, K. E., Chalk, P., Dobermann, A., Hammond, L., Heffer, P., Ladha, J., Nyamudeza, P., Maene, L., Ssali, H., and Freney, J. (2004). *Agriculture and the Nitrogen Cycle: Assessing the Impacts of Fertilizer Use on Food Production and the Environment*, chapter Emerging Technologies to Increase the Efficiency of Use of Fertilzer Nitrogen, pages 35 – 51. SCOPE Series, Island Press.

Giller, K. E., Rowe, E. C., de Ridder, N., and van Keulen, H. (2006). Resource use dynamics and interactions in the tropics: Scaling up in space and time. *Agricultural Systems*, 88:8 – 27.

Giller, K. E., Tittonell, P., Rufino, M. C., van Wijk, M. T., Zingore, S., Mapfumo, P., Adjei-Nsiah, S., Herrero, M., Chikowo, R., Corbeels, M., Rowe, E. C., Baijukya, F., Mwijage, A., Smith, J., Yeboah, E., van der Burg, W. J., Sanogo, O. M., Misiko, M., de Ridder, N., Karanja, S., Kaizzi, C., K'ungu, J., Mwale, M., Nwaga, D., Pacini, C., and Vanlauwe, B. (2011). Communicating complexity: Integrated assessment of trade-offs concerning soil fertility management within African farming systems to support innovation and development. *Agricultural Systems*, 104:191 – 203.

Gines, H. C., Redondo, G. O., Estigoy, A. P., and Dobermann, A. (2004). Site-specific nutrient management in irrigated rice systems of Central Luzon, Philippines. In Dobermann, A. and Witt, C., editors, *Increasing productivity of intensive rice systems through site-specific nutrient management*. Enfield, N.H. (USA) and Los Baños (Philippines): Science Publishers, Inc., and International Rice Research Institute (IRRI).

Gittleman, M., Jordan, K., and Brelsford, E. (2012). Using citizen science to quantify community garden crop yields. *Cities and the Environment (CATE)*, 5:4.

Goel, P. K., Prasher, S. O., Landry, J. A., Patel, R. M., Bonnell, R. B., Viau, A. A., and Miller, J. R. (2003). Potential of airborne hyperspectral remote sensing to detect nitrogen deficiency and weed infestation in corn. *Computers and Electronics in Agriculture*, 38:99 – 124.

GoK (2015). Economic review of agriculture. Technical report, Government of Kenya, Ministry of Agriculture, Livestock and Fisheries.

Gomez, C., Rossel, R. A. V., and McBratney, A. B. (2008). Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma*, 146:403 – 411.

Grace, K., Husak, G., and Bogle, S. (2014). Estimating agricultural production in marginal and food insecure areas in Kenya using very high resolution remotely sensed imagery. *Applied Geography*, 55:257 – 265.

Grassini, P. and Cassman, K. G. (2012). High-yield maize with large net energy yield and small global warming intensity. *Proceedings of the National Academy of Sciences*, 109:1074 – 1079.

Grassini, P., Thorburn, J., Burr, C., and Cassman, K. G. (2011a). High-yield irrigated maize in the Western U.S. Corn Belt: I. On-farm yield, yield potential, and impact of agronomic practices. *Field Crops Research*, 120:142 – 150.

Grassini, P., van Bussel, L. G. J., van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., van Ittersum, M. K., and Cassman, K. G. (2015). How good is good enough? Data requirements for reliable crop yield simulations and yield gap analysis. *Field Crops Research*, 177:49 – 63.

Grassini, P., Yang, H., Irmak, S., Thorburn, J., Burr, C., and Cassman, K. G. (2011b). High-yield irrigated maize in the Western U.S. Corn Belt: II. Irrigation management and crop water productivity. *Field Crops Research*, 120:133 – 141.

Guilpart, N., Grassini, P., Sadras, V. O., Timsina, J., and Cassman, K. G. (2017). Estimating yield gaps at the cropping system level. *Field Crops Research*, 206:21 – 32.

GYGA (2016). Application of the GYGA approach to Ethiopia (available online). Technical report, Wageningen University and University of Nebraska-Lincoln.

Habte, D., Tadesse, K., Admasu, W., Desalegn, T., and Mekonen, A. (2014). Agronomic and economic evaluation of the N and P response of bread wheat growin in the moist and humid mighighland vertisols areas of Arsi zone, Ethiopia. *African Journal of Agricultural Research*, 10:89 – 99.

Haefele, S. M., Jabbar, S. M. A., Siopongco, J. D. L. C., Tirol-Padre, A., Amarante, S. T., Cruz, P. C. S., and Cosico, W. C. (2008). Nitrogen use efficiency in selected rice (*Oryza sativa* L.) genotypes under different water regimes and nitrogen levels. *Field Crops Research*, 107:137 – 146.

Haefele, S. M., Wopereis, M. C. S., Donovan, C., and Maubuisson, J. (2001). Improving the productivity and profitability of irrigated rice production in Mauritania. *European Journal of Agronomy*, 14:181 – 196.

Haefele, S. M., Wopereis, M. C. S., Ndiaye, M. K., Barro, S. E., and Isselmou, M. O. (2003). Internal nutrient efficiencies, fertilizer recovery rates and indigenous nutrient supply of irrigated lowland rice in Sahelian West Africa. *Field Crops Research*, 80:19 – 32.

Haggblade, S., Hazell, P. B. R., and Reardon, T. (2007). *Transforming the Rural Nonfarm Economy: Opportunities and Threats in the Developing World*. IFPRI, The Johns Hopkins University Press.

Hall, A. J. and Richards, R. A. (2013). Prognosis for genetic improvement of yield potential and water-limited yield of major grain crops. *Field Crops Research*, 143:18 – 33.

Hassan, R. M., Murithi, F., and Kamau, G. (1998a). *Maize Technology Development and Transfer: A GIS Application for Research Planning in Kenya*, chapter Determinants of fertilizer use and the gap between farmers' yields and potential yields in Kenya, pages 137 – 162. CAB International.

Hassan, R. M., Onyango, R., and Rutto, J. K. (1998b). *Maize Technology Development and Transfer: A GIS Application for Research Planning in Kenya*, chapter Relevance of maize research in Kenya to maize production problems perceived by farmers, pages 71 – 88. CAB International.

Hazell, P., Poulton, C., Wiggins, S., and Dorward, A. (2010). The future of small farms: Trajectories and policy priorities. *World Development*, 38:1349 – 1361.

He, W., Struik, P. C., Wang, J., and Zhang, X. (1998). Potential and actual yields of potato at different elevations and in different seasons in subtropical Southwest China. *Journal of Agronomy and Crop Science*, 180:93 – 99.

Helfand, S. M. and Levine, E. S. (2004). Farm size and the determinants of productive efficiency in the Brazilian Center-West. *Agricultural Economics*, 31:241 – 249.

Henderson, B., Godde, C., Medina-Hidalgo, D., van Wijk, M. T., Silvestri, S., Douxchamps, S., Stephenson, E., Power, B., Rigolot, C., Cacho, O., and Herrero, M.

(2016). Closing system-wide yield gaps to increase food production and mitigate GHGs among mixed crop-livestock smallholders in Sub-Saharan Africa. *Agricultural Systems*, 143:106 – 113.

Hengsdijk, H., Franke, A. C., van Wijk, M. T., and Giller, K. E. (2014). How small is beautiful? Food self-sufficiency and land gap analysis of smallholders in humid and semi-arid sub-Saharan Africa. Technical report, Plant Research International, Wageningen UR.

Henkens, P. L. C. M. and van Keulen, H. (2001). Mineral policy in the Netherlands and nitrate policy within the European Community. *Netherlands Journal of Agricultural Sciences*, 49:117 – 134.

Heong, K. L. (2014). Corrigendum to ”Rice planthopper problems in Asia - Consequences of ecosystem breakdown induced by insecticide misuse” [Crop Prot. 61 (2014) 102 - 110]. *Crop Protection*, 65:247.

Herdt, R. W. (1979). *Farm-level constraints to high rice yields in Asia: 1974 - 1977*, chapter An overview of the Constraints project results, page 411. IRRI.

Herdt, R. W. and Mandac, A. M. (1981). Modern technology and economic efficiency of Philippine rice farmers. *University of Chicago*.

Herrick, J. E., Urama, K. C., Karl, J. W., Boos, J., Johnson, M.-V. V., Shepherd, K. D., Hempel, J., Bestelmeyer, B. T., Davies, J., Guerra, J. L., Kosnik, C., Kimiti, D. W., Ekai, A. L., Muller, K., Norfleet, L., Ozor, N., Reinsch, T., Sarukhan, J., and West, L. T. (2013). The global Land-Potential Knowledge System (LandPKS): Supporting evidence-based, site-specific land use and management through cloud computing, mobile applications, and crowdsourcing. *Journal of Soil and Water Conservation*, 68:5 – 12.

Hoang, V.-N. (2013). Analysis of productive performance of crop production systems: An integrated analytical framework. *Agricultural Systems*, 116:16 – 24.

Hochman, Z., Gobbett, D., Holzworth, D., McClelland, T., van Rees, H., Marinoni, O., Garcia, J. N., and Horan, H. (2012). Quantifying yield gaps in rainfed cropping systems: A case study of wheat in Australia. *Field Crops Research*, 136:85 – 96.

Hochman, Z., Prestwidge, D., and Carberry, P. S. (2014). Crop sequences in Australia’s northern grain zone are less agronomically efficient than implied by the sum of their parts. *Agricultural Systems*, 129:124 – 132.

Hoffmann, M. P., Haakana, M., Asseng, S., Hohn, J. G., Palosuo, T., Ruiz-Ramos, M., Fronzek, S., Ewert, F., Gaiser, T., Kassie, B. T., Paff, K., Rezaei, E. E., Rodriguez, A., Semenov, M., Srivastava, A. K., Stratonovitch, P., Tao, F., Chen, Y., and Roetter, R. P. (2017). How does inter-annual variability of attainable yield affect the magnitude of yield gaps for wheat and maize? An analysis at ten sites. *Agricultural Systems*, pages –.

Hossain, M. and Fischer, K. S. (1995). Rice research for food security and sustainable agricultural development in Asia: Achievements and future challenges. *GeoJournal*, 35:286 – 298.

Huat, J., Doré, T., and Aubry, C. (2013). Limiting factors for yields of field tomatoes grown by smallholders in tropical regions. *Crop Protection*, 44:120 – 127.

ILRI (2011). Gender, livestock and livelihood indicators. Technical report, International Livestock Research Institute, Nairobi.

Iráizoz, B., Rapún, M., and Zabaleta, I. (2003). Assessing the technical efficiency of horticultural production in Navarra, Spain. *Agricultural Systems*, 78:387 – 403.

Iversen, T. M., Grant, R., and Nielsen, K. (1998). Nitrogen enrichment of European inland and marine waters with special attention to Danish policy measures. *Environmental Pollution*, 102:771 – 780.

Jaetzold, R. and Schmidt, H. (1982). *Farm management handbook of Kenya Vol. II - Natural conditions and farm management information. Part A West Kenya (Nyanza and Western Provinces)*. Nairobi, Kenya: Ministry of Agriculture.

Jaetzold, R., Schmidt, H., Hornetz, B., and Shisanya, C. (2009). Farm management handbook of Kenya Vol. II - Part A west Kenya subpart A2 Nyanza province. Technical report, Ministry of Agriculture, Nairobi, Kenya.

Jain, M., Mondal, P., DeFries, R. S., Small, C., and Galford, G. L. (2013). Mapping cropping intensity of smallholder farms: A comparison of methods using multiple sensors. *Remote Sensing of Environment*, 134:210 – 223.

Janssen, B. H., Guiking, F. C. T., van der Eijk, D., Smaling, E. M. A., Wolf, J., and van Reuler, H. (1990). A system for quantitative evaluation of the fertility of tropical soils (QUEFTS). *Geoderma*, 46:299 – 318.

Janssen, S. and van Ittersum, M. K. (2007). Assessing farm innovations and responses to policies: A review of bio-economic farm models. *Agricultural Systems*, 94:622 – 636.

Jatoe, J. B. D., Yiridoe, E. K., Weersink, A., and Clark, J. S. (2008). Economic and environmental impacts of introducing land use policies and rotations on Prince Edward Island potato farms. *Land Use Policy*, 25:309 – 319.

Jayne, T. S., Chamberlin, J., and Headey, D. D. (2014). Land pressures, the evolution of farming systems, and development strategies in Africa: A synthesis. *Food Policy*, 48:1 – 17.

Jiji, G. W. and Nadar, P. (2016). Soil type identification using remotely sensed data for agricultural purpose. *Journal of The Institution of Engineers (India): Series B*, 97:405 – 411.

Kamanga, B. C. G., Waddington, S. R., Whitbread, A. M., Almekinders, C. J. M., and Giller, K. E. (2014). Improving the efficiency of use of small amounts of nitrogen and phosphorus fertiliser on smallholder maize in central Malawi. *Experimental Agriculture*, 50:229 – 249.

Kanellopoulos, A., Berentsen, P. B. M., van Ittersum, M. K., and Lansink, A. O. (2012). A method to select alternative agricultural activities for future-oriented land use studies. *European Journal of Agronomy*, 40:75 – 85.

Kanellopoulos, A., Reidsma, P., Wolf, J., and van Ittersum, M. K. (2014). Assessing climate change and associated socio-economic scenarios for arable farming in the Netherlands: An application of benchmarking and bio-economic farm modelling. *European Journal of Agronomy*, 52, Part A:69 – 80.

Kang'ethe, E. (2011). Situation analysis: Improving food safety in the maize value chain in Kenya. Technical report, FAO.

KARI (1990). The Fertilizer Use and Recommendations Project (FURP): Main Report. Technical report, National Agricultural Research Laboratories, Nairobi, Kenya.

KARI (1993). The Fertilizer Use and Recommendations Project (FURP): Extended Phase II - Vol. 9. Busia district. Technical report, National Agricultural Research Laboratories, Nairobi, Kenya.

Kassie, B. T., van Ittersum, M. K., Hengsdijk, H., Asseng, S., Wolf, J., and Rötter, R. P. (2014). Climate-induced yield variability and yield gaps of maize (*Zea mays* L.) in the Central Rift Valley of Ethiopia. *Field Crops Research*, 160:41 – 53.

Keating, B. A., Carberry, P. S., Bindraban, P. S., Asseng, S., Meinke, H., and Dixon, J. (2010). Eco-efficient agriculture: Concepts, challenges, and opportunities. *Crop Science*, 50:109 – 119.

Kebede, Y., Baudron, F., Bianchi, F., and Tittonell, P. (2015). A multi-scale analysis (plot, farm, landscape) of the factors controlling maize stem borer infestation in Southern Ethiopia. towards a multi-scale push-pull approach. In *5th International Symposium for Farming Systems Design, Montpellier, France*.

Kerkvliet, B. (1990). *Everyday politics in the Philippines: Class and status relations in a Central Luzon Village*. University of California Press.

Kersebaum, K. C., Hecker, J.-M., Mirschel, W., and Wegehenkel, M. (2007). *Modelling water and nutrient dynamics in soil-crop systems: a comparison of simulation models applied on common data sets*, pages 1 – 17. Springer Netherlands, Dordrecht.

Khush, G. S. (2001). Green revolution: the way forward. *Nature Reviews Genetics*, 2:815 – 822.

Kibaara, B., Ariga, J., Olwande, J., and Jayne, T. S. (2008). Trends in agricultural productivity: 1997-2007. Technical report, Tegemeo Institute, Egerton University.

Kimani, V. and Gruere, G. (2010). Implications of import regulations and information requirements under the Cartagena protocol on biosafety for GM commodities in Kenya. *AgBioForum*, 13:222 – 241.

Klapwijk, C. J., van Wijk, M. T., Rosenstock, T. S., van Asten, P. J. A., Thornton, P. K., and Giller, K. E. (2014). Analysis of trade-offs in agricultural systems: current status and way forward. *Current Opinion in Environmental Sustainability*, 6:110 – 115.

Koirala, K. H., Mishra, A., and Mohanty, S. (2016). Impact of land ownership on productivity and efficiency of rice farmers: The case of the Philippines. *Land Use Policy*, 50:371 – 378.

Koning, N., van Ittersum, M. K., Becx, G. A., Boekel, M. A. J. S. V., Brandenburg, W. A., van den Broek, J. A., Goudriaan, J., van Hofwegen, G., Jongeneel, R. A., Schiere, J. B., and Smies, M. (2008). Long-term global availability of food: continued abundance or new scarcity? *NJAS - Wageningen Journal of Life Sciences*, 55:229 – 292.

Kropff, M. J., Cassman, K. G., van Laar, H. H., and Peng, S. (1993). Nitrogen and yield potential of irrigated rice. *Plant and Soil*, 155/156:391 – 394.

Kumbhakar, S. C. and Lovell, C. A. K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press.

Laborte, A. G., de Bie, K. C., Smaling, E. M. A., Moya, P. F., Boling, A. A., and van Ittersum, M. K. (2012). Rice yields and yield gaps in Southeast Asia: Past trends and future outlook. *European Journal of Agronomy*, 36:9 – 20.

Laborte, A. G., Paguirigan, N. C., Moya, P. F., Nelson, A., Sparks, A. H., and Gregorio, G. B. (2015). Farmers' preference for rice traits: Insights from farm surveys in Central Luzon, Philippines, 1966-2012. *PLoS ONE*, 10:e0136562.

Lampayan, R. M., Bouman, B. A. M., de Dios, J. L., Espiritu, A. J., Soriano, J. B., Lactaoen, A. T., Faronilo, J. E., and Thant, K. M. (2010). Yield of aerobic rice in rainfed lowlands of the Philippines as affected by nitrogen management and row spacing. *Field Crops Research*, 116:165 – 174.

Lampayan, R. M., Rejesus, R. M., Singleton, G. R., and Bouman, B. A. M. (2015). Adoption and economics of alternate wetting and drying water management for irrigated lowland rice. *Field Crops Research*, 170:95 – 108.

Landis, D. A., Gardiner, M. M., van der Werf, W., and Swinton, S. M. (2008). Increasing corn for biofuel production reduces biocontrol services in agricultural landscapes. *Proceedings of the National Academy of Sciences*, 105:20552 – 20557.

Lansigan, F. P., de los Santos, W. L., and Coladilla, J. O. (2000). Agronomic impacts of climate variability on rice production in the Philippines. *Agriculture, Ecosystems & Environment*, 82:129 – 137.

Lantican, M. A., Lampayan, R. M., Bhuiyan, S. I., and Yadav, M. K. (1999). Determinants of improving productivity of dry-seeded rice in rainfed lowlands. *Experimental Agriculture*, 35:127 – 140.

Lassaletta, L., Billen, G., Grizzetti, B., Anglade, J., and Garnier, J. (2014). 50 year trends in nitrogen use efficiency of world cropping systems: the relationship between yield and nitrogen input to cropland. *Environmental Research Letters*, 9:105011.

Launio, C. C., Redondo, G. O., Beltran, J. C., and Morooka, Y. (2008). Adoption and spatial diversity of later generation modern rice varieties in the Philippines. *Agronomy Journal*, 100:1380 – 1389.

Ledesma, A. J. (1980). Land reform in East and Southeast Asia: A comparative approach. *Philippine Studies*, 28:451 – 481.

Leonardo, W. J., van de Ven, G. W. J., Udo, H., Kanellopoulos, A., Sitoe, A., and Giller, K. E. (2015). Labour not land constrains agricultural production and food self-sufficiency in maize-based smallholder farming systems in Mozambique. *Food Security*, 7:857 – 874.

li Liang, W., Carberry, P., yan Wang, G., hai Lu, R., zhan Lu, H., and ping Xia, A. (2011). Quantifying the yield gap in wheat-maize cropping systems of the Hebei Plain, China. *Field Crops Research*, 124:180 – 185.

Licker, R., Johnston, M., Foley, J. A., Barford, C., Kucharik, C. J., Monfreda, C., and Ramankutty, N. (2010). Mind the gap: how do climate and agricultural management explain the 'yield gap' of croplands around the world? *Global Ecology and Biogeography*, 19:769 – 782.

LNV (2009). Vierde Nederlandse Actieprogramma betreffende de Nitraatrichtlijn (2010-2013). Technical report, Ministerie van Landbouw, Natuur en Voedselkwaliteit. 's-Gravenhage, The Netherlands.

Lobell, D. B. (2013). The use of satellite data for crop yield gap analysis. *Field Crops Research*, 143:56 – 64.

Lobell, D. B., Asner, G. P., Ortiz-Monasterio, J. I., and Benning, T. L. (2003). Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. *Agriculture, Ecosystems & Environment*, 94:205 – 220.

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

Lobell, D. B., Ivan-Monasterio, J., Asner, G. P., Naylor, R. L., and Falcon, W. P. (2005). Combining field surveys, remote sensing, and regression trees to understand yield variations in an irrigated wheat landscape. *Agronomy Journal*, 97:241 – 249.

Lobell, D. B., Ortiz-Monasterio, J. I., Addams, C. L., and Asner, G. P. (2002). Soil, climate, and management impacts on regional wheat productivity in Mexico from remote sensing. *Agricultural and Forest Meteorology*, 114:31 – 43.

Loevinsohn, M. E., Bandong, J. B., and Alviola, A. A. (1993). Asynchrony in cultivation among Philippine rice farmers: Causes and prospects for change. *Agricultural Systems*, 41:419 – 439.

Lowder, S. K., Skoet, J., and Raney, T. (2016). The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Development*, 87:16 – 29.

Lu, C. and Fan, L. (2013). Winter wheat yield potentials and yield gaps in the North China Plain. *Field Crops Research*, 143:98 – 105.

Luo, D., Ganesh, S., and Koolaard, J. (2014). *predictmeans: Calculate Predicted Means for Linear Models*. R package version 0.99.

Ma, L., Feng, S., Reidsma, P., Qu, F., and Heerink, N. (2012). Identifying entry points to improve fertilizer use efficiency in Taihu Basin, China. *Land Use Policy*, 37:52 – 59.

Mandac, A. M. and Flinn, J. C. (1985). Determinants of fertiliser use on rainfed rice in Bicol region, Philippines. *Philippine Journal of Crop Science*, 10:123 – 128.

Mandemaker, M., Bakker, M., and Stoorvogel, J. (2011). The role of governance in agricultural expansion and intensification: a global study of arable agriculture. *Ecology and Society*, 16:8.

Mandryk, M., Reidsma, P., Kanellopoulos, A., Groot, J. C. J., and van Ittersum, M. K. (2014). The role of farmers' objectives in current farm practices and adaptation preferences: a case study in Flevoland, the Netherlands. *Regional Environmental Change*, 14(4):1463 – 1478.

Masters, W. A., Djurfeldt, A. A., Haan, C. D., Hazell, P., Jayne, T., Jirstrom, M., and Reardon, T. (2013). Urbanization and farm size in Asia and Africa: Implications for food security and agricultural research. *Global Food Security*, 2:156 – 165.

Mazzilli, S. R., Ernst, O. R., de Mello, V. P., and Pérez, C. A. (2016). Yield losses on wheat crops associated to the previous winter crop: Impact of agronomic practices based on on-farm analysis. *European Journal of Agronomy*, 75:99 – 104.

McCann, J. C. (1995). *People of the Plow: An Agricultural History of Ethiopia, 1800 - 1990*. The University of Wisconsin Press.

Meeusen, W. and van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18:435 – 444.

Mellisse, B. T., Descheemaeker, K. K. E., Giller, K. E., Abebe, T., and van de Ven, G. W. J. (2017). Are traditional home gardens in southern Ethiopia heading for extinction? Implications for productivity, plant species diversity and food security. *Agriculture, Ecosystems & Environment*, 252:1 – 13.

Meng, Q., Hou, P., Wu, L., Chen, X., Cui, Z., and Zhang, F. (2012). Understanding production potentials and yield gaps in intensive maize production in China. *Field Crops Research*, 143:91 – 97.

Meul, M., Nevens, F., Reheul, D., and Hofman, G. (2007). Energy use efficiency of specialised dairy, arable and pig farms in Flanders. *Agriculture, Ecosystems & Environment*, 119:135 – 144.

Monchuk, D. C., Chen, Z., and Bonaparte, Y. (2010). Explaining production inefficiency in China's agriculture using data envelopment analysis and semi-parametric bootstrapping. *China Economic Review*, 21:346 – 354.

Monjardino, M., McBeath, T., Ouzman, J., Llewellyn, R., and Jones, B. (2015). Farmer risk-aversion limits closure of yield and profit gaps: A study of nitrogen management in the southern Australian wheat belt. *Agricultural Systems*, 137:108 – 118.

Moya, P. F., Dawe, D., Pabale, D., Tiongco, M., Chien, N. V., Devarajan, S., Djatiharti, A., Lai, N. X., Niyomvit, L., Ping, H. X., Redondo, G., and Wardana, P. (2004). The economics of intensively irrigated rice in Asia. In Dobermann, A. and Witt, C., editors, *Increasing productivity of intensive rice systems through site-specific nutrient management*. Enfield, N.H. (USA) and Los Baños (Philippines): Science Publishers, Inc., and International Rice Research Institute (IRRI).

Moya, P. F., Kajisa, K., Barker, R., Mohanty, S., Gascon, F., and Valentin, M. S. (2015). *Changes in rice farming in the Philippines: Insights from five decades of a household level survey*. International Rice Research Institute (IRRI).

Moya, T. B. (2001). Biophysical characterization of Philippine crop production environment in support of technology intervention and research prioritization. Technical report, Philippines Institute for Development Studies. Makati City, Philippines.

Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., and Foley, J. A. (2012). Closing yield gaps through nutrient and water management. *Nature*, 490:254 – 257.

Mussgnug, F., Becker, M., Son, T. T., Buresh, R. J., and Vlek, P. L. G. (2006). Yield gaps and nutrient balances in intensive, rice-based cropping systems on degraded soils in the red river delta of vietnam. *Field Crops Research*, 98:127 – 140.

Muto, M. and Yamano, T. (2009). The impact of mobile phone coverage expansion on market participation: Panel data evidence from Uganda. *World Development*, 37:1887 – 1896.

Naab, J. B., Singh, P., Boote, K. J., Jones, J. W., and Marfo, K. O. (2004). Using the CROPGRO-peanut model to quantify yield gaps of peanut in the Guinean Savanna Zone of Ghana. *Agronomy Journal*, 96:1231 – 1242.

Neumann, K., Verburg, P. H., Stehfest, E., and Muller, C. (2010). The yield gap of global grain production: A spatial analysis. *Agricultural Systems*, 103:316 – 326.

Nijhof, K. (1987). *The concentration of macro-nutrients in economic products and residues of (sub)tropical field crops*. Center for World Food Studies, Amsterdam - Wageningen.

Nijland, G. O., Schouls, J., and Goudriaan, J. (2008). Integrating the production functions of Liebig, Michaelis-Menten, Mitscherlich and Liebscher into one system dynamics model. *NJAS - Wageningen Journal of Life Sciences*, 55:199 – 224.

Njoroge, K., Mugo, S. N., and Laboso, A. K. (1995). *Maize variety national performance trials*. Kenya Agricultural Research Institute, Nairobi, Kenya.

Nyoro, J. K., Kirimi, L., and Jayne, T. S. (2004). Competitiveness of Kenyan and Ugandan maize production: Challenges for the future. Technical report, Tegemeo Institute, Egerton University.

Nyssen, J., Govaerts, B., Araya, T., Cornelis, W. M., Bauer, H., Haile, M., Sayre, K., and Deckers, J. (2011). The use of the marasha ard plough for conservation agriculture in Northern Ethiopia. *Agronomy for Sustainable Development*, 31:287 – 297.

Okumu, M. O., van Asten, P. J. A., Kahangi, E., Okech, S. H., Jefwa, J., and Vanlauwe, B. (2011). Production gradients in smallholder banana (cv. Giant cavendish) farms in Central Kenya. *Scientia Horticulturae*, 127:475 – 481.

Oliver, Y. M. and Robertson, M. J. (2013). Quantifying the spatial pattern of the yield gap within a farm in a low rainfall Mediterranean climate. *Field Crops Research*, 150:29 – 41.

Olk, D. C., Cassman, K. G., Mahieu, N., and Randall, E. W. (1998). Conserved chemical properties of young humic acid fractions in tropical lowland soil under intensive irrigated rice cropping. *European Journal of Soil Science*, 49:337 – 349.

Ollenburger, M. H., Descheemaeker, K., Crane, T. A., Sanogo, O. M., and Giller, K. E. (2016). Waking the Sleeping Giant: Agricultural intensification, extensification or stagnation in Mali's Guinea savannah. *Agricultural Systems*, 148:58 – 70.

Ortiz-Monasterio, J. I. and Lobell, D. B. (2007). Remote sensing assessment of regional yield losses due to sub-optimal planting dates and fallow period weed management. *Field Crops Research*, 101:80 – 87.

Otsuka, K. (1991). Determinants and consequences of land reform implementation in the Philippines. *Journal of Development Economics*, 35:339 – 355.

Otsuka, K. (2000). Role of agricultural research in poverty reduction: Lessons from the Asian experience. *Food Policy*, 25:447 – 462.

Pacheco, A., Bannari, A., Staenz, K., and McNairn, H. (2008). Deriving percent crop cover over agriculture canopies using hyperspectral remote sensing. *Canadian Journal of Remote Sensing*, 34:110 – 123.

Pampolino, M. F., Laureles, E. V., Gines, H. C., and Buresh, R. J. (2008). Soil carbon and nitrogen changes in long-term continuous lowland rice cropping. *Soil Science Society of America Journal*, 72(3):798.

Pampolino, M. F., Manguiat, I. J., Ramanathan, S., Gines, H. C., Tan, P. S., Chi, T. T. N., Rajendran, R., and Buresh, R. J. (2007). Environmental impact and economic benefits of site-specific nutrient management (SSNM) in irrigated rice systems. *Agricultural Systems*, 93:1 – 24.

Pandey, S. and Velasco, L. (2002). *Direct Seeding: Research strategies and opportunities*, chapter Economics of direct seeding in Asia: Patterns of adoption and research priorities, pages 3 – 14. International Rice Research Institute (IRRI).

Passioura, J. B. (1996). Simulation models: Science, snake oil, education, or engineering? *Agronomy Journal*, 88:690 – 694.

Pasuquin, J. M., Pampolino, M. F., Witt, C., Dobermann, A., Oberthur, T., Fisher, M. J., and Inubushi, K. (2014). Closing yield gaps in maize production in Southeast Asia through site-specific nutrient management. *Field Crops Research*, 156:219 – 230.

Paustian, K. (2013). Bridging the data gap: engaging developing country farmers in greenhouse gas accounting. *Environmental Research Letters*, 8:021001.

PDA (2012). Food Staples Sufficiency Program 2011 - 2016. Technical report, Department of Agriculture, The Philippines.

Peng, S., Garcia, F. V., Gines, H. C., Laza, R. C., Samson, M. I., Sanico, A. L., Visperas, R. M., and Cassman, K. G. (1995). Nitrogen use efficiency of irrigated tropical rice established by broadcast wet-seeding and transplanting. *Fertilizer research*, 45:123 – 134.

Petropoulos, G. P., Ireland, G., and Barrett, B. (2015). Surface soil moisture retrievals from remote sensing: Current status, products & future trends. *Physics and Chemistry of the Earth, Parts A/B/C*, 83:36 – 56.

Pingali, P. L. (1997). From subsistence to commercial production systems: The transformation of Asian agriculture. *American Journal of Agricultural Economics*, 79:628 – 634.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and R Core Team (2016). *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-128.

Plaza-Bonilla, D., Nolot, J.-M., Raffaillac, D., and Justes, E. (2017). Innovative cropping systems to reduce N inputs and maintain wheat yields by inserting grain legumes and cover crops in southwestern France. *European Journal of Agronomy*, 82, Part B:331 – 341.

Pongratz, J., Reick, C., Raddatz, T., and Claussen, M. (2008). A reconstruction of global agricultural areas and land cover for the last millennium. *Global Biogeochemical Cycles*, 22:1 – 16.

Ponisio, L. C., M'Gonigle, L. K., Mace, K. C., Palomino, J., de Valpine, P., and Kremen, C. (2014). Diversification practices reduce organic to conventional yield gap. *Proceedings of the Royal Society of London B: Biological Sciences*, 282.

Ponsioen, T. C., Hengsdijk, H., Wolf, J., van Ittersum, M. K., Rötter, R. P., Son, T. T., and Laborte, A. G. (2006). TechnoGIN, a tool for exploring and evaluating resource use efficiency of cropping systems in East and Southeast Asia. *Agricultural Systems*, 87:80 – 100.

Pratihast, A. K., Herold, M., Avitabile, V., de Bruin, S., Bartholomeus, H., Souza Jr., C. M., and Ribbe, L. (2013). Mobile devices for community-based REDD+ monitoring: A case study for central Vietnam. *Sensors*, 13:21 – 38.

Pretty, J. and Bharucha, Z. P. (2014). Sustainable intensification in agricultural systems. *Annals of Botany*, 114:1571 – 1596.

Pretty, J., Toulmin, C., and Williams, S. (2011). Sustainable intensification in African agriculture. *International Journal of Agricultural Sustainability*, 9:5 – 24.

Prost, L., Makowski, D., and Jeuffroy, M.-H. (2008). Comparison of stepwise selection and bayesian model averaging for yield gap analysis. *Ecological Modelling*, 219:66 – 76.

PSA (2017). CountrySTAT Philippines. Available at: http://psa.gov.ph/ (accessed 15.01.2017).

PSA-BAS (2015). CountrySTAT Philippines. Available at: http://countrystat.psa.gov.ph/ (accessed 05.10.2015).

Qin, J., Ali, J., Quinones, C., Marcaida, M., Li, T., Tang, Q., and Jagadish, S. V. K. (2014). Nitrogen deficient, moderately sufficient and potential supply alters climatic yield potential, grain and head rice yield among tropical rice varieties. *Journal of Agronomy and Crop Science*, Submitted.

Quilty, J. R., McKinley, J., Pede, V. O., Buresh, R. J., Correa, T. Q., and Sandro, J. M. (2014). Energy efficiency of rice production in farmers' fields and intensively cropped research fields in the Philippines. *Field Crops Research*, 168:8 – 18.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rahman, M., Blackwell, B., Banerjee, N., and Saraswat, D. (2015). Smartphone-based hierarchical crowdsourcing for weed identification. *Computers and Electronics in Agriculture*, 113:14 – 23.

Ramisch, J. J. (2011). *Knowing Nature: Conversations at the intersection of political ecology and science studies*, chapter Experiments as 'performances': Interpreting farmers' soil fertility management practices in western Kenya, pages 280 – 295. University of Chicago Press.

Ramisch, J. J. (2014). 'They don't know what they are talking about': Learning from the dissonances in dialogue about soil fertility knowledge and experimental practice in western Kenya. *Geoforum*, 55:120 – 132.

Ramisch, J. J. (2016). 'Never at ease': cellphones, multilocational households, and the metabolic rift in western Kenya. *Agriculture and Human Values*, 33:979 – 995.

Ramisch, J. J., Misiko, M. T., Ekise, I. E., and Mukalama, J. B. (2006). Strengthening 'folk ecology': community-based learning for integrated soil fertility management, western Kenya. *International Journal of Agricultural Sustainability*, 4:154 – 168.

Reckling, M., Hecker, J.-M., Bergkvist, G., Watson, C. A., Zander, P., Schlafke, N., Stoddard, F. L., Eory, V., Topp, C. F. E., Maire, J., and Bachinger, J. (2016).

A cropping system assessment framework - Evaluating effects of introducing legumes into crop rotations. *European Journal of Agronomy*, 76:186 – 197.

Reidsma, P., Ewert, F., Boogaard, H., and van Diepen, K. (2009a). Regional crop modelling in Europe: The impact of climatic conditions and farm characteristics on maize yields. *Agricultural Systems*, 100:51 – 60.

Reidsma, P., Lansink, A. O., and Ewert, F. (2009b). Economic impacts of climatic variability and subsidies on European agriculture and observed adaptation strategies. *Mitigation and Adaptation Strategies for Global Change*, 14:35 – 59.

Reidsma, P., Rietema, J., Yan, Y., Kroes, J., and Silva, J. V. (2015a). Potential yield and yield gap at farm level are different from the field level: A case study on a large Dutch potato farm. In *5th International Symposium for Farming Systems Design, Montpellier, France*.

Reidsma, P., Wolf, J., Kanellopoulos, A., Schaap, B. F., Mandryk, M., Verhagen, J., and van Ittersum, M. K. (2015b). Climate change impact and adaptation research requires integrated assessment and farming systems analysis: a case study in the Netherlands. *Environmental Research Letters*, 10:045004.

Reinhard, S., Lovell, C. A. K., and Thijssen, G. (1999). Econometric estimation of technical and environmental efficiency: An application to Dutch dairy farms. *American Journal of Agricultural Economics*, 81:44 – 60.

Rejesus, R. M., Palis, F. G., Rodriguez, D. G. P., Lampayan, R. M., and Bouman, B. A. M. (2011). Impact of the alternate wetting and drying (AWD) water-saving irrigation technique: Evidence from rice producers in the Philippines. *Food Policy*, 36:280 – 288.

Rijk, B., van Ittersum, M. K., and Withagen, J. (2013). Genetic progress in Dutch crop yields. *Field Crops Research*, 149:262 – 268.

Rockström, J., Williams, J., Daily, G., Noble, A., Matthews, N., Gordon, L., Wetterstrand, H., DeClerck, F., Shah, M., Steduto, P., de Fraiture, C., Hatibu, N., Unver, O., Bird, J., Sibanda, L., and Smith, J. (2017). Sustainable intensification of agriculture for human prosperity and global sustainability. *Ambio*, 46(1):4 – 17.

Rojas, O. (2007). Operational maize yield model development and validation based on remote sensing and agro-meteorological data in Kenya. *International Journal of Remote Sensing*, 28:3775 – 3793.

Rötter, R. P. and 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*, 53:69 – 95.

Roy, H. E., Pocock, M. J. O., Preston, C. D., Roy, D. B., Savage, J., Tweddle, J. C., and Robinson, L. D. (2012). Understanding citizen science & environmental monitoring. Final report on behalf of UK-EOF. Technical report, NERC Centre for Ecology & Hydrology and Natural History Museum.

Ryu, Y., Verfaillie, J., Macfarlane, C., Kobayashi, H., Sonnentag, O., Vargas, R., Ma, S., and Baldocchi, D. D. (2012). Continuous observation of tree leaf area index at ecosystem scale using upward-pointing digital cameras. *Remote Sensing of Environment*, 126:116 – 125.

Sadoulet, E. and de Janvry, A. (1995). *Quantitative Development Policy Analysis*. The Johns Hopkins University Press.

Sadras, V., Roget, D., and O'Leary, G. (2002). On-farm assessment of environmental and management constraints to wheat yield and efficiency in the use of rainfall in the Mallee. *Australian Journal of Agricultural Research*, 53:587 – 598.

Sadras, V. O. and Angus, J. F. (2006). Benchmarking water-use efficiency of rainfed wheat in dry environments. *Australian Journal of Agricultural Research*, 57:847 – 856.

Sattari, S. Z., van Ittersum, M. K., Bouwman, A. F., Smit, A. L., and Janssen, B. H. (2014). Crop yield response to soil fertility and N, P, K inputs in different environments: Testing and improving the QUEFTS model. *Field Crops Research*, 157:35 – 46.

Savary, S., Horgan, F., Willocquet, L., and Heong, K. L. (2012). A review of principles for sustainable pest management in rice. *Crop Protection*, 32:54 – 63.

Schnug, E., Heym, J., and Achwan, F. (1996). Establishing critical values for soil and plant analysis by means of the boundary line development system (Bolides). *Communications in Soil Science and*, 27:2739 – 2748.

Scholte, K. and s'Jacob, J. J. (1990). Effect of crop rotation, cultivar and nematicide on growth and yield of potato (*Solanum tuberosum* L.) in short rotations on a marine clay soil. *Potato Research*, 33:191 – 200.

Schröder, J. J. (2014). The position of mineral nitrogen fertilizer in efficient use of nitrogen and land: A review. *Natural Resources*, 5:936 – 948.

Schröder, J. J. and Neeteson, J. J. (2008). Nutrient management regulations in the Netherlands. *Geoderma*, 144:418 – 425.

Schulthess, U., Timsina, J., Herrera, J. M., and McDonald, A. (2012). Mapping field-scale yield gaps for maize: An example from Bangladesh. *Field Crops Research*, 143:151 – 156.

Setiyono, T. D., Walters, D. T., Cassman, K. G., Witt, C., and Dobermann, A. (2010). Estimating maize nutrient uptake requirements. *Field Crops Research*, 118:158 – 168.

Seufert, V., Ramankutty, N., and Foley, J. A. (2012). Comparing the yields of organic and conventional agriculture. *Nature*, 485:229 – 232.

Shatar, T. M. and McBratney, A. B. (2004). Boundary-line analysis of field-scale yield response to soil properties. *The Journal of Agricultural Science*, 142:553 – 560.

Sheahan, M. and Barrett, C. B. (2017). Ten striking facts about agricultural input use in Sub-Saharan Africa. *Food Policy*, 67:12 – 25.

Shelley, W., Lawley, R., and Robinson, D. A. (2013). Technology: Crowd-sourced soil data for Europe. *Nature*, 496:300.

Siebert, S., Kummu, M., Porkka, M., Döll, P., Ramankutty, N., and Scanlon, B. R. (2015). A global data set of the extent of irrigated land from 1900 to 2005. *Hydrology and Earth System Sciences*, 19:1521 – 1545.

Silva, J. V., Corcuera, F., Quilty, J. R., Li, T., and Laborte, A. G. (2014). Description of methodology for sampling and analysing soil properties in farmers' fields of Central Luzon Loop Survey. Unpublished.

Sime, G. and Aune, J. B. (2014). Maize response to fertilizer dosing at three sites in the Central Rift Valley of Ethiopia. *Agronomy*, 4:436 – 451.

Smale, M. and Jayne, T. S. (2003). Maize in Eastern and Southern Africa: 'Seeds' of success in retrospect. Technical report, IFPRI, EPTD Discussion Paper #97.

Smaling, E. M. A. and Janssen, B. H. (1993). Calibration of QUEFTS, a model predicting nutrient uptake and yields from chemical soil fertility indices. *Geoderma*, 59:21 – 44.

Smith, A., Snapp, S., Chikowo, R., Thorne, P., Bekunda, M., and Glover, J. (2017). Measuring sustainable intensification in smallholder agroecosystems: A review. *Global Food Security*, 12:127 – 138.

Snyder, K. A., Miththapala, S., Sommer, R., and Braslow, J. (2016). The yield gap: Closing the gap by widening the approach. *Experimental Agriculture*, pages 1 – 15.

Sopheap, U., Patanothai, A., and Aye, T. M. (2012). Unveiling constraints to cassava production in Cambodia: An analysis from farmers' yield variations. *International Journal of Plant Production*, 6:409 – 428.

Stoate, C., Baldi, A., Beja, P., Boatman, N., Herzon, I., van Doorn, A., de Snoo, G. R., Rakosy, L., and Ramwell, C. (2009). Ecological impacts of early 21st century agricultural change in Europe - A review. *Journal of Environmental Management*, 91:22 – 46.

Stone, G. D., Netting, R. M., and Stone, M. P. (1990). Seasonality, labor scheduling, and agricultural intensification in the Nigerian savanna. *American Anthropologist*, 92:7 – 23.

Struik, P. C. and Bonciarelli, F. (1997). Resource use at the cropping system level. *European Journal of Agronomy*, 7:133 – 143.

Struik, P. C., Kuyper, T. W., Brussaard, L., and Leeuwis, C. (2014). Deconstructing and unpacking scientific controversies in intensification and sustainability: why the tensions in concepts and values? *Current Opinion in Environmental Sustainability*, 8:80 – 88.

Stuart, A. M., Pame, A. R. P., Silva, J. V., Dikitanan, R. C., Rutsaert, P., Malabayabas, A. J. B., Lampayan, R. M., Radanielson, A. M., and Singleton, G. R. (2016). Yield gaps in rice-based farming systems: Insights from local studies and prospects for future analysis. *Field Crops Research*, 194:43 – 56.

Studwell, J. (2013). *How Asia Works: Success and failure in the world's most dynamic region*. Grove Press, New York.

Subedi, K. D. and Ma, B. L. (2009). Assessment of some major yield-limiting factors on maize production in a humid temperate environment. *Field Crops Research*, 110:21 – 26.

Sui, R., Thomasson, J. A., Hanks, J., and Wooten, J. (2008). Ground-based sensing system for weed mapping in cotton. *Computers and Electronics in Agriculture*, 60:31 – 38.

Sumberg, J. (2012). Mind the (yield) gap(s). *Food Security*, 4:509 – 518.

Taa, A., Tanner, D. G., and Bennie, A. T. P. (2004). Effects of stubble management, tillage and cropping sequence on wheat production in the south-eastern highlands of Ethiopia. *Soil and Tillage Research*, 76:69 – 82.

Tabbal, D. F., Bouman, B. A. M., Bhuiyan, S. I., Sibayan, E. B., and Sattar, M. A. (2002). On-farm strategies for reducing water input in irrigated rice; case studies in the Philippines. *Agricultural Water Management*, 56:93 – 112.

Takahashi, K. and Otsuka, K. (2009). The increasing importance of nonfarm income and the changing use of labor and capital in rice farming: the case of Central Luzon, 1979 - 2003. *Agricultural Economics*, 40:231 – 242.

Tanaka, A., Saito, K., Azoma, K., and Kobayashi, K. (2013). Factors affecting variation in farm yields of irrigated lowland rice in southern-central Benin. *European Journal of Agronomy*, 44:46 – 53.

Tanner, D. G., Gorfu, A., and Taa, A. (1993). Fertiliser effects on sustainability in the wheat-based small-holder farming systems of southeastern Ethiopia. *Field Crops Research*, 33:235 – 248.

Thiam, A., Bravo-Ureta, B. E., and Rivas, T. E. (2001). Technical efficiency in developing country agriculture: a meta-analysis. *Agricultural Economics*, 25:235 – 243.

Thorp, K. R., Steward, B. L., Kaleita, A. L., , and Batchelor, W. D. (2008). Using aerial hyperspectral remote sensing imagery to estimate corn plant stand density. *Transactions of the ASABE*, 51:311 – 320.

Tilman, D., Balzer, C., Hill, J., and Befort, B. L. (2011). Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences*, 108:20260 – 20264.

Tilman, D., Cassman, K. G., Matson, P. A., Naylor, R. L., and Polasky, S. (2002). Agricultural sustainability and intensive production practices. *Nature*, 418:671 – 677.

Timmer, C. P. (1988). Chapter 8 The agricultural transformation. *Handbook of Development Economics*, 1:275 – 331.

Timmer, C. P. (2009). *A World Without Agriculture: The Structural Transformation in Historical Perspective*. The AEI Press.

Tittonell, P. (2014). Ecological intensification of agriculture - sustainable by nature. *Current Opinion in Environmental Sustainability*, 8:53 – 61.

Tittonell, P., Gérard, B., and Erenstein, O. (2015). Tradeoffs around crop residue biomass in smallholder crop-livestock systems - What's next? *Agricultural Systems*, 134:119 – 128.

Tittonell, P. and Giller, K. E. (2013). When yield gaps are poverty traps: The paradigm of ecological intensification in African smallholder agriculture. *Field Crops Research*, 143:76 – 90.

Tittonell, P., Shepherd, K. D., Vanlauwe, B., and Giller, K. E. (2008a). Unravelling the effects of soil and crop management on maize productivity in smallholder agricultural systems of western Kenya - An application of classification and regression tree analysis. *Agriculture, Ecosystems & Environment*, 123:137 – 150.

Tittonell, P., van Wijk, M. T., Herrero, M., Rufino, M. C., de Ridder, N., and Giller, K. E. (2009). Beyond resource constraints - Exploring the biophysical feasibility of options for the intensification of smallholder crop-livestock systems in Vihiga district, Kenya. *Agricultural Systems*, 101:1 – 19.

Tittonell, P., Vanlauwe, B., Corbeels, M., and Giller, K. E. (2008b). Yield gaps, nutrient use efficiencies and response to fertilisers by maize across heterogeneous smallholder farms of western Kenya. *Plant and Soil*, 313:19 – 37.

Tittonell, P., Zingore, S., van Wijk, M. T., Corbeels, M., and Giller, K. E. (2007). Nutrient use efficiencies and crop responses to N, P and manure applications in Zimbabwean soils: Exploring management strategies across soil fertility gradients. *Field Crops Research*, 100:348 – 368.

Valbuena, D., Tui, S. H.-K., Erenstein, O., Teufel, N., Duncan, A., Abdoulaye, T., Swain, B., Mekonnen, K., Germaine, I., and Gérard, B. (2015). Identifying determinants, pressures and trade-offs of crop residue use in mixed smallholder farms in sub-Saharan Africa and South Asia. *Agricultural Systems*, 134:107 – 118.

van Asten, P. J. A., Wopereis, M. C. S., Haefele, S., Isselmou, M. O., and Kropff, M. J. (2003). Explaining yield gaps on farmer-identified degraded and non-degraded soils in a Sahelian irrigated rice scheme. *NJAS - Wageningen Journal of Life Sciences*, 50:277 – 296.

van Bussel, L. G. J., Grassini, P., van Wart, J., Wolf, J., Claessens, L., Yang, H., Boogaard, H., de Groot, H., Saito, K., Cassman, K. G., and van Ittersum, M. K. (2015). From field to atlas: Upscaling of location-specific yield gap estimates. *Field Crops Research*, 177:98 – 108.

van den Berg, M. and Singels, A. (2013). Modelling and monitoring for strategic yield gap diagnosis in the South African sugar belt. *Field Crops Research*, 143:143 – 150.

van der Veen, H. B., Ge, L., van der Meer, R. W., and Vrolijk, H. C. J. (2014). Sample of Dutch FADN 2012: Design principles and quality of the sample of agricultural and horticultural holdings. Technical report, LEI Wageningen UR.

van Dijk, M., Morley, T., Jongeneel, R., van Ittersum, M. K., Reidsma, P., and Ruben, R. (2017). Disentangling agronomic and economic yield gaps: An integrated framework and application. *Agricultural Systems*, 154:90 – 99.

van Eerdewijk, A. and Danielsen, K. (2015). Gender matters in farm power. Technical report, Royal Tropical Institute of the Netherlands (KIT).

van Ittersum, M. K. and Cassman, K. G. (2013). Yield gap analysis - Rationale, methods and applications - introduction to the Special Issue. *Field Crops Research*, 143:1 – 3.

van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., and Hochman, Z. (2013). Yield gap analysis with local to global relevance - A review. *Field Crops Research*, 143:4 – 17.

van Ittersum, M. K., Leffelaar, P. A., van Keulen, H., Kropff, M. J., Bastiaans, L., and Goudriaan, J. (2003). On approaches and applications of the Wageningen crop models. *European Journal of Agronomy*, 18:201 – 234.

van Ittersum, M. K. and Rabbinge, R. (1997). Concepts in production ecology for analysis and quantification of agricultural input-output combinations. *Field Crops Research*, 52:197 – 208.

van Ittersum, M. K., van Bussel, L. G. J., Wolf, J., Grassini, P., van Wart, J., Guilpart, N., Claessens, L., de Groot, H., Wiebe, K., Mason-D'Croz, D., Yang, H., Boogaard, H., van Oort, P. A. J., van Loon, M. P., Saito, K., Adimo, O., Adjei-Nsiah, S., Agali, A., Bala, A., Chikowo, R., Kaizzi, K., Kouressy, M., Makoi, J. H. J. R., Ouattara, K., Tesfaye, K., and Cassman, K. G. (2016). Can sub-Saharan Africa feed itself? *Proceedings of the National Academy of Sciences*, 113:14964 – 14969.

van Keulen, H. (1982). Graphical analysis of annual crop response to fertiliser application. *Agricultural Systems*, 9:113 – 126.

van Noordwijk, M. and Brussaard, L. (2014). Minimizing the ecological footprint of food: closing yield and efficiency gaps simultaneously? *Current Opinion in Environmental Sustainability*, 8:62 – 70.

van Oort, P. A. J., Saito, K., Dieng, I., Grassini, P., Cassman, K. G., and van Ittersum, M. K. (2016). Can yield gap analysis be used to inform R&D prioritisation? *Global Food Security*, 12:109 – 118.

van Oort, P. A. J., Timmermans, B. G. H., Meinke, H., and van Ittersum, M. K. (2012). Key weather extremes affecting potato production in The Netherlands. *European Journal of Agronomy*, 37:11 – 22.

van Vliet, J. A., Schut, A. G. T., Reidsma, P., Descheemaeker, K., Slingerland, M., van de Ven, G. W. J., and Giller, K. E. (2015). De-mystifying family farming: Features, diversity and trends across the globe. *Global Food Security*, 5:11 – 18.

van Wart, J., van Bussel, L. G. J., Wolf, J., Licker, R., Grassini, P., Nelson, A., Boogaard, H., Gerber, J., Mueller, N. D., Claessens, L., van Ittersum, M. K., and Cassman, K. G. (2013). Use of agro-climatic zones to upscale simulated crop yield potential. *Field Crops Research*, 143:44 – 55.

van Wijk, M. T., Rufino, M. C., Enahoro, D., Parsons, D., Silvestri, S., Valdivia, R. O., and Herrero, M. (2014). Farm household models to analyse food security in a changing climate: A review. *Global Food Security*, 3:77 – 84.

Vanlauwe, B., Coe, R., and Giller, K. E. (2016). Beyond averages: New approaches to understand heterogeneity and risk of technology success or failure in smallholder farming. *Experimental Agriculture*, pages 1 – 23.

Vanlauwe, B., Coyne, D., Gockowski, J., Hauser, S., Huising, J., Masso, C., Nziguheba, G., Schut, M., and van Asten, P. J. A. (2014). Sustainable intensification and the African smallholder farmer. *Current Opinion in Environmental Sustainability*, 8:15 – 22.

Vanlauwe, B., Descheemaeker, K., Giller, K. E., Huising, J., Merckx, R., Nziguheba, G., Wendt, J., and Zingore, S. (2015). Integrated soil fertility management in sub-Saharan Africa: unravelling local adaptation. *SOIL*, 1:491 – 508.

Vanlauwe, B., Tittonell, P., and Mukalama, J. (2007). *Advances in Integrated Soil Fertility Management in sub-Saharan Africa: Challenges and Opportunities*, chapter Within-farm soil fertility gradients affect response of maize to fertiliser application in western Kenya, pages 121–132. Springer Netherlands, Dordrecht.

Vereijken, P. (1997). A methodical way of prototyping integrated and ecological arable farming systems (I/EAFS) in interaction with pilot farms. *European Journal of Agronomy*, 7:235 – 250.

Verma, R. (2001). *Gender, land and livelihoods: Through farmers' eyes*. International Development Research Centre.

Villano, R., Bravo-Ureta, B., Solis, D., and Fleming, E. (2015). Modern rice technologies and productivity in the Philippines: Disentangling technology from managerial gaps. *Journal of Agricultural Economics*, 66:129 – 154.

Vitousek, P. M., Naylor, R., Crews, T., David, M. B., Drinkwater, L. E., Holland, E., Johnes, P. J., Katzenberger, J., Martinelli, L. A., Matson, P. A., Nziguheba, G., Ojima, D., Palm, C. A., Robertson, G. P., Sanchez, P. A., Townsend, A. R., and Zhang, F. S. (2009). Nutrient imbalances in agricultural development. *Science*, 324:1519 – 1520.

Wairegi, L. W. I., van Asten, P. J. A., Tenywa, M. M., and Bekunda, M. A. (2010). Abiotic constraints override biotic constraints in East African highland banana systems. *Field Crops Research*, 117:146 – 153.

Wandera, F., Akulumuka, V., Maina, J., Otiang, J., Mubiru, J. F., Mbugua, D., and Muinga, R. (2015). Regional specialisation for enhanced agricultural productivity and transformation. Proceedings of the eastern Africa agricultural productivity program end-of-phase one project conference and exhibition. Technical report, Association for Strengthening Agricultural Research in Eastern and Central Africa (ASARECA).

WB (2017). World development indicators 2011. Poverty headcount ratio at \$1.90 a day is the percentage of the population living on less than \$1.90 a day at 2011 international prices. The World Bank. Retrieved from http://databank.worldbank.org/data/home.aspx.

Webb, R. A. (1972). Use of the boundary line in the analysis of biological data. *Journal Horticultural Science*, 47:309 – 319.

Wijnands, F. G., Diaz, A. G., and de Haan, J. J. (2002). *Manual on prototyping methodology and multifunctional crop rotation*, chapter Multifunctional Crop Rotation (MCR), pages 42 – 50. Applied Plant Research B.V. (VEGINECO report 2).

Witt, C., Dobermann, A., Abdulrachman, S., Gines, H. C., Guanghuo, W., Nagarajan, R., Satawatananont, S., Son, T. T., Tan, P. S., Tiem, L. V., Simbahan, G. C., and

Olk, D. C. (1999). Internal nutrient efficiencies of irrigated lowland rice in tropical and subtropical Asia. *Field Crops Research*, 63:113 – 138.

Wolf, J., Mandryk, M., Kanellopoulos, A., van Oort, P., Schaap, B. F., Reidsma, P., and van Ittersum, M. K. (2011). Integrated assessment of adaptation to climate change in Flevoland at farm and regional levels. AgriAdapt Reports 4&5. Technical report, Wageningen University and Research Centre.

Woodhouse, P. (2010). Beyond industrial agriculture? Some questions about farm size, productivity and sustainability. *Journal of Agrarian Change*, 10:437 – 453.

Wopereis, M. C. S., Donovan, C., Nebie, B., Guindo, D., and N'Diaye, M. K. (1999). Soil fertility management in irrigated rice systems in the Sahel and Savanna regions of West Africa: Part I. Agronomic analysis. *Field Crops Research*, 61:125 – 145.

Workayehu, T. and Wortmann, C. S. (2011). Maize - bean intercrop weed suppression and profitability in Southern Ethiopia. *Agronomy Journal*, 103:1058 – 1063.

Xiao, D., Feng, J., Wang, N., Luo, X., and Hu, Y. (2013). Integrated soil moisture and water depth sensor for paddy fields. *Computers and Electronics in Agriculture*, 98:214 – 221.

Yan, L. and Roy, D. P. (2014). Automated crop field extraction from multi-temporal Web Enabled Landsat Data. *Remote Sensing of Environment*, 144:42 – 64.

Yanalak, M., Musaoglu, N., Ipbuker, C., Sertel, E., and Kaya, S. (2012). *DEM Accuracy of High Resolution Satellite Images*, pages 471 – 480. Springer Berlin Heidelberg.

Zhang, C., Walters, D., and Kovacs, J. M. (2014). Applications of low altitude remote sensing in agriculture upon farmers' requests - A case study in Northeastern Ontario, Canada. *PLOS ONE*, 9:1 – 9.

Zhang, X., Davidson, E. A., Mauzerall, D. L., Searchinger, T. D., Dumas, P., and Shen, Y. (2015). Managing nitrogen for sustainable development. *Nature*, 528:51 – 59.

Zheng, B., Campbell, J. B., Serbin, G., and Galbraith, J. M. (2014). Remote sensing of crop residue and tillage practices: Present capabilities and future prospects. *Soil and Tillage Research*, 138:26 – 34.

Zhengfei, G. and Lansink, A. O. (2006). The source of productivity growth in Dutch agriculture: A perspective from finance. *American Journal of Agricultural Economics*, 88:644 – 656.

Zhengfei, G., Lansink, A. O., van Ittersum, M. K., and Wossink, A. (2006). Integrating agronomic principles into production function specification: A dichotomy of growth inputs and facilitating inputs. *American Journal of Agricultural Economics*, 88:203 – 214.

Annex A. Review of yield gap explaining factors (Chapter 2)

Table A1: Studies included in the meta-analysis, showing the crop types, the scale at which the yield gap was estimated, the scale at which the purpose (yield gap or yield variability) was explained, the level at which data was collected, the country where the study was conducted, years included in the analysis, benchmark yield and magnitude of the yield gap (only recorded if explicitly provided in the paper). *Codes*: Ya = actual yield, Yp = Potential yield, Yw = water-limited-yield, Yatt = attainable yield.

Crop	*Scale Yg estimation estimation*	*Scale Yg explanation*	*Scale of data collection*	*Location Country*	*Period analysed*	*Benchmark yield*	*Yield gap (%)*	*Reference*
Rice	Irrigation Scheme	Irrigation Scheme	Field	Mauritania	1998 - 1999	Yp	49 - 62	van Asten et al. (2003)
Rice	Watershed	Watershed	Field	Côte d'Ivoire	1995 - 1996	Yp	–	Becker and Johnson (1999)
Rice	Irrigation Scheme	Adm. Region	Field	Benin	2011	–	–	Tanaka et al. (2013)
Rice	AE Region	AE Region	Field	Côte d'Ivoire	1995 - 1999	Yp	40 - 63	Becker et al. (2003)
Rice	–	Irrigation scheme	Field	Mauritania	1997 - 1998	Yp, Yatt	14 - 80	Haefele et al. (2001)
Rice	Irrigation Scheme	Irrigation Scheme	Plot	West Africa	1995 - 1996	Yp, Yatt	29 - 47	Wopereis et al. (1999)
Rice	–	Site	Plot	Africa	1994 - 1998	Yatt	–	Audebert and Fofana (2009)
Rice	Adm. Region	Adm. Region	Farm	Philippines	1994 - 1999	Yp, Yecon, Yatt	38 - 55	Laborte et al. (2012)
Rice	Adm. Region	Adm. Region	Farm	Vietnam	1994 - 1999	Yp, Yecon, Yatt	21 - 37	Laborte et al. (2012)
Rice	Adm. Region	Adm. Region	Farm	Thailand	1994 - 1999	Yp, Yecon, Yatt	20 - 38	Laborte et al. (2012)
Rice	Adm. Region	Adm. Region	Farm	Indonesia	1994 - 1999	Yp, Yecon, Yatt	24 - 39	Laborte et al. (2012)
Rice	Province	Province	Province	Philippines	1985 - 2002	Yp	35 - 63	Angulo et al. (2012)
Rice	Toposequence	Toposequence	Plot	Indonesia	2001 - 2002	Yp	–	Boling et al. (2010)
Rice	Toposequence	Toposequence	Plot	Thailand	2000 - 2002	Yatt	–	Boling et al. (2011)
Rice	Delta	Delta	Field	Spain	1995 - 1996	Yp	34	Casanova et al. (1999)
Rice	AE Region	AE Region	Field	France	1992 & 2009	Yatt	37 - 57	Delmotte et al. (2011)
Rice	District	District	Plot	Vietnam	1999 - 2003	Yatt	–	Mussgnug et al. (2006)
Rice	Region	Region	5 arc min	Globe	2000	Yatt	–	Neumann et al. (2010)
Rice	Globe	Globe	5 arc min	Globe	2000	Yatt	40	Licker et al. (2010)
Rice	AE Region	AE Region	Field	Vietnam	1990 - 1991	Yp, Yw, Yatt	47 - 68	Affholder et al. (2012)
Wheat	Experimental site	–	Field	China	2003 - 2007	Yp	–	Lu and Fan (2013)
Wheat	AE Region	AE Region	District	India	1971 - 1993	Yp, Yatt	–	Aggarwal and Kalra (1994)
Wheat	Village	Village	Field	Spain	2003 - 2006	Yp, Yatt	–	Abeledo et al. (2008)
Wheat	Field	Country	Field	France	1995 - 2003	Yp	–	Prost et al. (2008)
Wheat	EU	–	NUTS	Europe	1990 - 2006	Yp, Yw	–	Boogaard et al. (2013)

Table A1: Studies included in the meta-analysis, showing the crop types, the scale at which the yield gap was estimated, the scale at which the purpose (yield gap or yield variability) was explained, the level at which data was collected, the country where the study was conducted, years included in the analysis, benchmark yield and magnitude of the yield gap (only recorded if explicitly provided in the paper). *Codes*: Ya = actual yield, Yp = Potential yield, Yw = water-limited-yield, Yatt = attainable yield. (*continued*)

Crop	***Scale Yg estimation estimation***	***Scale Yg explanation***	***Scale of data collection***	***Location Country***	***Period analysed***	***Benchmark yield***	***Yield gap (%)***	***Reference***
Wheat	–	AE Region	Field	Mexico	2001 & 2003	–	–	Lobell et al. (2005)
Wheat	–	Zone	Field	Argentina	1994 - 1999	Yp, Yatt	–	Calvino and Sadras (2002)
Wheat	Farm	–	Field	Australia	2004 - 2009	Yw	31 - 71	Oliver and Robertson (2013)
Wheat	–	AE Region	Farm	Australia	1998 - 2000	–	–	Sadras et al. (2002)
Wheat	Adm. Region	–	Farm	Australia	1996 - 2010	Yw	25 - 65	Hochman et al. (2012)
Wheat	County	AE Region	Farm	China	2004 - 2005	Yp, Yatt	26 - 46	li Liang et al. (2011)
Wheat	Region	Region	5 arc min	Globe	2000	Yatt	–	Neumann et al. (2010)
Wheat	Globe	Globe	5 arc min	Globe	2000	Yatt	60	Licker et al. (2010)
Maize	–	AE Region	Field	Kenya	2002	–	–	Tittonell et al. (2008a)
Maize	Site	Field	Field	Kenya	2002	Yw	10 - 82	Tittonell et al. (2008b)
Maize	Site	Site	Site	Ethiopia	1992 - 2003	Yp, Yw, Yatt	36 - 77	Kassie et al. (2014)
Maize	District	–	Field	Bangladesh	2010 - 2011	Yp	34 - 41	Schulthess et al. (2012)
Maize	Region	–	Farm	China	2007 - 2008	Yp	32 - 59	Meng et al. (2012)
Maize	Site	Site	Field	Philippines	2004 - 2008	Yp, Yw	10 - 16	Pasuquin et al. (2014)
Maize	Site	Site	Field	Indonesia	2004 - 2008	Yp, Yw	5 - 19	Pasuquin et al. (2014)
Maize	Site	Site	Field	Vietnam	2004 - 2008	Yp, Yw	3 - 27	Pasuquin et al. (2014)
Maize	EU	EU	Subnational	Europe	1990 - 2003	Yp, Yw	–	Reidsma et al. (2009a)
Maize	Country	District	Field	USA	2005 - 2007	Yp	–	Grassini et al. (2011a)
Maize	–	Experimental site	Field	Canada	2003 - 2005	Yatt	–	Subedi and Ma (2009)
Maize	District	District	Plot	Vietnam	1999 - 2003	Yatt	–	Mussgnug et al. (2006)
Maize	County	AE Region	Farm	China	2004 - 2005	Yp, Yatt	27 - 46	li Liang et al. (2011)
Maize	Region	Region	5 arc min	Globe	2000	Yatt	–	Neumann et al. (2010)
Maize	Globe	Globe	5 arc min	Globe	2000	Yatt	50	Licker et al. (2010)
Maize	AE Region	AE Region	Field	Brazil	1990 - 1991	Yp, Yw, Yatt	28 - 45	Affholder et al. (2012)
Maize	AE Region	AE Region	Field	Vietnam	1990 - 1991	Yp, Yw, Yatt	38 - 70	Affholder et al. (2012)

Table A1: Studies included in the meta-analysis, showing the crop types, the scale at which the yield gap was estimated, the scale at which the purpose (yield gap or yield variability) was explained, the level at which data was collected, the country where the study was conducted, years included in the analysis, benchmark yield and magnitude of the yield gap (only recorded if explicitly provided in the paper). *Codes*: Ya = actual yield, Yp = Potential yield, Yw = water-limited-yield, Yatt = attainable yield. (*continued*)

Crop	*Scale Yg estimation estimation*	*Scale Yg explanation*	*Scale of data collection*	*Location Country*	*Period analysed*	*Benchmark yield*	*Yield gap (%)*	*Reference*
Soybean	Village	Village	Field	India	1989 - 2003	Yp, Yw	15 - 77	Bhatia et al. (2008)
Soybean	–	AE Region	Field	Argentina	2001 - 2005	–	–	Bacigaluppo et al. (2011)
Soybean	District	District	Plot	Vietnam	1999 - 2003	Yatt	–	Mussgnug et al. (2006)
Soybean	Globe	Globe	5 arc min	Globe	2000	Yatt	20	Licker et al. (2010)
Cassava	Village	Country	Farm	Uganda	2004 - 2005	Yatt	64	Fermont et al. (2009)
Cassava	Village	Country	Farm	Kenya	2004 - 2005	Yatt	60 - 68	Fermont et al. (2009)
Cassava	AE Region	AE Region	Field	Cambodia	2009 - 1010	Yatt	43	Sopheap et al. (2012)
Banana	AE Region	AE Region	Field	Uganda	2006 - 2007	Yatt	40 - 58	Wairegi et al. (2010)
Banana	District	District	Field	Kenya	2005-2006	Yatt	–	Okumu et al. (2011)
Millet	AE Region	AE Region	Field	Senegal	1990 - 1991	Yp, Yw, Yatt	49 - 73	Affholder et al. (2012)
Tomato	n.a	Island	Plot	France	2003 & 2005	–	–	Huat et al. (2013)
Peanut	Field	Field	Field	Ghana	1997 - 1998	Yp, Yw	–	Naab et al. (2004)
Quinoa	–	AE Region	Site	Bolivia	1970 - 2003	Yp	50 - 87	Geerts et al. (2009)
Sunflower	AE Region	–	Field	Argentina	1999 - 2007	Yatt	17 - 67	Hall and Richards (2013)
Sugarcane	AE Region	AE Region	Mill area	South Africa	1988 - 2010	Yp	–	van den Berg and Singels (2013)
Mango	District	District	Field	Thailand	1993	Yatt	91	de Bie (2004)
Potato	Experimental site	–	Site	China	Not specified	Yp	44 - 75	He et al. (1998)

Table A2: Overview of the different methods used to explain yield gap and/or yield variability (var.), showing management and edaphic factors used by the different methods. *Codes*: Yg = Yield gap, Ya = Actual yield, Yw = Water-limited yield, Yp att. = Attainable potential yield, P = Planting, F = Fertilization, Cc = Crop characteristics, Cp = Crop protection, W = Weeding, I = Irrigation, Lp = Land preparation, O = Others, Sf = Soil fertility, St = Soil type, Sw = Soil water, S = Slope.

Explanatory approach	Purpose of the study	Number of studies	Number of records	Management								Edaphic				Farm characteristics					Socio-economic		
				P	F	Cc	Cp	W	I	Lp	O	Sf	St	Sw	S	S	L	I	In	T	Ins	Te	P
ANOVA	Yw var., Yg & Ya var.	9	26	+	+	+	+	+	+	−	−	+	+	+	+	+	−	−	−	−	−	−	−
Chi−square test	Ya var.	1	2	+	−	+	−	+	−	+	−	+	−	−	−	+	+	+	−	−	−	−	−
Comparative performance	Yg	1	1	−	−	+	+	+	+	−	−	+	−	+	+	−	−	−	−	−	−	−	−
Difference between groups	Yg, Ya var.	6	46	+	+	−	+	+	+	−	−	+	−	−	−	+	+	−	−	+	+	+	−
Spearman test	Ya var.	1	3	+	+	−	+	+	−	+	−	+	+	−	−	−	−	−	−	−	−	−	−
Wilcoxon test	Ya var.	1	1	+	+	+	+	−	+	+	−	−	−	−	−	−	−	−	−	−	−	−	−
AquaCrop	Water productivity	1	3	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−
CERES−Wheat	Yg, Yp var. & Yatt var.	1	4	−	+	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
CROPGRO−Peanut	Yg	1	2	+	−	+	+	−	−	−	−	+	+	+	−	−	−	−	−	−	−	−	−
CROPGRO−Soybean	Yg	1	42	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
DSSAT CSM−IXIM	Yw var. & Yp var.	1	13	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
ORYZA2000	Yg	2	8	−	+	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−
QUEFTS	Yg	1	3	+	+	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−
WTGROWS	Yp var., Yatt var.	1	6	+	+	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Correlation	Ya var., Yg & Yw var.	3	8	+	−	+	−	+	−	−	−	+	+	−	−	+	−	−	−	+	−	−	−
Pearson bivariate correlation	Ya var.	1	2	+	+	+	−	−	−	−	−	+	−	−	−	+	−	−	−	−	−	−	−
Pearson correlation	Ya var.	3	3	+	−	+	+	−	+	−	−	+	+	+	+	+	−	+	−	−	−	−	−
Linear correlation	Ya var.	1	1	−	−	−	−	−	−	−	−	+	+	+	−	−	−	−	−	−	−	−	−
Forward stepwise regression	Yg	1	6	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	+	−	+	−	+
Inefficiency model	Yg	1	3	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	+	−	+	−	+
Linear regression	Ya var.	6	9	+	+	+	+	+	+	+	−	+	+	+	−	−	−	−	−	−	−	−	−
Multiple linear regression	Yp var., Ya var. & Yg	4	6	+	+	+	+	−	−	−	−	+	+	+	−	−	−	−	−	−	−	−	−
Multiple regression	Ya var.	2	5	+	+	+	−	+	+	−	−	+	−	+	−	−	−	−	−	−	−	−	−
Stepwise linear regression	Ya var. & Yg	2	5	−	+	−	+	−	−	+	−	+	+	+	−	+	−	+	+	−	−	−	−
Stepwise multiple regression	Ya var. & Yg	5	7	+	+	+	+	+	+	+	−	+	+	+	+	+	+	−	−	+	−	−	+
Boundary line	Yg	6	12	+	+	−	+	+	−	+	−	+	+	+	−	−	−	−	−	−	−	−	−
Bayesian model averaging	Yg	1	2	+	+	+	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
CART	Ya var.	3	7	+	+	+	+	+	+	+	+	+	+	+	+	+	−	+	+	−	−	−	−
Qualitative	Yg & Ya var.	5	11	+	+	+	+	+	+	+	+	+	+	−	−	+	+	+	−	−	+	+	−
Principal components analysis	Ya var.	1	2	+	+	+	+	+	−	−	−	+	−	+	−	−	−	−	−	−	−	−	−

Annex B. Rice farming systems in Central Luzon, the Philippines (Chapters 3 and 4)

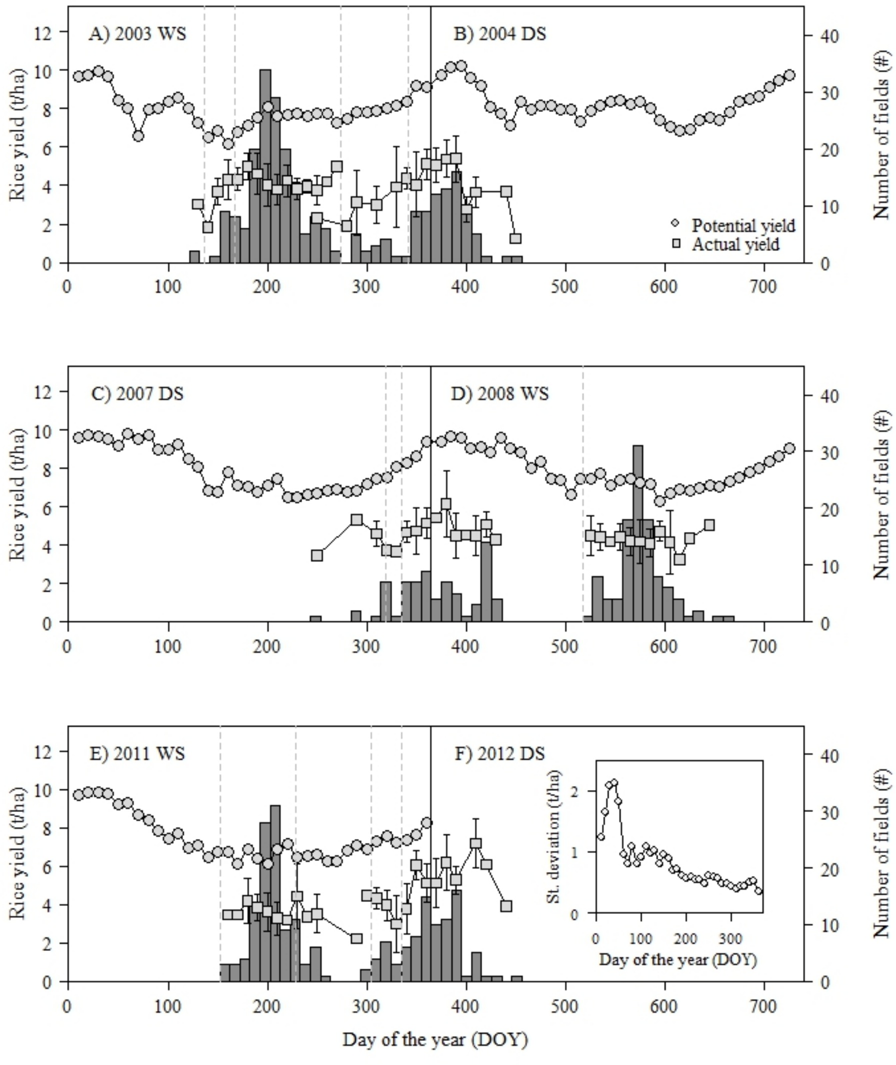

Figure B1: Annual variability of the climatic potential and actual yields (t ha^{-1}) and distribution of the transplanting dates observed in the household survey during the period A) 2003 WS, B) 2004 DS, C) 2007 DS, D) 2008 WS, E) 2011 WS and F) 2012 DS. Dashed lines represent the first and last date of water release by the National Irrigation Administration (NIA) in a particular season. Error bars show the standard deviation of Ya for a particular transplanting date and the inset in F) shows the standard deviation of Yp over the period 1991 - 2012.

Table B1: Descriptive statistics of the Central Luzon Loop Survey for the wet seasons (WS) surveyed between 1979 - 2011. Extreme observations of yield and input use (± 2STD) were removed prior to the analysis. Standard deviations are presented between brackets. [a] At 2005 prices; 1 US$ = 44.8 ₱(Central Bank of the Philippines, 19.10.2014). [b] Application of insecticides, herbicides, molluscicides, fungicides and hand weeding.

	1979		**1982**		**1986**		**1990**		**1994**		**1999**		**2003**		**2008**		**2011**	
Sample size (n)																		
Rice fields	203		218		137		160		131		101		155		92		103	
Households	134		133		109		102		95		78		108		91		81	
Cultivated land (ha)																		
Farm size	1.1	(0.7)	0.9	(0.6)	1.3	(0.7)	1.1	(0.6)	1.2	(0.7)	1.1	(0.5)	1.1	(0.6)	1.0	(0.5)	1.1	(0.6)
Rice yield (t ha^{-1})																		
Variety type Mv1-3	3.3	(1.4)	3.8	(1.3)	3.2	(0.9)	3.2	(1.1)	3.8	(1.3)	2.9	(1.2)	3.8	(0.9)	3.8	(1.0)	3.3	(1.3)
Variety type Mv4	n.a.		n.a.		n.a.		n.a.		n.a.		3.3	(1.1)	4.0	(0.9)	4.2	(0.8)	3.5	(0.9)
Input use (kg ha^{-1})																		
Seeds	98.3	(33.0)	110.9	(44.8)	126.4	(50.7)	131.9	(44.6)	148.5	(54.2)	151.0	(52.4)	130.0	(48.9)	111.9	(34.5)	102.2	(35.6)
Nitrogen (N)	60.9	(30.8)	68.4	(37.4)	68.1	(32.5)	70.3	(32.7)	95.1	(39.8)	90.9	(40.7)	91.2	(35.8)	89.7	(34.4)	103.6	(50.7)
Phosphorus (P)	14.2	(12.0)	12.9	(12.8)	12.9	(11.9)	18.1	(13.9)	19.0	(15.0)	22.5	(15.8)	28.1	(17.6)	22.4	(14.6)	28.5	(18.2)
Potassium (K)	7.3	(10.6)	4.3	(7.7)	3.9	(7.0)	10.9	(11.1)	12.6	(13.6)	13.7	(12.1)	18.7	(15.3)	13.1	(10.2)	19.6	(16.2)
Insecticide (a.i.)	0.6	(0.6)	0.4	(0.4)	0.3	(0.3)	0.3	(0.3)	0.3	(0.3)	0.1	(0.2)	0.2	(0.2)	0.2	(0.2)	0.2	(0.2)
Herbicide (a.i.)	0.3	(0.2)	0.3	(0.3)	0.3	(0.2)	0.3	(0.2)	0.5	(0.3)	0.4	(0.2)	0.4	(0.2)	0.4	(0.3)	0.4	(0.3)
Input costs[a] (k₱ha^{-1})																		
Seeds	1.2	(0.7)	1.2	(0.5)	1.6	(1.0)	2.2	(0.9)	2.1	(0.9)	2.1	(0.8)	1.8	(0.7)	2.1	(1.0)	1.9	(0.9)
Irrigation water	0.0	(0.2)	0.0	(0.1)	1.3	(1.4)	1.1	(1.2)	0.8	(0.9)	0.6	(0.9)	0.8	(0.7)	1.1	(0.9)	1.0	(1.6)
Fertilisers	4.0	(2.0)	3.9	(2.0)	2.4	(1.3)	3.2	(1.4)	3.2	(1.3)	2.8	(1.3)	3.7	(1.4)	9.3	(3.3)	6.5	(2.7)
Insecticide	0.9	(0.5)	0.8	(0.5)	0.9	(0.6)	0.7	(0.5)	0.8	(0.5)	0.7	(0.5)	0.5	(0.3)	0.5	(0.3)	0.5	(0.3)
Herbicide	0.6	(0.3)	0.5	(0.3)	0.4	(0.3)	0.5	(0.4)	0.6	(0.4)	0.6	(0.3)	0.5	(0.4)	0.5	(0.3)	0.5	(0.3)
No. of operations (#)																		
Land preparation	3.0	(0.0)	3.0	(0.0)	3.3	(0.4)	4.0	(0.8)	3.8	(0.9)	4.0	(0.8)	3.7	(0.7)	4.6	(0.5)	4.5	(0.5)
Crop establishment	3.3	(0.5)	3.2	(0.6)	4.3	(0.5)	3.9	(0.9)	3.9	(1.0)	4.0	(0.9)	4.2	(0.7)	4.4	(0.6)	4.3	(0.5)
Fertiliser application	1.4	(0.5)	1.4	(0.5)	1.6	(0.5)	1.5	(0.5)	1.5	(0.5)	1.7	(0.5)	1.8	(0.6)	1.9	(0.6)	2.0	(0.7)
Pest & weed control[b]	3.2	(1.7)	3.2	(1.5)	3.4	(1.3)	2.7	(1.2)	2.9	(1.2)	2.1	(1.0)	2.6	(1.2)	2.9	(1.2)	2.6	(1.2)
Total labour (ld ha^{-1})																		
Land preparation	14.1	(7.2)	14.2	(6.8)	14.4	(5.6)	12.5	(4.9)	11.1	(4.1)	9.8	(3.3)	9.4	(4.5)	11.1	(3.9)	11.8	(4.8)
Crop establishment	26.5	(7.7)	25.6	(11.8)	26.4	(8.7)	25.0	(13.7)	22.8	(13.3)	20.1	(10.4)	21.0	(10.8)	25.3	(5.5)	27.4	(10.2)
Fertiliser application	1.3	(0.8)	1.2	(0.7)	1.3	(0.7)	1.3	(0.7)	1.5	(0.9)	1.1	(0.5)	1.5	(0.8)	1.6	(0.9)	1.4	(0.7)
Pest & weed control[b]	7.6	(6.5)	4.9	(5.1)	3.9	(3.6)	2.9	(2.9)	3.2	(3.0)	1.3	(0.8)	2.5	(2.2)	1.9	(1.7)	2.2	(2.5)
Harvest & threshing	28.3	(11.1)	28.5	(13.6)	21.1	(6.9)	27.3	(11.2)	27.4	(12.2)	26.0	(7.4)	23.5	(8.8)	25.9	(4.5)	23.5	(10.5)

Table B2: Descriptive statistics of the Central Luzon Loop Survey for the dry seasons (DS) surveyed between 1980 - 2012. Extreme observations of yield and input use (± 2STD) were removed prior to the analysis. Standard deviations are presented between brackets. [a] At 2005 prices; 1 US$ = 44.8 ₱(Central Bank of the Philippines, 19.10.2014). [b] Application of insecticides, herbicides, molluscicides, fungicides and hand weeding.

	1980		**1987**		**1991**		**1995**		**1998**		**2004**		**2007**		**2012**	
Sample size (n)																
Rice fields	109		82		83		70		58		94		50		83	
Households	80		63		57		53		43		68		50		61	
Cultivated land (ha)																
Farm size	1.3	(0.8)	1.3	(0.7)	1.2	(0.6)	1.2	(0.7)	1.1	(0.5)	1.2	(0.7)	1.1	(0.6)	1.2	(0.7)
Rice yield ($t\ ha^{-1}$)																
Variety type Mv1-3	3.9	(1.5)	3.8	(1.2)	4.0	(1.3)	4.6	(1.6)	4.0	(1.5)	3.9	(1.6)	4.7	(1.0)	4.2	(1.4)
Variety type Mv4	n.a.		n.a.		n.a.		n.a.		4.8	(0.8)	4.8	(1.0)	4.8	(1.0)	5.2	(1.2)
Input use ($kg\ ha^{-1}$)																
Seeds	108.0	(36.7)	174.7	(62.1)	179.5	(63.2)	179.7	(51.2)	167.6	(50.5)	141.7	(42.7)	125.7	(42.3)	89.4	(50.5)
Nitrogen (N)	88.6	(44.3)	96.4	(40.3)	102.1	(45.1)	126.6	(48.1)	98.8	(39.5)	106.8	(32.6)	101.4	(31.8)	108.1	(40.1)
Phosphorus (P)	16.5	(15.2)	15.0	(13.3)	19.7	(18.8)	29.0	(18.1)	23.9	(16.6)	35.6	(20.7)	29.8	(17.1)	29.6	(17.4)
Potassium (K)	9.4	(12.9)	7.6	(10.9)	14.8	(16.3)	19.4	(17.9)	14.0	(13.7)	20.5	(16.7)	18.1	(13.8)	23.3	(16.0)
Insecticide (a.i.)	0.4	(0.4)	0.3	(0.2)	0.3	(0.2)	0.2	(0.2)	0.1	(0.2)	0.2	(0.2)	0.2	(0.2)	0.2	(0.2)
Herbicide (a.i.)	0.2	(0.3)	0.3	(0.2)	0.3	(0.2)	0.5	(0.3)	0.5	(0.3)	0.5	(0.3)	0.4	(0.3)	0.4	(0.3)
Input costs[a] (k₱ha^{-1})																
Seeds	1.3	(0.5)	2.2	(0.8)	2.3	(0.9)	2.3	(0.9)	2.5	(0.8)	1.9	(0.6)	2.3	(0.9)	2.2	(1.0)
Irrigation water	1.6	(1.8)	2.1	(1.7)	1.9	(1.6)	0.8	(0.9)	1.8	(2.4)	1.5	(1.2)	2.2	(2.1)	1.7	(2.0)
Fertilisers	4.7	(2.2)	3.0	(1.4)	5.1	(2.2)	4.6	(1.8)	3.1	(1.3)	4.7	(1.3)	6.8	(2.6)	6.5	(2.3)
Insecticide	1.0	(0.6)	1.1	(0.6)	0.7	(0.3)	0.6	(0.4)	0.7	(0.5)	0.6	(0.4)	0.6	(0.4)	0.4	(0.3)
Herbicide	0.6	(0.3)	0.6	(0.3)	0.8	(0.3)	0.7	(0.3)	0.6	(0.3)	0.6	(0.3)	0.6	(0.3)	0.5	(0.3)
No. of operations (#)																
Land preparation	3.0	(0.0)	3.4	(0.5)	4.8	(0.4)	4.3	(0.8)	4.1	(0.8)	4.3	(0.8)	4.6	(0.5)	4.7	(0.5)
Crop establishment	3.2	(0.4)	3.8	(1.1)	3.1	(1.1)	2.5	(1.3)	2.8	(1.6)	2.8	(1.4)	3.1	(1.6)	3.7	(1.4)
Fertiliser application	1.7	(0.5)	1.2	(0.4)	2.0	(0.7)	2.0	(0.7)	2.1	(0.7)	2.3	(0.7)	2.4	(0.5)	2.2	(0.6)
Pest & weed control[b]	3.6	(1.6)	1.7	(1.2)	3.0	(1.3)	2.5	(1.2)	2.4	(1.0)	2.7	(1.2)	3.3	(1.4)	3.0	(1.4)
Total labour ($ld\ ha^{-1}$)																
Land preparation	12.0	(5.3)	12.7	(5.6)	12.2	(4.2)	12.9	(6.0)	9.8	(3.6)	10.3	(3.2)	9.5	(3.1)	9.4	(4.3)
Crop establishment	30.6	(9.4)	17.1	(13.1)	9.3	(9.2)	11.3	(12.3)	11.5	(10.7)	9.9	(9.2)	11.7	(11.2)	16.7	(11.3)
Fertiliser application	1.5	(0.7)	1.4	(0.8)	1.5	(0.8)	1.9	(1.1)	1.2	(0.6)	1.8	(0.9)	1.5	(0.7)	1.7	(0.9)
Pest & weed control[b]	8.6	(7.6)	4.1	(3.3)	3.1	(3.7)	2.4	(4.3)	1.4	(0.8)	2.2	(2.2)	2.1	(1.4)	1.5	(1.2)
Harvest & threshing	29.8	(12.3)	27.4	(8.3)	29.8	(11.0)	34.5	(17.4)	24.9	(4.9)	24.8	(7.1)	25.2	(5.7)	23.3	(8.6)

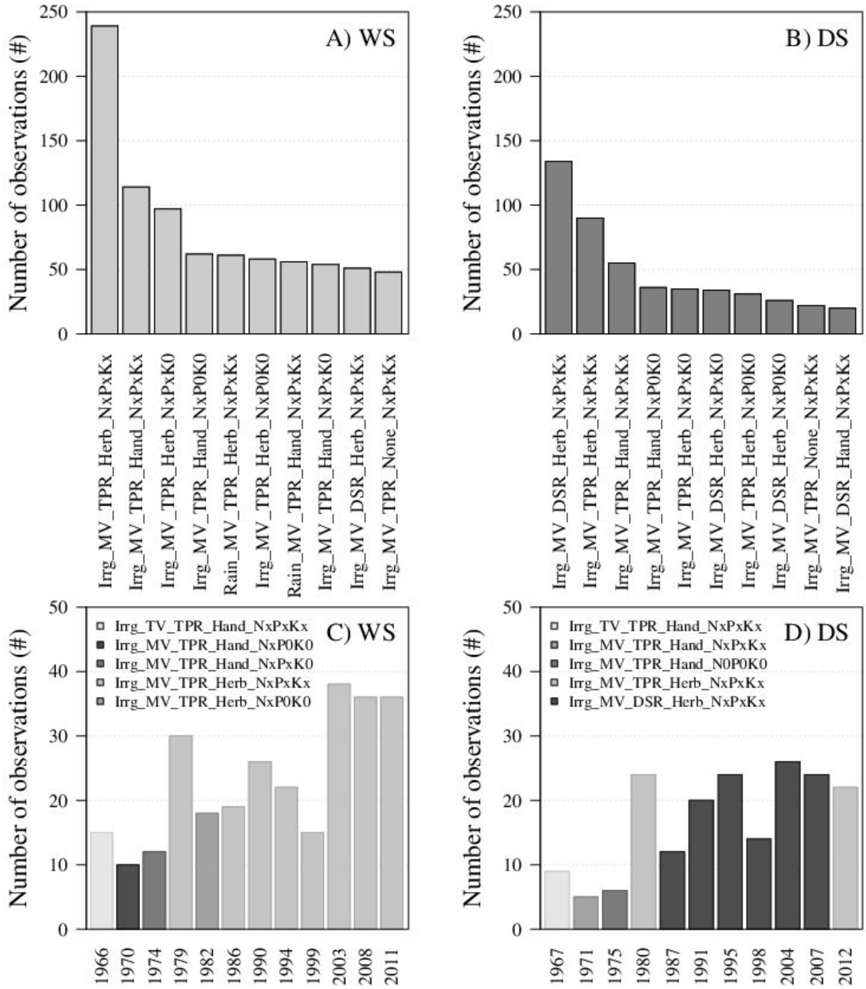

Figure B2: Overview of the management packages used in rice-based farming systems during the past half-century in Central Luzon, Philippines. A) and B) the 10 most used management packages in the WS and DS, respectively. C) and D) show the most commonly used package in each year during the WS and DS, respectively. Codes: 'Irrg' = Irrigation, 'Rain' = rainfed, 'MV' = modern variety, 'TV' = traditional variety, 'DSR' = direct-seeding, 'TPR' = transplanting, 'Hand' = hand-weeding, 'Herb' = herbicide use. *Data source*: Central Luzon Loop Survey.

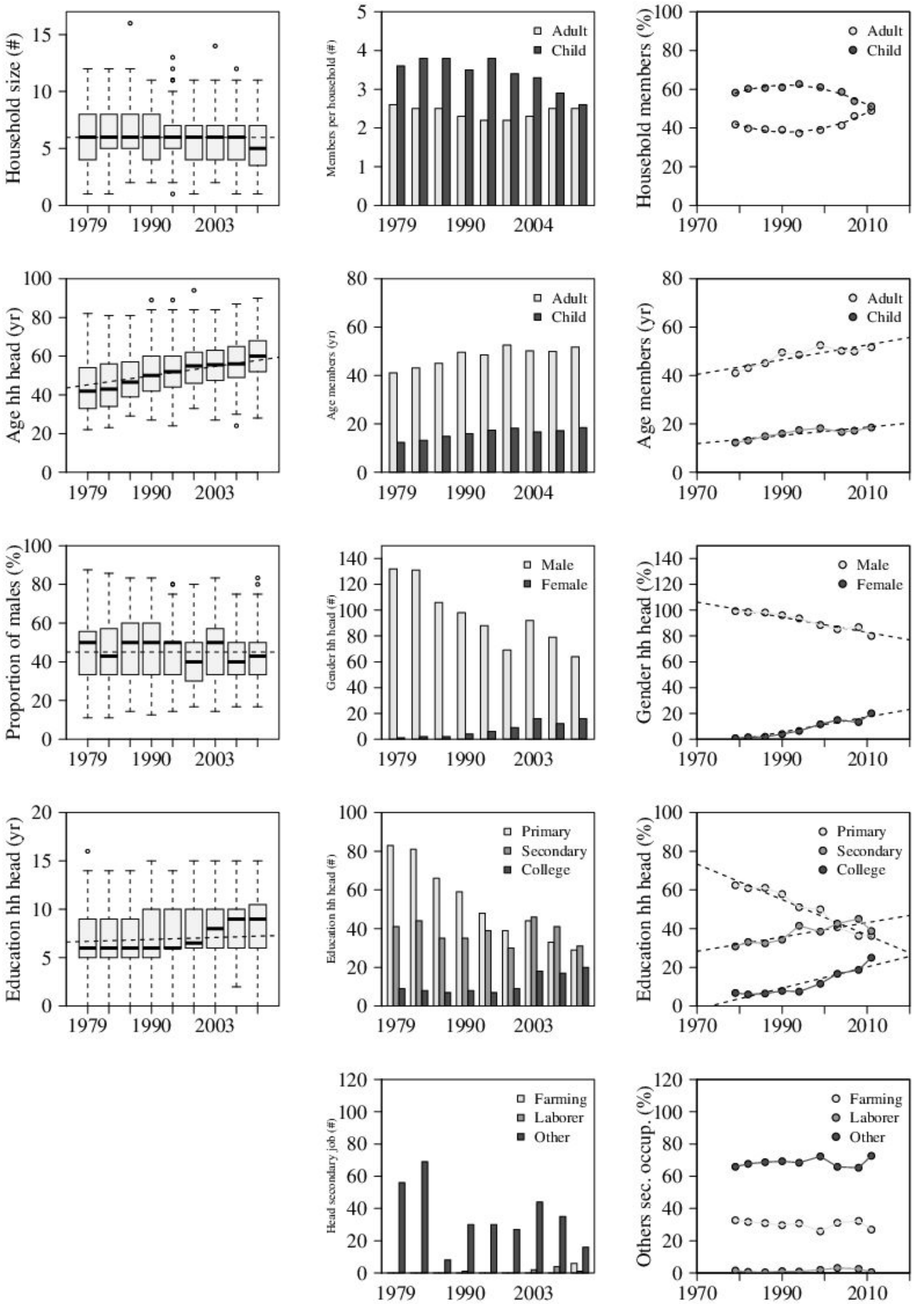

Figure B3: Descriptive statistics of household composition and characteristics over time. Dashed lines are fitted regressions with significant effects at 5% level. *Data source*: Central Luzon Loop Survey.

Table B3: Detailed characterization of the most used ($n > 15$) management packages in the WS (1966 - 2011). For each variable, the mean value across years is shown per management package and the standard deviation is provided in italics. *Data source*: Central Luzon Loop Survey.

	Irrigated Modern variety Direct-seeding Hand-weeding With fertiliser		Irrigated Modern variety Direct-seeding With herbicide With fertiliser		Irrigated Modern variety Transplanting Hand-weeding With fertiliser		Irrigated Modern variety Transplanting With herbicide With fertiliser		Irrigated Traditional variety Transplanting Hand-weeding With fertiliser		Rainfed Modern variety Transplanting Hand-weeding With fertiliser		Rainfed Modern variety Transplanting With herbicide With fertiliser		Rainfed Traditional variety Transplanting Hand-weeding With fertiliser	
Yield (ton/ha)	3.8	*1.2*	3.4	*1.1*	3.5	*1.3*	3.7	*1.2*	2.3	*0.9*	3.0	*1.3*	3.3	*1.2*	2.0	*0.7*
Seed use (kg/ha)	176.6	*50.9*	187.8	*69.0*	112.1	*55.9*	122.9	*53.3*	47.4	*22.8*	91.1	*45.5*	115.5	*53.1*	47.7	*8.7*
N use (kg/ha)	79.4	*27.5*	78.1	*34.7*	79.8	*43.0*	88.7	*45.4*	35.3	*27.6*	56.5	*38.0*	75.8	*56.0*	22.2	*13.5*
P use (kg/ha)	21.0	*10.0*	19.0	*13.9*	18.7	*17.4*	24.0	*19.5*	13.7	*11.8*	18.7	*19.5*	20.4	*19.4*	10.3	*11.9*
K use (kg/ha)	14.3	*10.7*	13.4	*13.8*	10.2	*13.5*	14.2	*16.9*	8.7	*10.9*	8.1	*12.2*	10.6	*13.4*	6.1	*8.7*
Irrigated (%)	100.0	*0.0*	99.3	*6.4*	98.5	*9.0*	99.2	*6.3*	100.0	*0.0*	3.1	*12.7*	2.8	*11.8*	0.0	*0.0*
Transplanted (%)	5.5	*14.4*	6.3	*17.7*	99.2	*7.4*	98.4	*9.8*	100.0	*0.0*	99.7	*3.1*	99.0	*6.6*	100.0	*0.0*
No. land preparation	3.7	*0.7*	4.0	*0.9*	3.4	*0.7*	3.8	*0.8*	3.0	*0.3*	3.3	*0.7*	3.5	*0.7*	3.0	*0.3*
No. crop establishment	1.8	*0.7*	2.1	*0.8*	3.9	*0.8*	4.0	*0.7*	3.3	*0.7*	3.8	*0.7*	3.9	*0.7*	3.4	*0.7*
Fertiliser applications	1.9	*0.5*	1.8	*0.5*	1.6	*0.6*	1.8	*0.7*	1.2	*0.4*	1.4	*0.6*	1.6	*0.7*	1.1	*0.3*
Farm size (ha)	2.0	*1.3*	2.1	*1.2*	1.9	*1.3*	1.9	*1.2*	2.5	*1.8*	1.7	*1.0*	1.7	*1.0*	2.2	*0.9*
Owned (%)	53.1	*50.0*	47.9	*47.9*	38.5	*47.8*	49.1	*48.2*	15.8	*37.0*	23.0	*42.0*	41.9	*48.6*	16.7	*38.1*
Rented (%)	46.9	*50.0*	47.1	*48.9*	48.7	*49.1*	42.0	*47.8*	35.5	*47.8*	59.5	*48.9*	46.1	*48.7*	25.0	*44.2*
Shared (%)	0.0	*0.0*	5.0	*19.6*	12.8	*33.1*	8.9	*27.5*	48.7	*50.0*	17.5	*37.9*	12.0	*32.1*	58.3	*50.4*
Labour use (ld/ha)	54.3	*22.6*	47.7	*17.0*	79.5	*26.8*	71.8	*26.3*	72.8	*29.7*	78.3	*28.1*	70.9	*24.7*	74.6	*23.4*
Family labour (ld/ha)	18.7	*19.1*	13.0	*12.3*	26.4	*21.4*	18.3	*18.8*	32.6	*16.2*	29.7	*21.4*	19.7	*16.0*	32.9	*19.8*
Hired labour (ld/ha)	35.5	*15.8*	34.7	*13.2*	53.1	*23.0*	53.6	*21.6*	40.2	*23.7*	48.6	*20.8*	51.3	*18.7*	41.7	*18.6*
Land preparation (ld/ha)	23.3	*17.2*	21.1	*13.5*	24.1	*22.2*	21.7	*18.9*	37.5	*22.5*	25.6	*22.8*	21.5	*17.1*	42.3	*27.3*
Crop establishment (ld/ha)	11.7	*14.4*	12.5	*15.6*	48.2	*37.1*	48.7	*35.7*	56.0	*42.5*	42.4	*30.6*	46.2	*32.9*	47.9	*29.3*
Fertiliser application (ld/ha)	3.5	*3.3*	2.7	*2.2*	2.5	*2.5*	2.6	*2.3*	3.1	*3.0*	1.9	*1.9*	2.4	*2.5*	1.9	*1.3*
Pesticide application (ld/ha)	14.5	*15.2*	6.8	*9.4*	17.7	*20.5*	8.5	*15.3*	24.6	*25.0*	16.0	*21.7*	7.5	*18.6*	22.1	*23.6*
Harvesting (ld/ha)	53.8	*36.2*	56.5	*30.8*	47.4	*38.5*	47.1	*36.8*	49.0	*40.0*	38.6	*26.8*	39.6	*28.5*	52.4	*48.5*
Revenues (P/ha)	47094	*47394*	44732	*33842*	43804	*30950*	45437	*32478*	38027	*21662*	36149	*20583*	39360	*28595*	31718	*18856*
Costs (P/ha)	40205	*35302*	37017	*24718*	33661	*22823*	35559	*19794*	29402	*18781*	28020	*17170*	30908	*18055*	30210	*16668*
Household size	5.6	*1.6*	6.4	*2.1*	5.4	*2.6*	5.8	*2.4*	2.2	*2.3*	5.1	*2.9*	5.3	*2.5*	2.0	*2.0*
Age HH head (yr)	51.0	*14.1*	52.3	*15.2*	51.1	*13.4*	53.4	*14.3*	47.6	*12.7*	46.7	*11.9*	50.9	*13.8*	43.9	*12.3*
Education HH head (yr)	8.2	*3.8*	7.3	*3.6*	6.7	*3.2*	7.1	*3.2*	4.6	*2.6*	6.5	*2.9*	7.3	*2.7*	5.3	*3.4*

Table B4: Detailed characterization of the most used ($n > 15$) management packages in the DS (1967 - 2012). For each variable, the mean value across years is shown per management package and the standard deviation is provided in italics. *Data source*: Central Luzon Loop Survey.

	Irrigated Modern variety Direct-seeding Hand-weeding With fertiliser		**Irrigated Modern variety Direct-seeding With herbicide With fertiliser**		**Irrigated Modern variety Transplanting Hand-weeding With fertiliser**		**Irrigated Modern variety Transplanting With herbicide With fertiliser**		**Irrigated Traditional variety Transplanting Hand-weeding With fertiliser**	
Yield (ton/ha)	4.3	*1.2*	4.2	*1.3*	3.9	*1.5*	4.4	*1.5*	2.0	*1.1*
Seed use (kg/ha)	180.4	*67.6*	178.5	*69.5*	124.9	*97.5*	120.8	*51.4*	59.4	*31.0*
N use (kg/ha)	113.3	*37.9*	104.4	*38.2*	99.1	*46.3*	105.5	*41.0*	34.3	*27.5*
P use (kg/ha)	20.0	*16.0*	26.7	*19.5*	21.8	*20.5*	25.4	*19.2*	16.4	*16.0*
K use (kg/ha)	13.0	*13.7*	18.1	*18.0*	13.0	*16.6*	15.0	*16.7*	9.8	*9.3*
Irrigated (%)	100.0	*0.0*	100.0	*0.0*	100.0	*0.0*	100.0	*0.0*	100.0	*0.0*
Transplanted (%)	3.2	*13.9*	1.1	*7.9*	99.2	*6.8*	98.7	*8.1*	100.0	*0.0*
No. land preparation	4.1	*0.9*	4.3	*0.8*	3.5	*0.8*	3.8	*0.9*	3.1	*0.2*
No. crop establishment	2.2	*1.1*	2.0	*0.8*	3.8	*0.8*	4.0	*0.8*	3.3	*0.6*
Fertiliser applications	2.1	*0.7*	2.2	*0.8*	1.8	*0.8*	1.9	*0.7*	1.1	*0.3*
Farm size (ha)	1.8	*1.1*	1.9	*1.1*	1.7	*1.3*	1.8	*1.6*	1.7	*1.0*
Owned (%)	53.8	*50.0*	51.0	*49.1*	32.2	*46.7*	37.1	*46.3*	5.0	*22.4*
Rented (%)	42.6	*49.4*	44.1	*49.1*	55.2	*49.4*	53.8	*48.0*	35.0	*48.9*
Shared (%)	3.6	*17.5*	4.9	*20.1*	12.7	*32.9*	9.1	*28.4*	60.0	*50.3*
Labour use (ld/ha)	54.2	*22.9*	46.2	*18.1*	86.1	*30.7*	72.4	*24.9*	75.1	*26.1*
Family labour (ld/ha)	19.9	*17.0*	12.8	*11.2*	28.6	*26.6*	17.3	*18.1*	31.9	*26.7*
Hired labour (ld/ha)	34.3	*20.7*	33.4	*15.3*	57.5	*19.6*	55.0	*19.0*	43.2	*14.5*
Land preparation (ld/ha)	20.0	*14.3*	20.4	*14.0*	20.1	*14.6*	19.6	*19.8*	21.0	*11.3*
Crop establishment (ld/ha)	7.5	*9.8*	7.3	*11.6*	46.8	*36.2*	45.9	*36.3*	40.1	*21.8*
Fertiliser application (ld/ha)	3.3	*2.8*	3.3	*3.3*	2.6	*2.3*	2.8	*3.2*	2.1	*2.7*
Pesticide application (ld/ha)	15.9	*19.3*	4.9	*9.7*	17.3	*21.4*	9.5	*18.1*	12.6	*8.6*
Harvesting (ld/ha)	46.9	*32.3*	50.0	*32.1*	46.6	*35.8*	49.0	*42.7*	33.1	*21.4*
Revenues (P/ha)	46168	*27279*	49159	*29297*	49178	*30886*	53519	*34604*	30989	*24850*
Costs (P/ha)	34371	*13637*	36426	*19608*	35866	*15901*	38537	*19725*	22548	*14652*
Household size	5.6	*1.9*	6.1	*2.2*	5.6	*2.6*	5.9	*2.3*	3.1	*3.3*
Age HH head (yr)	55.9	*12.6*	54.9	*13.8*	52.5	*13.6*	50.4	*14.6*	50.2	*14.4*
Education HH head (yr)	6.7	*3.5*	6.8	*3.4*	6.5	*3.4*	7.3	*3.3*	5.1	*2.0*

Table B5: Farm level indicators, management practices, farm resources (land and labour) and household characteristics across different management practices during the WS in Central Luzon, Philippines. The columns 'Sign.' report the number of years for which the difference between two contrasting management practices is significant at 5% (the maximum number of years is presented in after the '/'). Standard deviations (SD) are presented in italics. *Data source*: Central Luzon Loop Survey.

	Direct-seeding		Transplanting			Herbicide		Hand-weeding			NxPxKx		Other NPK		
	Mean	SD	Mean	SD	Signif.	Mean	SD	Mean	SD	Signif.	Mean	SD	Mean	SD	Signif.
Indicators															
Yield (ton/ha)	3.5	*1.1*	3.4	*1.3*	2/9	3.2	*1.4*	3.6	*1.3*	1/11	3.5	*1.3*	3.0	*1.4*	1/12
Labour prod. (kg/ld)	79.3	*35.3*	56.2	*53.7*	8/9	49.0	*34.6*	64.3	*63.0*	3/11	59.6	*42.6*	53.1	*77.3*	3/12
Gross margin (kPhP/ha)	7.5	*17.3*	10.7	*23.7*	1/9	9.6	*22.3*	9.9	*23.4*	0/11	10.3	*22.5*	10.9	*25.5*	1/12
NUE (kg N/kg N)	0.3	*0.1*	0.3	*0.2*	2/9	0.3	*0.2*	0.3	*0.2*	0/11	0.3	*0.1*	0.3	*0.2*	4/12
Management															
Seed (kg/ha)	186.8	*65.3*	110.0	*56.4*	7/9	100.1	*57.8*	128.4	*62.8*	1/11	119.4	*60.3*	105.6	*61.6*	1/12
N (kg/ha)	78.6	*33.2*	73.5	*47.4*	3/9	63.2	*43.4*	81.3	*47.7*	1/11	78.0	*46.4*	60.7	*43.7*	3/12
P (kg/ha)	19.9	*13.0*	19.6	*18.8*	0/9	17.0	*17.1*	21.6	*18.8*	0/11	25.6	*16.9*	0.0	*0.0*	12/12
K (kg/ha)	14.1	*13.3*	11.1	*14.5*	3/9	9.0	*12.5*	12.8	*15.6*	0/11	14.7	*14.8*	0.0	*0.0*	12/12
Irrigated area (%)	99.4	*5.7*	71.4	*44.3*	3/9	68.0	*45.9*	80.4	*39.0*	1/11	75.3	*42.3*	68.7	*45.8*	2/12
Transplanted area (%)	6.2	*17.1*	98.7	*8.7*	3/9	95.2	*20.5*	87.2	*32.2*	0/11	89.9	*28.7*	95.3	*20.7*	3/12
Land															
Farm size (ha)	2.1	*1.2*	1.9	*1.5*	6/9	1.9	*1.5*	2.0	*1.4*	1/11	2.0	*1.4*	1.9	*1.6*	4/12
Own area (%)	49.7	*47.9*	39.8	*47.5*	1/9	31.1	*45.6*	46.7	*48.0*	2/11	44.7	*48.3*	26.8	*42.3*	2/12
Rented area (%)	45.5	*48.7*	45.9	*48.4*	3/9	49.6	*49.2*	43.1	*47.9*	1/11	44.1	*48.3*	51.9	*48.3*	1/12
Shared area (%)	4.8	*18.9*	14.3	*34.3*	0/9	19.3	*39.2*	10.2	*29.1*	1/11	11.2	*30.5*	21.3	*40.6*	1/12
Labour															
Labour use (ld/ha)	48.8	*18.7*	72.2	*27.4*	7/9	75.5	*28.7*	67.7	*27.5*	2/11	70.0	*27.1*	71.2	*29.1*	0/12
Family labour (ld/ha)	13.9	*14.2*	21.9	*19.4*	5/9	27.1	*20.7*	17.8	*17.7*	3/11	19.4	*18.0*	27.3	*21.6*	5/12
Hired labour (ld/ha)	34.9	*13.8*	50.3	*22.0*	7/9	48.4	*23.0*	49.9	*22.0*	1/11	50.6	*21.9*	43.9	*20.6*	2/12
Household															
Household size (#)	6.2	*2.1*	5.3	*2.7*	2/9	4.9	*2.8*	5.7	*2.4*	0/11	5.4	*2.6*	5.5	*2.7*	3/12
Male members (%)	44.1	*15.5*	45.5	*16.3*	2/9	46.9	*15.9*	44.5	*16.2*	0/11	44.8	*16.2*	47.2	*16.4*	0/12
Age hh head (yr)	51.7	*14.8*	51.4	*13.7*	2/9	49.4	*13.1*	52.6	*14.2*	0/11	52.0	*13.8*	49.8	*13.6*	0/12
Education hh head (yr)	7.7	*3.5*	6.9	*3.2*	2/9	6.5	*3.2*	7.2	*3.2*	0/11	7.2	*3.2*	6.1	*3.2*	0/12

Table B6: Farm level indicators, management practices, farm resources (land and labour) and household characteristics across different management practices during the DS in Central Luzon, Philippines. The columns 'Sign.' report the number of years for which the difference between two contrasting management practices is significant at 5% (the maximum number of years is presented in after the '/'). Standard deviations (SD) are presented in italics. *Data source*: Central Luzon Loop Survey.

	Direct-seeding		**Transplanting**			**Herbicide**		**Hand-weeding**			**NxPxKx**		**Other NPK**		
	Mean	**SD**	**Mean**	**SD**	**Signif.**	**Mean**	**SD**	**Mean**	**SD**	**Signif.**	**Mean**	**SD**	**Mean**	**SD**	**Signif.**
Indicators															
Yield (ton/ha)	4.3	*1.3*	4.1	*1.7*	2/8	3.9	*1.7*	4.3	*1.5*	1/7	4.4	*1.5*	3.5	*1.7*	4/10
Labour prod. (kg/ld)	110.0	*69.8*	69.0	*69.2*	4/8	63.7	*63.3*	94.3	*70.4*	2/7	91.4	*67.0*	65.8	*83.9*	4/10
Gross margin (kPhP/ha)	14.9	*35.3*	19.2	*76.4*	1/8	12.7	*22.9*	19.4	*78.0*	0/7	20.7	*71.6*	7.5	*23.8*	1/10
NUE (kg N/kg N)	0.3	*0.1*	0.3	*0.2*	1/8	0.3	*0.2*	0.3	*0.1*	0/7	0.3	*0.1*	0.3	*0.2*	3/10
Management															
Seed (kg/ha)	183.2	*74.2*	116.1	*71.3*	7/8	126.0	*93.4*	151.6	*72.3*	1/7	148.4	*69.2*	124.2	*103.2*	2/10
N (kg/ha)	108.1	*37.9*	94.8	*56.0*	1/8	87.7	*54.1*	103.1	*44.3*	1/7	105.0	*43.7*	84.5	*63.9*	1/10
P (kg/ha)	27.8	*23.6*	22.0	*20.1*	0/8	19.2	*19.3*	26.4	*22.6*	0/7	32.2	*19.2*	0.0	*0.0*	10/10
K (kg/ha)	18.8	*19.2*	13.3	*16.7*	0/8	11.9	*15.4*	17.2	*18.8*	0/7	20.5	*18.0*	0.0	*0.0*	10/10
Irrigated area (%)	100.0	*0.0*	98.6	*11.8*	0/8	98.3	*12.8*	99.7	*5.2*	0/7	99.5	*6.7*	97.9	*14.4*	0/10
Transplanted area (%)	1.4	*9.0*	98.5	*8.5*	8/8	80.8	*39.0*	47.0	*49.5*	0/7	55.3	*48.9*	76.2	*42.7*	1/10
Land															
Farm size (ha)	2.0	*1.3*	1.8	*1.7*	1/8	1.8	*1.5*	1.9	*1.5*	0/7	2.0	*1.5*	1.7	*1.7*	3/10
Own area (%)	53.3	*49.0*	32.8	*45.5*	1/8	33.2	*47.0*	45.5	*48.3*	1/7	47.3	*48.4*	21.0	*40.5*	3/10
Rented area (%)	42.1	*48.6*	54.6	*48.4*	0/8	52.0	*49.6*	47.7	*48.6*	1/7	45.9	*48.5*	61.2	*48.3*	2/10
Shared area (%)	4.6	*19.5*	12.6	*32.8*	0/8	14.9	*35.2*	6.8	*24.3*	0/7	6.8	*24.2*	17.7	*38.1*	0/10
Labour															
Labour use (ld/ha)	58.9	*192.4*	78.6	*57.6*	8/8	78.4	*36.1*	67.9	*160.3*	4/7	70.0	*146.9*	73.5	*34.0*	2/10
Family labour (ld/ha)	14.2	*19.5*	23.0	*26.9*	0/8	26.7	*25.5*	16.3	*24.1*	0/7	17.0	*25.3*	27.3	*20.6*	6/10
Hired labour (ld/ha)	44.8	*177.1*	55.6	*40.4*	8/8	51.6	*25.4*	51.6	*144.4*	0/7	53.0	*131.9*	46.3	*24.9*	0/10
Household															
Household size (#)	6.0	*2.2*	5.4	*2.6*	0/8	5.1	*2.7*	6.0	*2.3*	0/7	5.8	*2.4*	5.4	*2.5*	1/10
Male members (%)	42.9	*15.2*	46.9	*16.5*	3/8	46.9	*16.6*	44.0	*15.8*	0/7	44.9	*15.8*	46.4	*16.9*	1/10
Age hh head (yr)	54.8	*13.6*	51.5	*14.0*	0/8	53.3	*13.5*	52.6	*14.0*	2/7	53.4	*14.2*	51.0	*13.0*	2/10
Education hh head (yr)	7.1	*3.4*	7.0	*3.3*	1/8	6.5	*3.3*	7.2	*3.3*	0/7	7.1	*3.3*	6.7	*3.3*	5/10

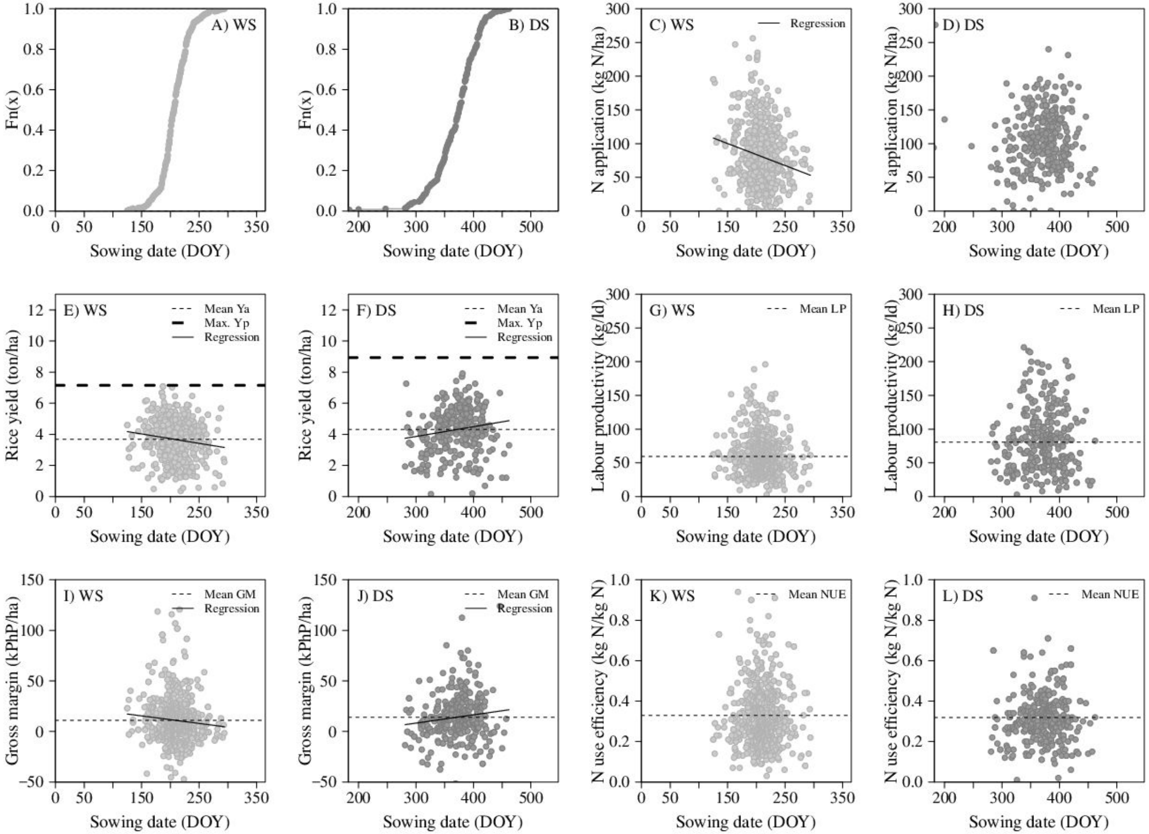

Figure B4: Cumulative probability of sowing date (DOY, day of the year; A - B) and its relationship with N applied (C - D), rice yield (E - F), labour productivity (G - H), gross margin (I - J) and N use efficiency (K - L) in rice-based farming systems in Central Luzon, Philippines (1966 - 2012). Vertical dashed lines in A) - D) reflect the average time of active tillering (T) and panicle initiation (PI). Statistically significant fitted regressions are shown as solid lines. *Data source*: Central Luzon Loop Survey.

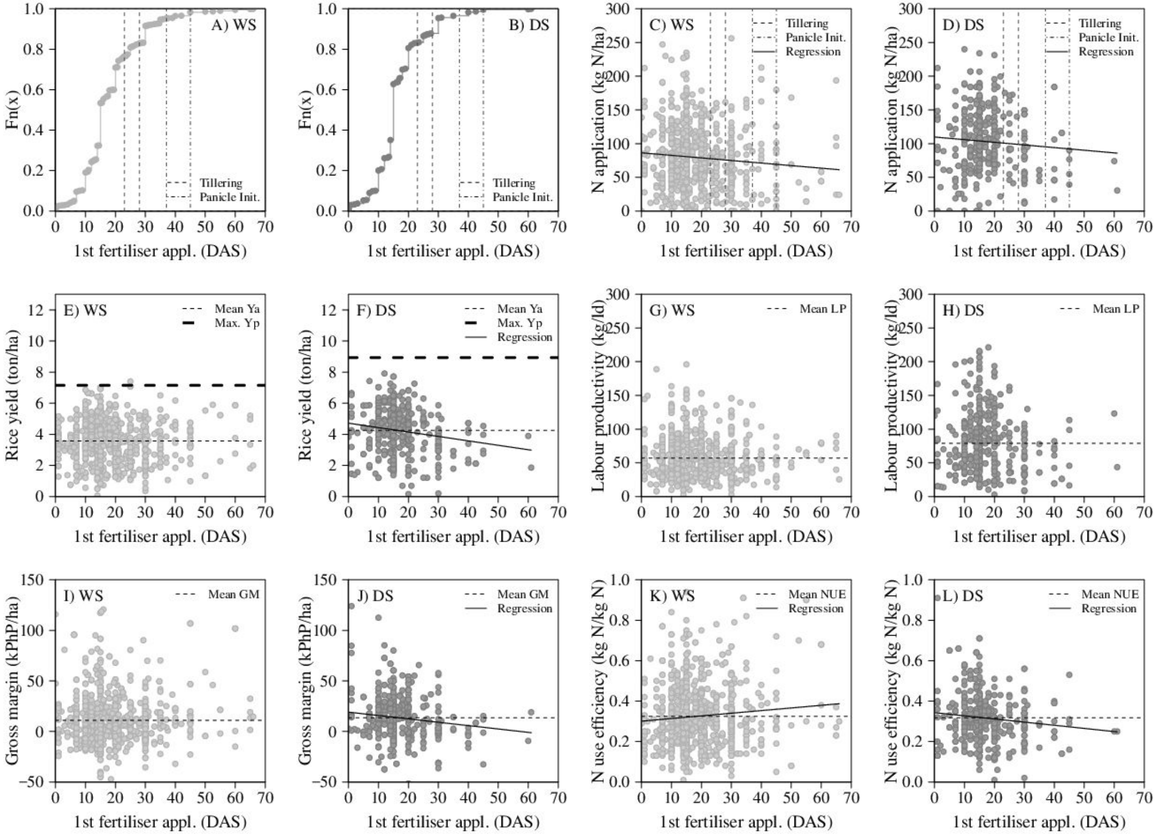

Figure B5: Cumulative probability of timing of the 1st fertiliser application (DAS, days after sowing; A - B) and its relationship with N applied (C - D), rice yield (E - F), labour productivity (G - H), gross margin (I - J) and N use efficiency (K - L) in rice-based farming systems in Central Luzon, Philippines (1966 - 2012). Vertical dashed lines in A) - D) reflect the average time of active tillering (T) and panicle initiation (PI). Statistically significant fitted regressions are shown as solid lines. *Data source*: Central Luzon Loop Survey.

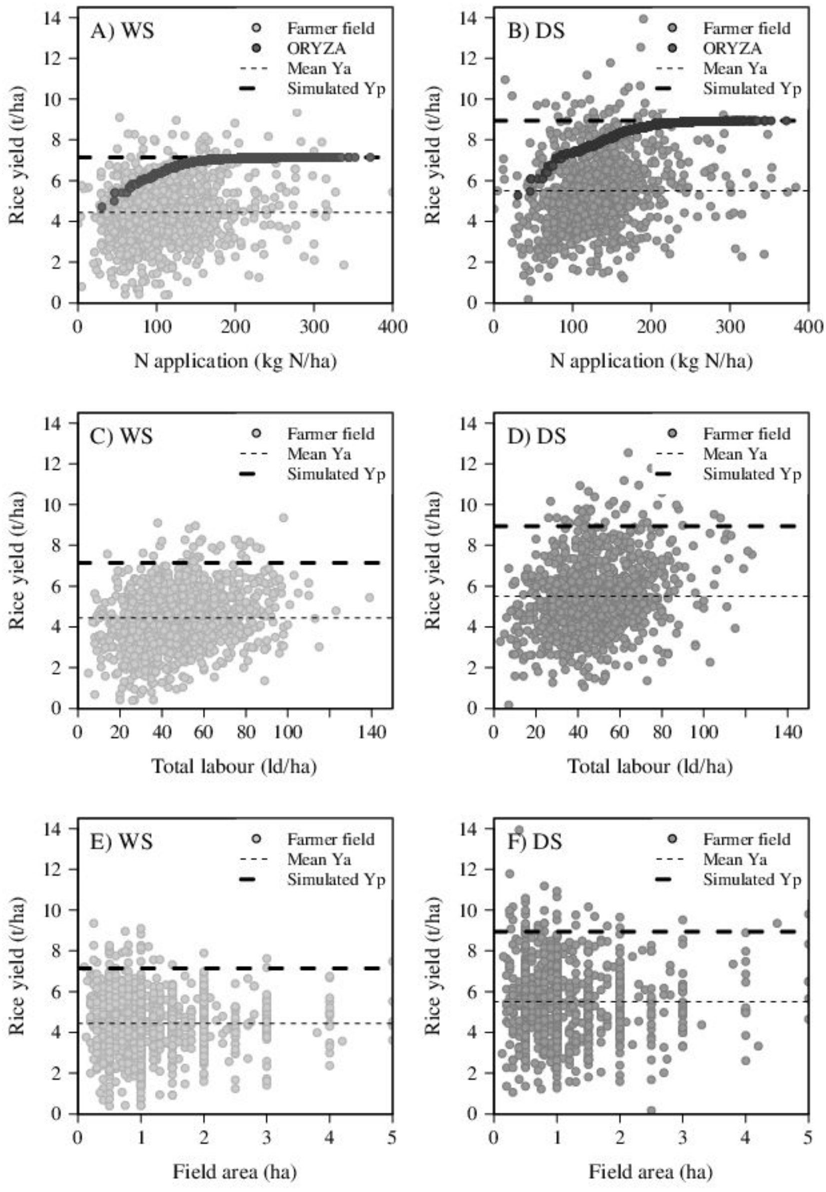

Figure B6: Relationship between rice yields and A - B) N applied, C - D) total labour; E - F) field area for rice-based farming systems in Central Luzon, Philippines (2013 DS - 2014 WS). WS data includes both irrigated and rainfed observations while DS data include irrigated observations only. Each dot represents the largest parcel cultivated with rice within a single farm. *Data source*: MISTIG Survey.

Annex C. Arable farming systems in the Netherlands (Chapter 5)

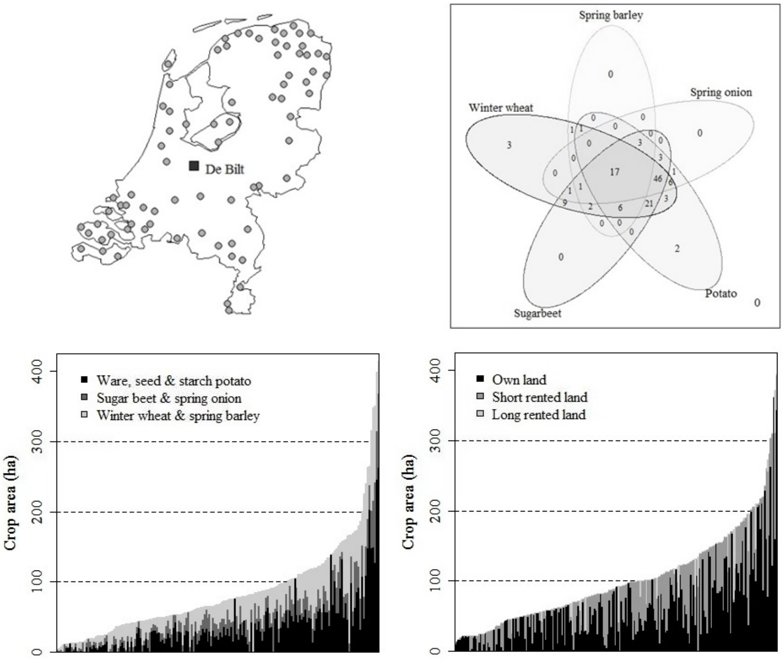

Figure C1: Top left: Location of villages. **Top right**: Crops grown per farm. **Bottom left**: Land cultivated with different crops by each individual farm. **Bottom right**: Owned and rented land for each individual farm.

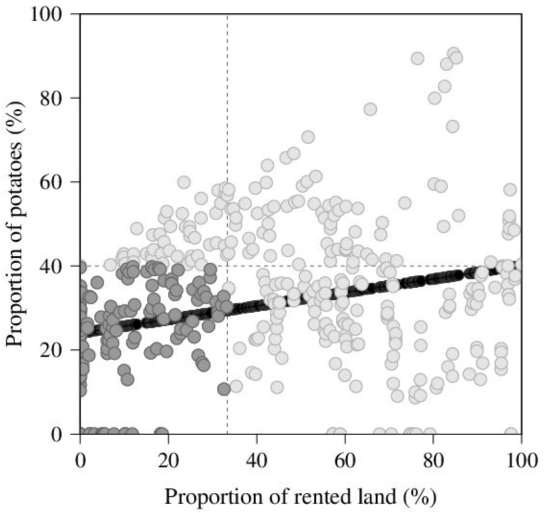

Figure C2: Relationship between proportion of rented land and potato area share. Vertical dashed line show the 33.3% of rented land and horizontal dashed line the 40% potato area share. The black line on the background shows the fitted linear regression between both variables ($y = 0.16x$+24.1, R^2 = 8%). Dark grey points were the observations used in the analysis (n = 189) and light grey dots were the observations excluded (n = 285) due to confounding effects of rented land and potato area share.

Table C1: Crop-specific descriptive statistics of inputs and outputs for arable farming systems in The Netherlands during the period 2008 - 2012. Standard deviations are presented between brackets. 'Nitrogen' refers to plant available N from mineral fertilisers and organic amendments.

	2008		**2009**		**2010**		**2011**		**2012**	
Ware potato										
Sample size (n)	81		80		83		83		86	
Clay soil (n)	79		77		73		77		76	
FM yield (ton ha^{-1})	53.5	(10.0)	52.0	(10.2)	46.3	(11.6)	52.9	(12.0)	50.5	(11.3)
Nitrogen (kg N ha^{-1})	221.9	(126.1)	226.0	(145.6)	281.4	(283.7)	226.5	(94.1)	225.4	(93.6)
Phosphorus (kg P ha^{-1})	35.8	(50.7)	29.4	(28.3)	43.7	(86.5)	33.9	(42.6)	29.7	(33.1)
Seed potato										
Sample size (n)	91		94		97		91		88	
Clay soil (n)	71		71		66		65		62	
FM yield (ton ha^{-1})	33.8	(7.2)	35.8	(8.6)	34.9	(8.4)	32.8	(8.2)	35.2	(8.1)
Nitrogen (kg N ha^{-1})	119.5	(73.8)	129.3	(64.9)	125.3	(71.9)	130.5	(77.0)	123.3	(88.2)
Phosphorus (kg P ha^{-1})	42.0	(53.3)	37.9	(28.8)	36.3	(31.1)	40.2	(44.3)	31.7	(18.7)
Starch potato										
Sample size (n)	23		24		32		29		29	
FM yield (ton ha^{-1})	44.6	(4.9)	43.4	(4.3)	41.4	(4.2)	43.4	(4.3)	46.6	(3.9)
Nitrogen (kg N ha^{-1})	204.4	(50.2)	186.4	(36.7)	206.0	(39.2)	188.8	(39.1)	185.5	(38.3)
Phosphorus (kg P ha^{-1})	42.6	(15.4)	40.6	(15.0)	40.6	(13.1)	39.8	(11.7)	34.3	(12.0)

Table C1: Crop-specific descriptive statistics of inputs and outputs for arable farming systems in The Netherlands during the period 2008 - 2012. Standard deviations are presented between brackets. 'Nitrogen' refers to plant available N from mineral fertilisers and organic amendments. *(continued)*

	2008		**2009**		**2010**		**2011**		**2012**	
Sugar beet										
Sample size (n)	142		149		160		152		153	
Clay soil (n)	106		109		110		108		105	
FM yield (ton ha^{-1})	74.2	(11.8)	81.6	(11.7)	77.4	(9.8)	80.8	(10.0)	80.1	(9.4)
Sugar content (%)	17.2	(0.6)	17.7	(0.7)	17.0	(0.6)	17.1	(0.6)	17.2	(0.6)
Nitrogen (kg N ha^{-1})	152.9	(168.1)	134.9	(52.6)	149.6	(171.1)	145.7	(77.2)	139.5	(61.2)
Phosphorus (kg P ha^{-1})	28.3	(46.9)	22.5	(27.7)	18.4	(22.1)	20.8	(28.9)	17.6	(21.4)
Spring onion										
Sample size (n)	58		58		62		61		62	
Clay soil (n)	58		58		59		58		59	
FM yield (ton ha^{-1})	56.8	(13.8)	61.4	(10.3)	52.3	(15.4)	56.7	(16.0)	57.3	(14.3)
Nitrogen (kg N ha^{-1})	135.6	(90.9)	166.5	(78.3)	223.3	(274.3)	149.7	(66.3)	136.3	(50.2)
Phosphorus (kg P ha^{-1})	19.6	(20.7)	25.1	(30.0)	27.2	(49.3)	19.0	(19.4)	15.8	(16.8)
Winter wheat										
Sample size (n)	104		98		107		99		103	
Clay soil (n)	99		95		102		95		98	
FM yield (ton ha^{-1})	9.2	(1.3)	9.6	(1.2)	9.2	(1.4)	8.1	(1.2)	8.7	(1.0)
Nitrogen (kg N ha^{-1})	234.1	(177.8)	222.9	(211.2)	237.0	(224.6)	211.6	(71.2)	220.2	(154.2)
Phosphorus (kg P ha^{-1})	19.1	(33.0)	23.7	(33.8)	25.5	(48.3)	22.6	(22.5)	20.4	(23.3)
Spring barley										
Sample size (n)	82		69		66		72		51	
Clay soil (n)	51		42		32		45		24	
FM yield (ton ha^{-1})	6.4	(1.4)	7.2	(1.1)	6.3	(1.4)	6.2	(1.3)	6.8	(0.9)
Nitrogen (kg N ha^{-1})	95.3	(58.8)	89.9	(36.4)	122.2	(161.7)	93.2	(34.9)	107.1	(53.6)
Phosphorus (kg P ha^{-1})	15.5	(35.9)	11.0	(13.9)	16.9	(27.8)	10.3	(15.7)	12.0	(13.4)

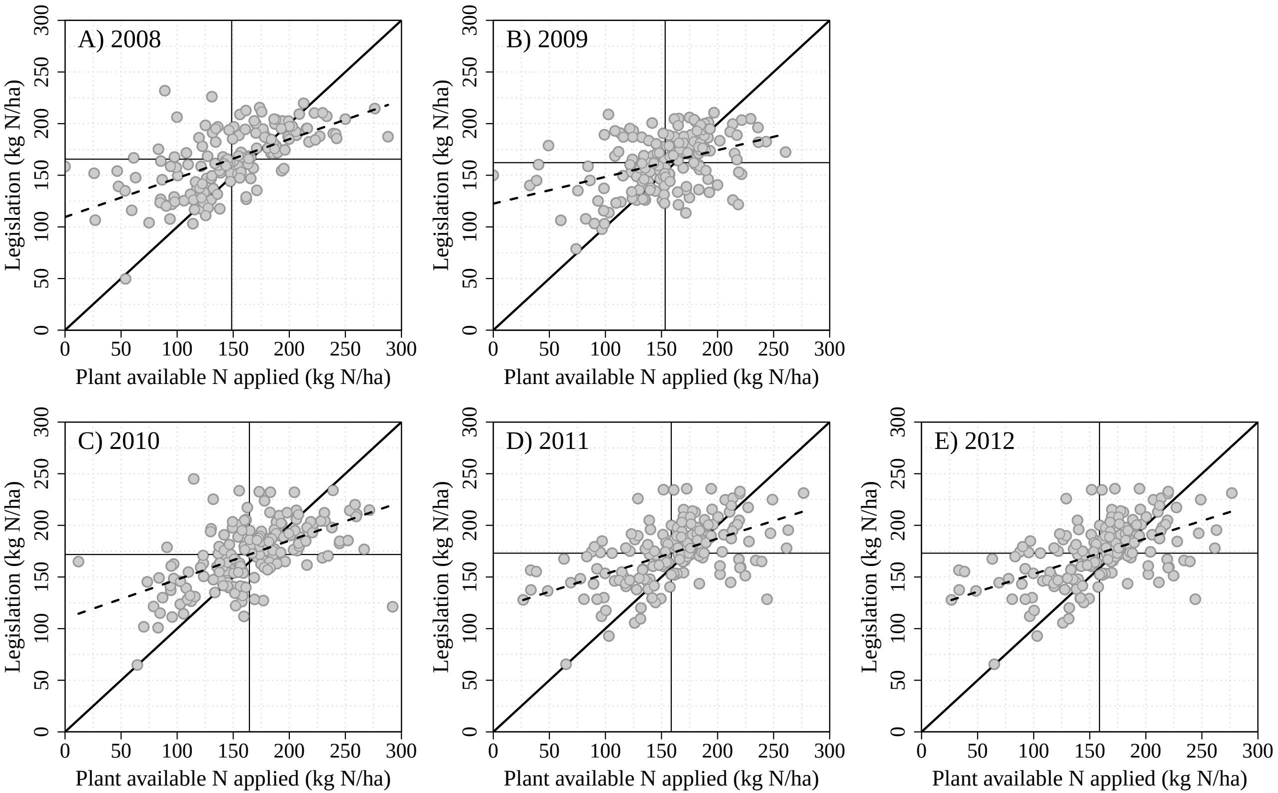

Figure C3: Farmer compliance at the farm level with N legislation enforced in the Netherlands for the period 2008 - 2012. Plant available N applied (x-axis) refers to the sum of N applied with mineral fertilisers together with the N fertiliser replacement value of organic amendments. Legislation (y-axis) corresponds to the maximum amount of plant available N which can be applied at the farm level by each single farm. The 1:1 line is presented as a solid line and vertical and horizontal solid lines show the average plant available N applied and the average amount of N which can be applied according to the legislation, respectively. Dashed lines are fitted (and statistically significant) linear regressions. As the N legislation is controlled at farm level, N levels for individual crops can be higher than the legal level per crop. A high N level for one crop can be compensated by a lower level for another crop.

Table C2: Summary results of regression analysis.

	Estimate	Std. Error	t value	Pr(>|t|)
y = *ware potato yield*				
Intercept	15.0752	10.7838	1.40	0.1652
Area share of potato	2.4705	0.8718	2.83	0.0055
Area share of potato^2	-0.0375	0.0166	-2.26	0.0262
Intercept	50.4967	3.8927	12.97	0.0000
Area share of sugar beet	0.3983	0.6017	0.66	0.5086
Area share of sugar beet^2	-0.0212	0.0220	-0.97	0.3348
Intercept	46.9979	1.0622	44.25	0.0000
Area share of spring onion	0.4699	0.1000	4.70	0.0000
Intercept	49.1455	2.7033	18.18	0.0000
Area share of cereals	0.1588	0.1914	0.83	0.4075
Area share of cereals^2	-0.0028	0.0032	-0.88	0.3774
y = *sugar beet yield*				
Intercept	52.9932	11.4055	4.65	0.0000
Area share of potato	2.4122	0.8868	2.72	0.0079
Area share of potato^2	-0.0429	0.0165	-2.60	0.0109
Intercept	82.2294	2.7209	30.22	0.0000
Area share of sugar beet	0.0973	0.4029	0.24	0.8090
Area share of sugar beet^2	-0.0071	0.0141	-0.50	0.6160
Intercept	78.6689	0.6387	123.16	0.0000
Area share of spring onion	0.5147	0.0675	7.62	0.0000
Intercept	85.1259	1.0044	84.75	0.0000
Area share of cereals	-0.0905	0.0264	-3.43	0.0007
y = *winter wheat yield*				
Intercept	5.3765	1.4327	3.75	0.0003
Area share of potato	0.2720	0.1144	2.38	0.0194
Area share of potato^2	-0.0046	0.0022	-2.12	0.0368
Intercept	8.4546	0.1726	48.97	0.0000
Area share of sugar beet	0.0858	0.0313	2.74	0.0063
Area share of sugar beet^2	-0.0031	0.0014	-2.24	0.0253
Intercept	8.6975	0.0925	94.02	0.0000
Area share of spring onion	0.0338	0.0102	3.32	0.0010
Intercept	9.2188	0.1519	60.69	0.0000
Area share of cereals	-0.0084	0.0038	-2.23	0.0262

Table C3: Performance of cropping systems in terms of N and energy production, profitability, labour productivity and N use efficiency at farm level. All data were compiled for the period 2008 - 2012 and clay soils. Production under 'Actual' and 'Highest' was estimated based on observed cropping sequences and crop-specific Y_A and Y_{HF}, respectively. Crop yields are reported in ton FM ha^{-1} and N and P applied are expressed in kg ha^{-1}. N refers to total N applied in mineral and organic forms. n.a. = not applicable

Year / Soil 2008-12/Clay	Actual N prod. (kg N ha^{-1})	Actual MJ prod. (10^3 MJ ha^{-1})	Actual Gr. margin (€ ha^{-1})	Actual Labour (hr ha^{-1})	Actual NUE (kg kg^{-1})	Highest N prod. (kg N ha^{-1})	Highest Gr. margin (€ ha^{-1})
Ware potato							
Obj. yield	**200.2**	**199.9**	**4551.4**	**n.a.**	**n.a.**	**212.0**	**3861.2**
FM yield	60.7	58.8	54.4	49.8	44.0	64.3	64.3
N applied	346.4	310.1	252.5	298.1	244.3	248.2	248.2
P applied	39.7	29.8	17.6	25.8	28.3	33.8	33.8
Seed potato							
Obj. yield	**106.5**	**135.2**	**7296.7**	**n.a.**	**n.a.**	**138.0**	**8452.9**
FM yield	35.5	39.8	42.2	35.9	38.9	46.0	46.0
N applied	196.6	228.4	225.1	942.2	162.6	197.5	197.5
P applied	27.3	41.6	36.7	202.4	35.3	31.6	31.6
Sugar beet							
Obj. yield	**157.9**	**468.8**	**3120.4**	**n.a.**	**n.a.**	**170.6**	**3614.4**
FM yield	87.7	88.4	85.3	82.7	80.0	94.8	94.8
N applied	169.1	170.6	194.6	145.6	154.8	170.8	170.8
P applied	16.5	19.2	14.1	4.7	16.7	14.5	14.5
Spring onion							
Obj. yield	**141.2**	**84.9**	**4465.7**	**n.a.**	**n.a.**	**164.5**	**3822.7**
FM yield	64.2	65.3	68.3	57.0	61.3	74.8	74.8
N applied	169.5	157.2	198.3	173.0	150.5	188.8	188.8
P applied	16.8	15.3	21.9	10.9	20.8	23.8	23.8

Table C3: Performance of cropping systems in terms of N and energy production, profitability, labour productivity and N use efficiency at farm level. All data were compiled for the period 2008 - 2012 and clay soils. Production under 'Actual' and 'Highest' was estimated based on observed cropping sequences and crop-specific Y_A and Y_{HF}, respectively. Crop yields are reported in ton FM ha^{-1} and N and P applied are expressed in kg ha^{-1}. N refers to total N applied in mineral and organic forms. n.a. = not applicable *(continued)*

Year / Soil 2008-12/Clay	**Actual N prod. (kg N ha^{-1})**	**Actual MJ prod. (10^3 MJ ha^{-1})**	**Actual Gr. margin (€ ha^{-1})**	**Actual Labour (hr ha^{-1})**	**Actual NUE (kg kg^{-1})**	**Highest N prod. (kg N ha^{-1})**	**Highest Gr. margin (€ ha^{-1})**
Winter wheat							
Obj. yield	**162.6**	**150.9**	**1080.4**	**n.a.**	**n.a.**	**173.5**	**1215.6**
FM yield	9.7	9.7	9.6	9.0	9.0	10.3	10.3
N applied	299.3	295.9	293.3	324.1	257.2	300.6	300.6
P applied	23.0	22.8	29.3	34.5	13.5	22.9	22.9
Spring barley							
Obj. yield	**77.7**	**108.1**	**772.7**	**n.a.**	**n.a.**	**107.6**	**851.4**
FM yield	6.2	6.8	6.5	7.1	6.9	8.5	8.5
N applied	118.6	126.3	127.1	131.4	105.4	109.9	109.9
P applied	9.5	7.5	2.7	8.4	3.4	0.5	0.5
Farm total							
N prod.	**166.0**	163.7	137.3	110.4	130.6		
Energy prod.	205.0	**209.6**	172.6	140.4	166.7		
Gross margin	3170.3	3564.7	**5056.2**	1733.7	3900.3		
Labour use	44.8	51.6	68.8	**21.5**	66.2		
NUE	0.7	0.7	0.6	0.5	**0.9**		

Table C4: Summary results of regression analysis.

	Estimate	Std. Error	t value	Pr(>|t|)
y = *nitrogen production*				
Intercept	9.7469	0.5873	16.60	0.0000
Gross margin	0.9935	0.1866	5.33	0.0000
Intercept	10.4660	1.1714	8.93	0.0000
NUE	3.5490	1.8666	1.90	0.0603
Intercept	12.2470	0.6038	20.28	0.0000
Labour use	0.0100	0.0111	0.90	0.3721
y = *ware potato yield*				
Intercept	41.3800	3.4279	12.07	0.0000
Gross margin	3.0025	1.0321	2.91	0.0048
Intercept	48.6134	4.6513	10.45	0.0000
NUE	3.4403	7.0886	0.49	0.6289
Intercept	55.4989	2.7643	20.08	0.0000
Labour use	-0.0777	0.0505	-1.54	0.1279
y = *sugar beet yield*				
Intercept	73.5089	2.2904	32.09	0.0000
Gross margin	2.2842	0.7243	3.15	0.0022
Intercept	83.9254	3.4976	24.00	0.0000
NUE	-5.8499	5.3425	-1.09	0.2766
Intercept	82.0288	2.1837	37.56	0.0000
Labour use	-0.0357	0.0395	-0.90	0.3682
y = *winter wheat yield*				
Intercept	8.3114	0.2363	35.17	0.0000
Gross margin	0.2591	0.0759	3.42	0.0009
Intercept	9.2768	0.3730	24.87	0.0000
NUE	-0.3562	0.5820	-0.61	0.5420
Intercept	9.1553	0.1952	46.90	0.0000
Labour use	-0.0005	0.0036	-0.13	0.8948

Annex D. Mixed crop-livestock systems in Southern Ethiopia (Chapter 6)

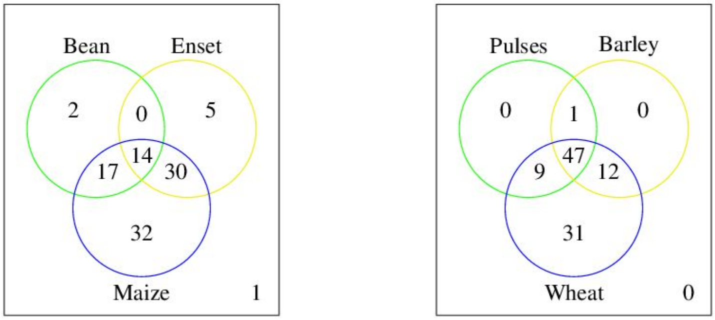

Figure D1: Number of farms cultivating maize, bean and enset in Hawassa (left) and wheat, barley (or other small grains) and pulses in Asella (right).

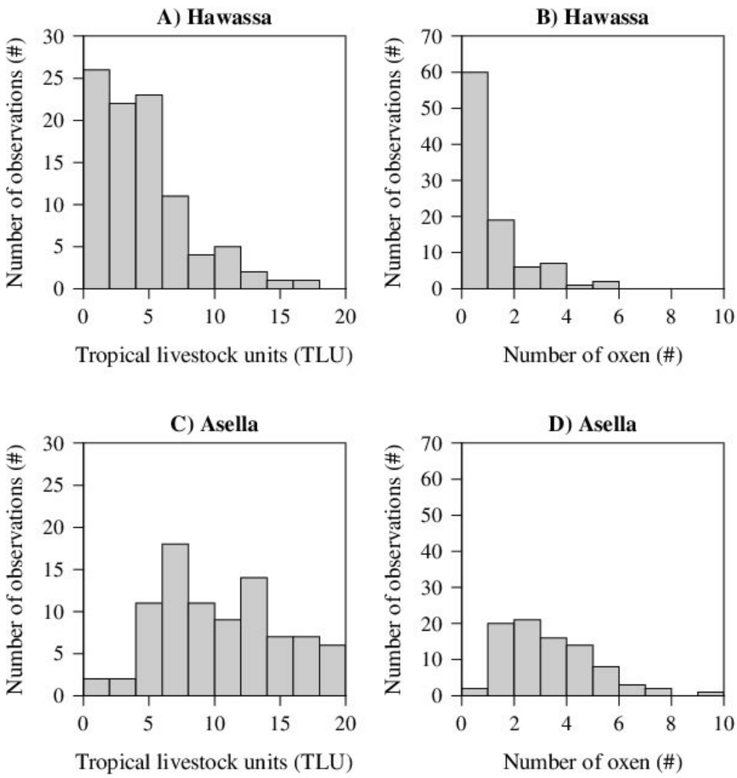

Figure D2: Histograms of tropical livestock unit (TLU) and oxen ownership. Panels A) and B) refer to farms in Hawassa and panels C) and D) to farms in Asella.

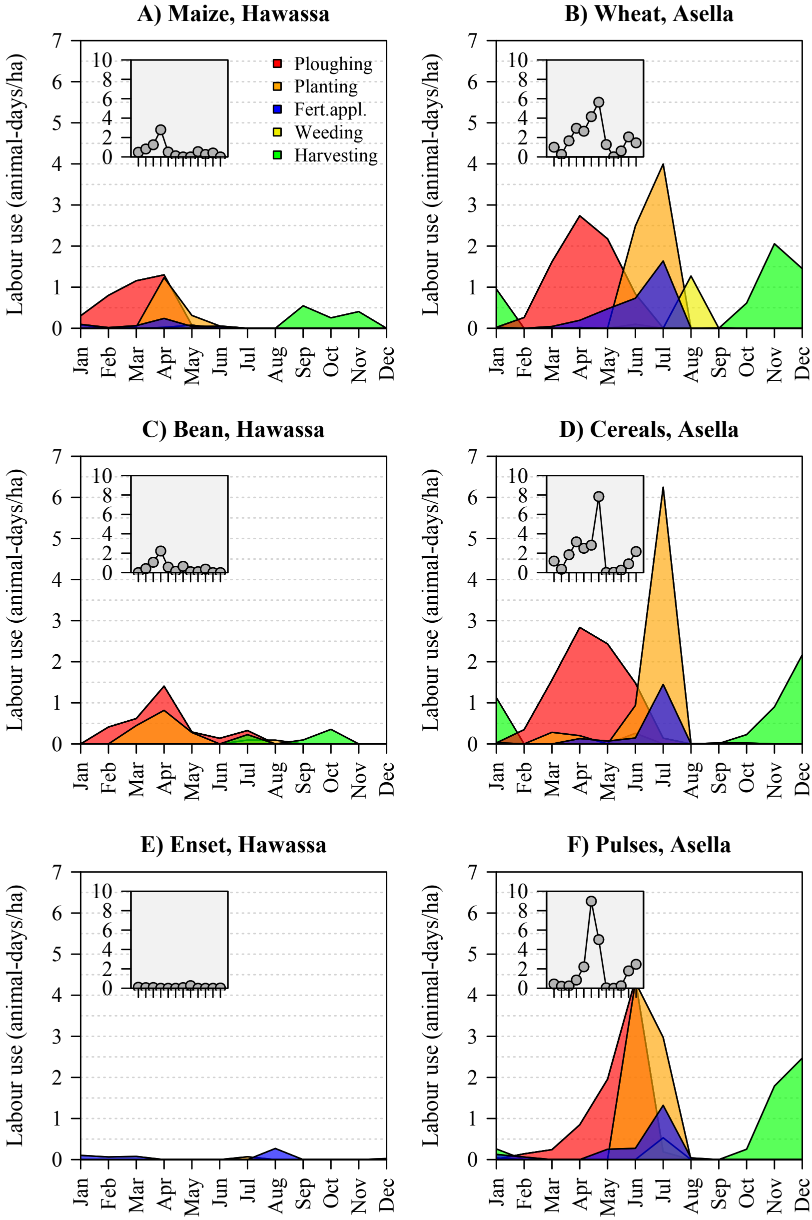

Figure D3: Labour calendar (animal draught power) for crop-specific management operations across smallholder farms in Hawassa and Asella, Southern Ethiopia. The main crops in Hawassa are A) maize, C) bean and E) enset, in Asella B) wheat, D) other cereals such as barley and tef and F) pulses such as pea and faba bean. The inset figures show the total labour used per month.

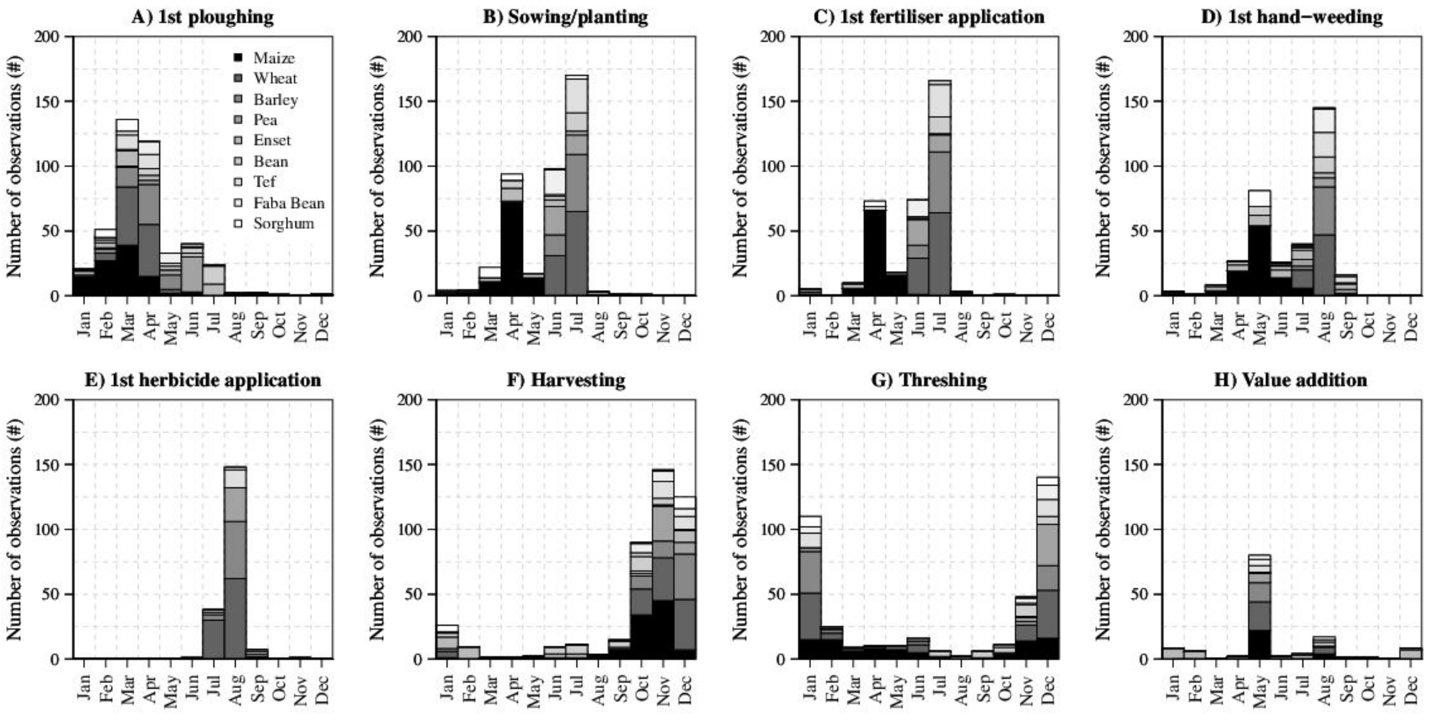

Figure D4: Timing (month) of management operations per crop across smallholder farms in Hawassa (maize, enset and bean) and Asella (wheat, barley, tef, sorghum, pea and faba bean), Southern Ethiopia: A) 1^{st} ploughing, B) sowing/planting, C) 1^{st} fertiliser application, D) 1^{st} hand-weeding, E) 1^{st} herbicide application, F) harvesting, G) threshing and H) value addition. Legend of grey scale is provide in A).

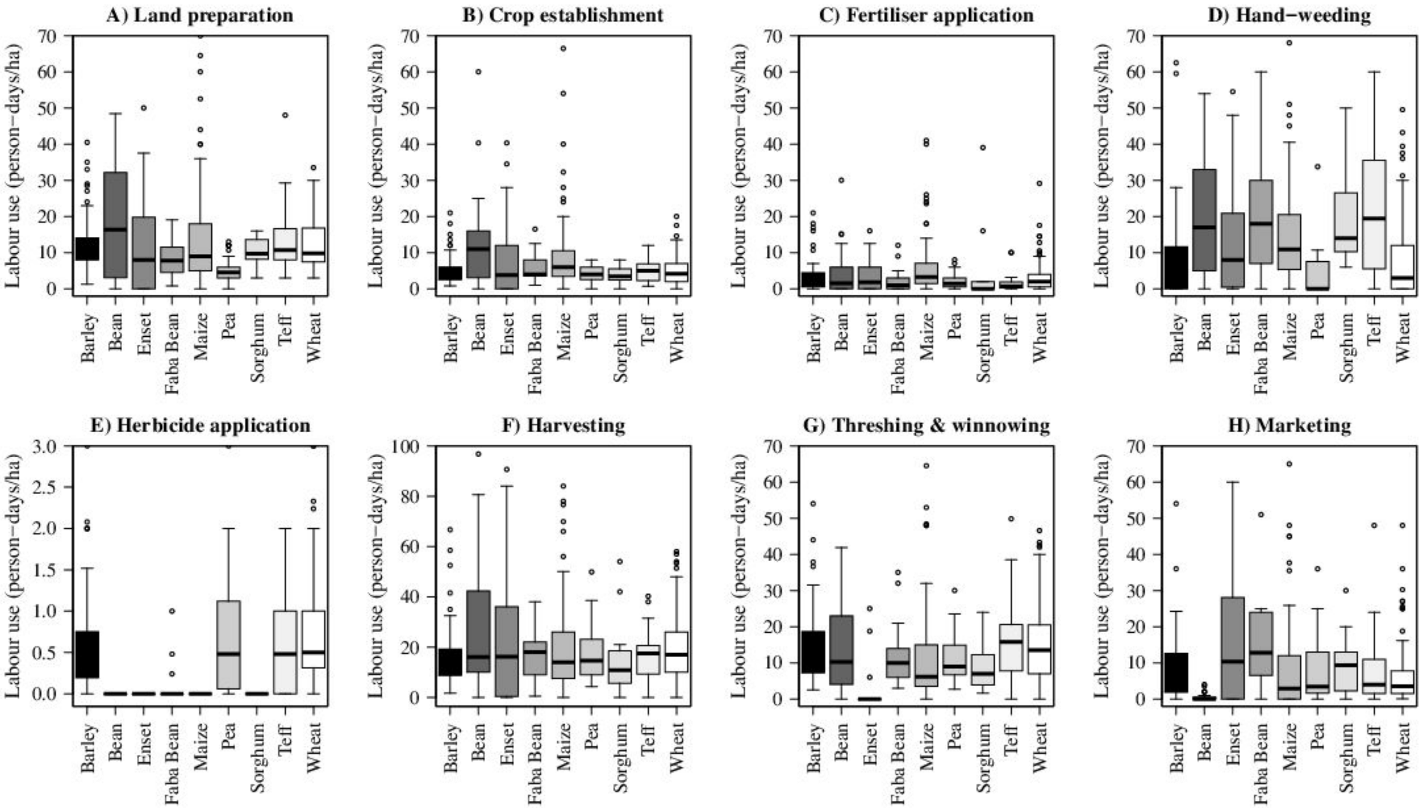

Figure D5: Labour use (person-days ha^{-1}) per crop across smallholder farms in Hawassa (maize, enset and bean) and Asella (wheat, barley, tef, sorghum, pea and faba bean), Southern Ethiopia: A) land preparation, B) crop establishment, C) fertiliser application, D) hand-weeding, E) herbicide application, F) harvesting, G) threshing and winnowing and H) marketing.

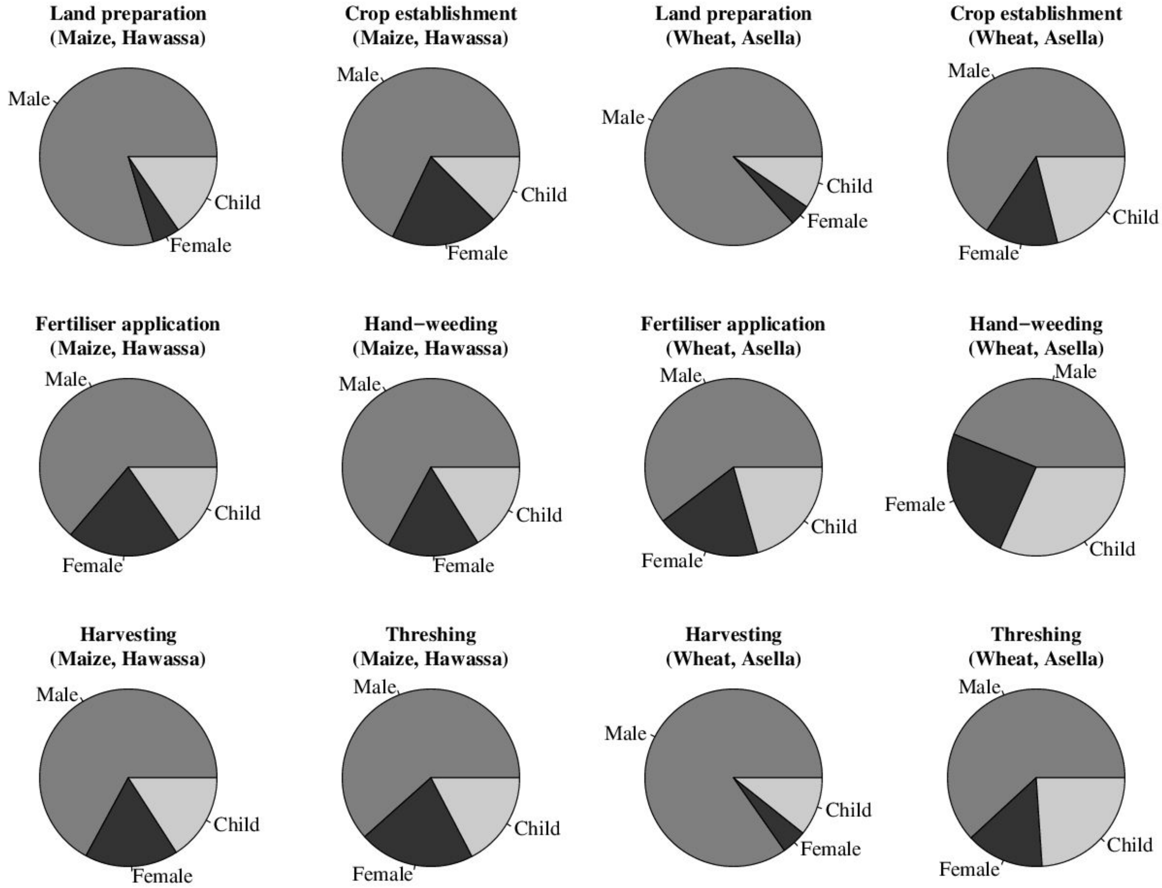

Figure D6: Proportion of male, female and child labour to crop management operations for maize in Hawassa (pie charts on the two most left columns) and wheat in Asella (pie charts on the two most right columns), Southern Ethiopia.

Table D1: Descriptive statistics of family adult labour, hired adult labour, family child labour (person-days ha^{-1}), animal draught (animal-days ha^{-1}) and machinery use (machine-days ha^{-1}) per crop in Asella and Hawassa, Southern Ethiopia. Standard deviations are presented in italics. 'n.a.' = not applicable

	Asella Wheat		**Asella Barley**		**Asella Tef**		**Asella Sorghum**		**Asella Pea**		**Asella Faba Bean**		**Hawassa Maize**		**Hawassa Bean**		**Hawassa Enset**	
Land preparation																		
1st ploughing (month)	3.51	*1.01*	3.93	*1.02*	3.59	*0.89*	2.50	*0.63*	5.57	*0.95*	4.39	*0.84*	2.77	*1.47*	5.60	*1.87*	4.06	*2.08*
Adult family	9.46	*7.40*	9.28	*6.67*	7.50	*5.56*	7.37	*4.84*	3.94	*2.90*	5.43	*3.88*	11.70	*14.41*	17.81	*19.56*	10.27	*15.30*
Adult hired	2.55	*5.23*	1.98	*4.60*	2.99	*4.51*	1.53	*3.57*	0.90	*1.82*	1.61	*3.44*	3.66	*12.95*	0.38	*2.12*	2.32	*7.65*
Child family	1.26	*2.37*	1.46	*3.02*	2.18	*3.20*	1.60	*2.53*	0.46	*0.95*	1.35	*3.29*	2.92	*10.06*	2.60	*8.12*	1.71	*4.13*
Animal draught	21.43	*14.54*	22.34	*18.20*	21.27	*17.03*	18.19	*12.34*	8.95	*6.38*	12.23	*8.18*	7.06	*18.11*	4.18	*8.86*	0.00	*0.00*
Machinery	0.04	*0.31*	0.00	*0.00*	0.01	*0.03*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.13	*1.26*	0.00	*0.00*	0.00	*0.00*
Crop establishment																		
Sowing (month)	6.59	*0.78*	6.75	*0.47*	6.96	*0.20*	3.57	*0.85*	6.37	*0.49*	5.91	*1.12*	3.96	*0.66*	5.65	*1.80*	4.95	*1.68*
Adult family	3.20	*2.44*	3.76	*3.40*	2.96	*2.24*	2.92	*2.20*	3.05	*2.05*	3.95	*3.58*	7.50	*10.37*	10.01	*11.62*	5.67	*7.23*
Adult hired	0.69	*1.86*	0.48	*0.98*	1.17	*1.71*	0.19	*0.54*	0.55	*0.99*	0.84	*1.97*	0.63	*2.44*	0.23	*1.31*	0.78	*2.76*
Child family	1.04	*1.86*	0.97	*1.52*	1.06	*1.38*	0.56	*0.98*	0.48	*0.69*	1.24	*2.06*	0.93	*1.54*	1.89	*4.71*	1.22	*3.09*
Animal draught	7.11	*6.37*	6.56	*4.60*	10.49	*9.35*	5.51	*5.81*	7.02	*5.45*	5.98	*3.72*	2.22	*4.11*	1.81	*4.08*	0.20	*1.04*
Machinery	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*
Weeding																		
1st weeding (month)	7.79	*0.45*	7.93	*0.27*	8.08	*0.64*	5.31	*0.95*	7.80	*0.86*	7.90	*0.54*	4.95	*1.15*	6.48	*1.67*	5.84	*1.98*
Adult family	5.30	*8.25*	4.98	*7.02*	12.73	*15.17*	15.78	*23.59*	2.78	*4.37*	9.92	*13.20*	9.90	*12.52*	18.84	*20.36*	11.14	*16.72*
Adult hired	0.89	*3.20*	1.33	*5.26*	7.73	*16.52*	3.54	*7.34*	0.58	*1.35*	7.49	*13.48*	3.35	*8.88*	0.15	*0.87*	6.30	*35.25*
Child family	2.87	*6.87*	1.69	*3.55*	4.07	*7.97*	5.92	*9.42*	0.72	*2.42*	5.42	*6.85*	1.58	*3.56*	2.77	*4.76*	1.51	*3.80*
Animal draught	0.68	*6.53*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.68	*3.20*	0.13	*0.92*	0.00	*0.00*	0.00	*0.00*
Machinery	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*
Herbicide application																		
1st application (month)	7.69	*0.49*	7.94	*0.32*	7.89	*0.66*	n.a.	*n.a.*	8.00	*0.38*	8.33	*0.58*	n.a.	*n.a.*	n.a.	*n.a.*	n.a.	*n.a.*
Adult family	0.51	*0.56*	0.46	*0.63*	0.74	*1.61*	0.00	*0.00*	0.72	*1.06*	0.07	*0.24*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*
Adult hired	0.11	*0.33*	0.06	*0.23*	0.18	*0.44*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*
Child family	0.11	*0.26*	0.07	*0.20*	0.01	*0.05*	0.00	*0.00*	0.19	*0.78*	0.01	*0.05*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*
Animal draught	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*
Machinery	0.35	*0.77*	0.51	*1.34*	0.28	*0.80*	0.00	*0.00*	0.46	*1.16*	0.27	*0.94*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*

Table D1: Descriptive statistics of family adult labour, hired adult labour, family child labour (person-days ha^{-1}), animal draught (animal-days ha^{-1}) and machinery use (machine-days ha^{-1}) per crop in Asella and Hawassa, Southern Ethiopia. Standard deviations are presented in italics. 'n.a.' = not applicable (*continued*)

	Asella Wheat		**Asella Barley**		**Asella Tef**		**Asella Sorghum**		**Asella Pea**		**Asella Faba Bean**		**Hawassa Maize**		**Hawassa Bean**		**Hawassa Enset**	
Fertiliser application																		
1st application (month)	6.56	*0.98*	6.84	*0.41*	6.96	*0.20*	3.80	*0.45*	6.03	*1.38*	6.19	*0.40*	4.11	*0.63*	5.91	*1.62*	n.a.	*n.a.*
Adult family	2.43	*3.44*	2.34	*3.68*	1.48	*2.67*	3.88	*10.17*	1.79	*2.04*	1.94	*3.24*	4.49	*6.28*	3.20	*5.68*	3.06	*3.99*
Adult hired	0.35	*1.46*	0.29	*1.14*	0.07	*0.24*	0.00	*0.00*	0.02	*0.09*	0.00	*0.00*	0.55	*2.66*	0.08	*0.43*	0.15	*0.60*
Child family	0.73	*1.66*	0.78	*2.27*	0.08	*0.20*	0.03	*0.13*	0.26	*0.49*	0.33	*0.84*	0.82	*1.49*	0.90	*2.71*	0.38	*1.22*
Animal draught	3.07	*6.95*	2.08	*4.54*	0.42	*2.10*	3.07	*9.80*	1.73	*4.02*	1.43	*3.57*	0.78	*2.58*	0.00	*0.00*	0.48	*1.72*
Machinery	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.11	*0.85*	0.00	*0.00*	0.00	*0.00*
Harvesting																		
Harvesting (month)	10.63	*2.43*	11.05	*2.05*	10.15	*3.35*	8.38	*5.16*	11.17	*0.51*	10.43	*2.23*	10.31	*1.04*	8.76	*1.79*	5.24	*4.36*
Adult family	6.30	*6.65*	6.63	*6.33*	6.87	*8.52*	8.91	*9.96*	11.25	*7.94*	9.54	*7.21*	12.18	*14.23*	24.33	*24.33*	17.89	*28.60*
Adult hired	11.30	*12.17*	7.48	*11.15*	8.23	*9.35*	2.38	*4.65*	2.80	*4.61*	3.47	*5.62*	4.11	*10.82*	0.23	*0.96*	4.71	*11.33*
Child family	2.11	*5.19*	2.25	*4.12*	2.11	*4.94*	3.60	*5.13*	3.15	*3.01*	3.87	*5.87*	2.51	*4.57*	5.49	*14.45*	2.38	*6.37*
Animal draught	8.14	*9.43*	6.05	*5.21*	4.57	*5.41*	6.57	*5.89*	7.21	*5.35*	5.50	*5.77*	1.15	*3.60*	1.05	*2.99*	0.00	*0.00*
Machinery	0.16	*1.09*	0.00	*0.00*	0.00	*0.00*	0.06	*0.25*	0.80	*4.73*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*
Threshing																		
Threshing (month)	6.85	*5.17*	5.25	*5.13*	7.00	*5.31*	5.88	*5.48*	11.00	*3.07*	8.87	*4.63*	6.10	*4.18*	9.68	*1.89*	2.00	*n.a.*
Adult family	8.06	*6.07*	8.91	*7.90*	8.47	*6.89*	4.83	*3.56*	7.09	*5.30*	7.41	*6.88*	6.03	*7.78*	9.93	*8.64*	0.89	*4.03*
Adult hired	3.75	*6.22*	2.32	*3.91*	3.98	*8.43*	1.67	*3.44*	1.73	*2.85*	2.36	*3.92*	2.24	*7.26*	0.73	*3.53*	2.30	*16.13*
Child family	3.72	*5.24*	3.68	*4.48*	4.77	*7.37*	2.12	*2.67*	2.43	*2.39*	2.47	*3.82*	1.42	*3.35*	2.97	*6.01*	1.28	*8.09*
Animal draught	22.90	*23.68*	20.79	*15.09*	27.11	*23.84*	13.05	*12.80*	18.42	*15.95*	15.82	*9.95*	0.09	*0.83*	0.18	*1.04*	0.00	*0.00*
Machinery	0.01	*0.10*	0.00	*0.00*	1.15	*5.88*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	0.00	*0.00*	10.85	*65.29*
Marketing																		
Marketing (month)	5.56	*1.19*	5.19	*0.75*	7.00	*2.35*	6.20	*1.64*	5.00	*0.00*	5.86	*1.46*	5.41	*1.05*	5.00	*n.a.*	5.24	*4.36*
Adult family	6.06	*8.75*	9.93	*18.68*	5.60	*6.60*	8.17	*8.84*	7.03	*8.68*	20.69	*29.71*	8.15	*18.39*	0.53	*1.04*	14.34	*17.58*
Adult hired	0.17	*1.24*	0.26	*1.57*	0.34	*1.08*	0.00	*0.00*	0.05	*0.26*	0.01	*0.05*	1.71	*14.28*	0.00	*0.00*	2.78	*10.92*
Child family	0.45	*2.62*	0.53	*2.40*	0.18	*0.80*	0.31	*1.15*	0.51	*3.04*	2.20	*10.23*	0.76	*4.10*	0.00	*0.00*	0.31	*2.14*
Animal draught	4.30	*8.33*	5.47	*10.85*	3.56	*5.61*	2.07	*4.20*	4.01	*7.44*	7.74	*25.52*	0.22	*1.03*	0.26	*0.78*	0.62	*4.28*
Machinery	0.10	*0.83*	0.01	*0.04*	0.00	*0.00*	0.00	*0.00*	0.03	*0.17*	0.73	*3.41*	0.01	*0.05*	0.00	*0.00*	48.04	*109.35*

Table D2: Significant differences in input use between Y_{HF}, Y_{AF} and Y_{LF} for maize in Hawassa and wheat in Asella. Weed control refers to the labour used for the 1st hand-weeding in maize and herbicide application in wheat.

	Yield level (percentile)	N applied (kg N ha^{-1})	P applied (kg P ha^{-1})	Ploughing & sowing (animal-days ha^{-1})	Weed control (person-days ha^{-1})
Hawassa					
Maize	Y_{HF}	111.2 a	40.9 a	18.3 a	16.6 a
Maize	Y_{AF}	58.8 b	19.2 b	7.3 a	9.7 a
Maize	Y_{LF}	41.3 b	13.3 b	2.3 a	6.2 a
Asella					
Wheat	Y_{HF}	57.1 a	41.7 a	49.2 a	1.7 a
Wheat	Y_{AF}	42.8 a	30.1 b	25.7 b	0.7 b
Wheat	Y_{LF}	35.5 a	27.8 b	20.1 b	0.6 b

Table D3: Significant differences in farm resources, crop yield and input use between farm types differing in oxen ownership. n.a. = not applicable.

	Zero pair	One pair	Two or more pairs
Hawassa			
Farm assets (#)	4.6 c	7.2 b	12.7 a
Cultivated land (ha)	0.7 a	0.7 a	0.8 a
Household size (#)	6.6 a	6.7 a	5.2 a
Maize yield (t ha^{-1})	1.7 b	2.7 ab	3.2 a
Maize N use (kg N ha^{-1})	52.7 b	72.3 ab	120.1 a
Maize ploughing & sowing (animal-days ha^{-1})	1.8 b	11.6 a	5.5 ab
Maize ploughing & sowing (person-days ha^{-1})	14.1 a	13.6 a	13.4 a
Maize 1st hand-weeding (person-days ha^{-1})	10.4 a	10.3 a	14.9 a
Maize 2nd hand-weeding (person-days ha^{-1})	5.2 a	10.2 a	7.6 a
Asella			
Farm assets (#)	n.a.	10.9 b	16.2 a
Cultivated land (ha)	n.a.	1.3 b	2.1 a
Household size (#)	n.a.	5.3 b	6.8 a
Wheat yield (t ha^{-1})	n.a.	2.9 a	3.5 a
Wheat N use (kg N ha^{-1})	n.a.	44.3 a	50.6 a
Wheat ploughing & sowing (animal-days ha^{-1})	n.a.	12.6 a	14.7 a
Wheat ploughing & sowing (person-days ha^{-1})	n.a.	8.8 a	10.1 a
Wheat 1st herbicide app. (person-days ha^{-1})	n.a.	0.7 a	0.7 a
Wheat 1st hand-weeding (person-days ha^{-1})	n.a.	13.4 a	11.7 a

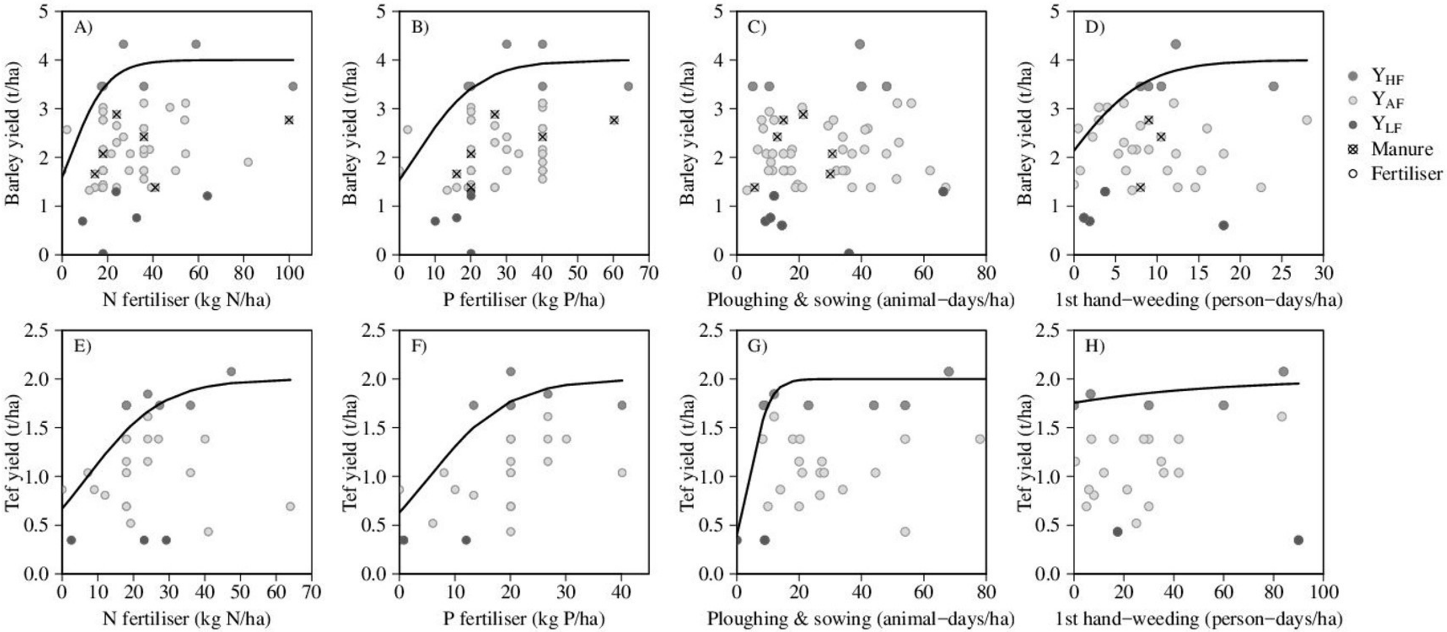

Figure D7: Yield response to N, P and labour for barley (A - D) and tef in Asella (E - H), Southern Ethiopia. Each dot refers to a single farm in the year 2012. Solid lines are fitted boundary lines using the model of Fermont et al. (2009).

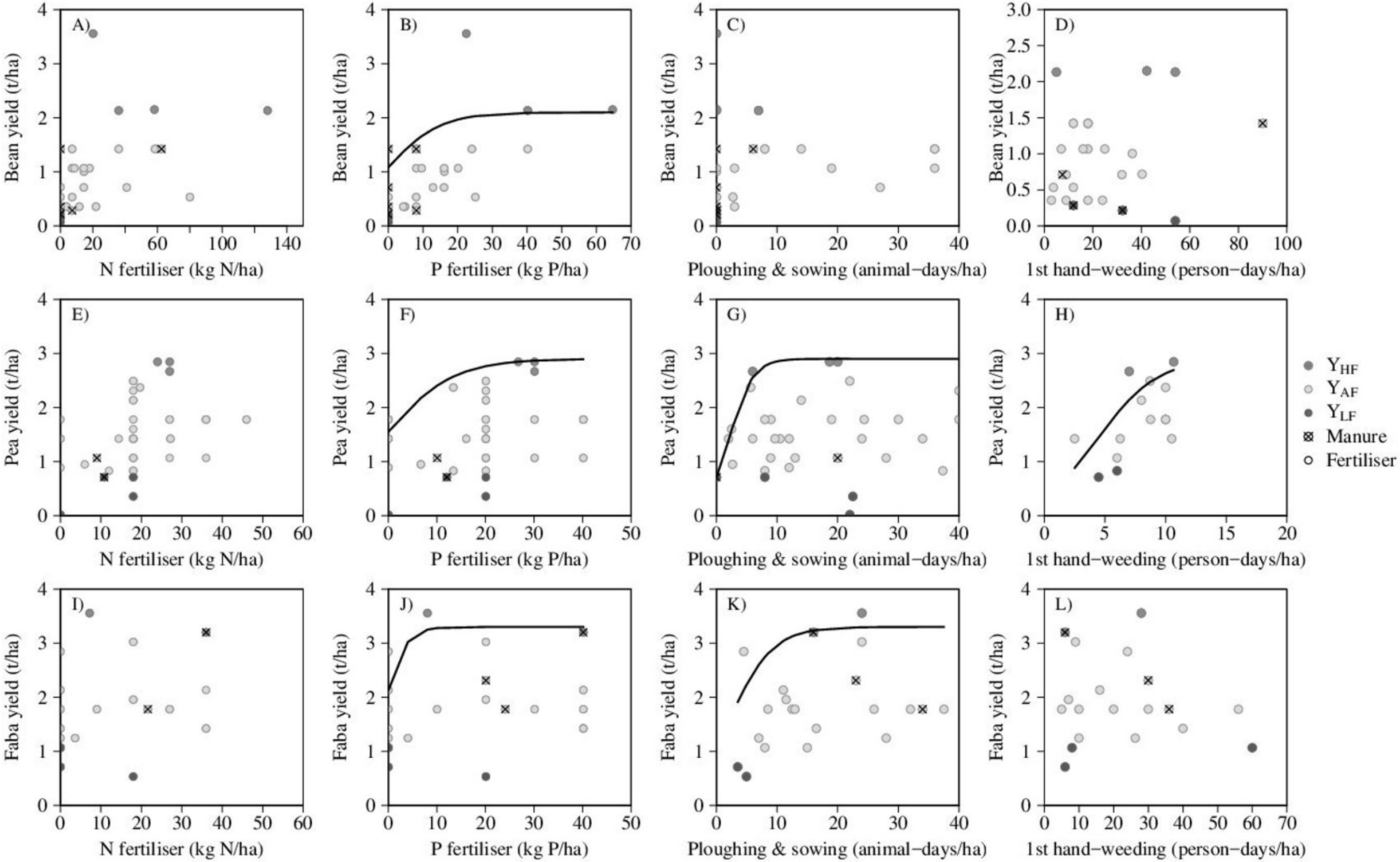

Figure D8: Yield response to N, P and labour for bean in Hawassa (A - D), pea in Asella (E - H) and faba bean in Asella (I - L), Southern Ethiopia. Each dot refers to a single farm in the year 2012. Solid lines are fitted boundary lines using the model of Fermont et al. (2009).

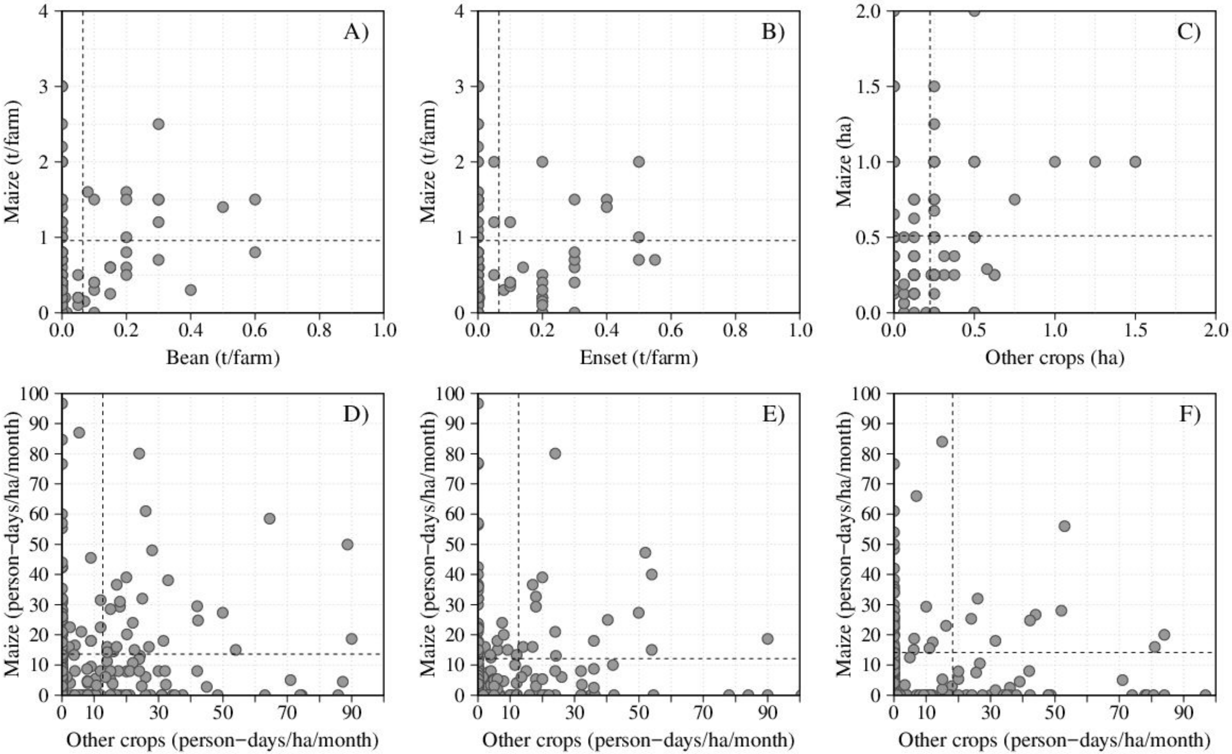

Figure D9: Land and labour productivity of maize as a function of land and labour productivity of bean and/or enset in Hawassa, Southern Ethiopia: A) relationship between production of maize and bean, B) relationship between production of maize and enset, C) crop area shares, D) person-days used in the months of ploughing and planting, E) in the months of hand-weeding and F) in the months of harvesting.

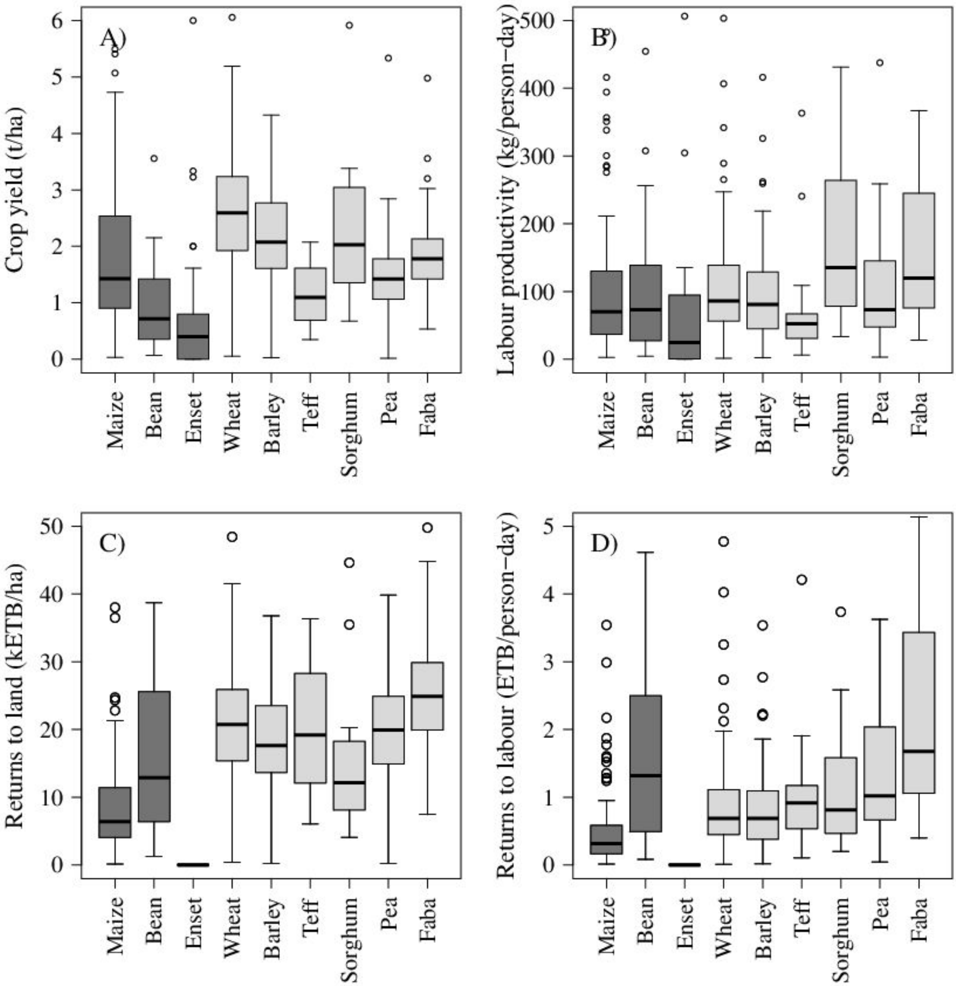

Figure D10: Crop performance in smallholder farms of Southern Ethiopia. The performance indicators considered were A) crop yields (t ha^{-1}), B) labour productivity (kg $person\text{-}day^{-1}$), C) returns to land (ETB ha^{-1}) and D) returns to labour (ETB $person\text{-}day^{-1}$). Crops cultivated in Hawassa are highlighted in dark grey and in Asella in light grey.

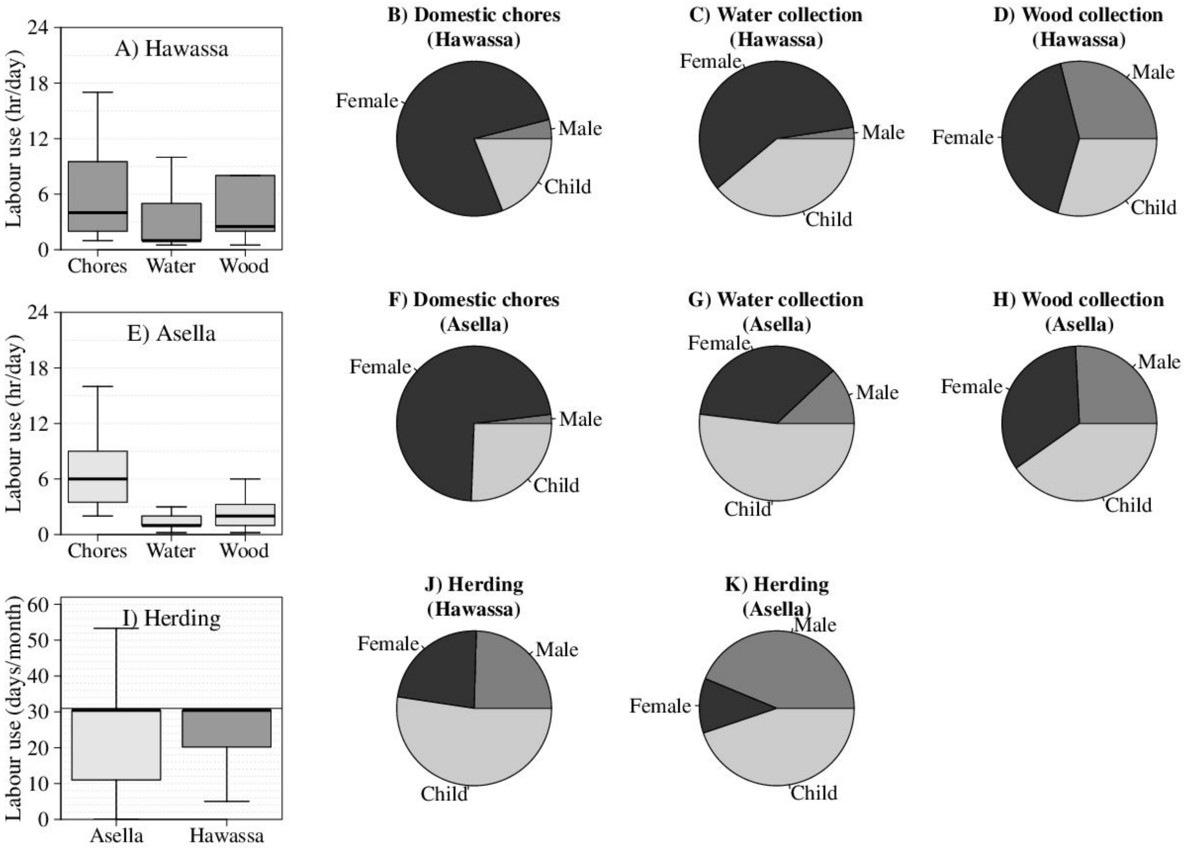

Figure D11: Human labour used by households in Hawassa and Asella for household activities (domestic chores, water collection and wood collection) and herding. Pie charts show the relative contribution of male, female and child labour to each activity.

Summary

Yield gap analyses have been performed traditionally with mechanistic crop models to understand the relative contribution of growth-defining, -limiting and -reducing factors to actual yields. These were conducted mostly at field level and directly up-scaled to the regional level, without considering explanatory factors at farm(ing) system level. However, farmers make decisions on resource allocation and prioritization of crop management across the entire farm and these decisions determine crop and farm performance as well as possible trade-off between both. The objective of this thesis was to gain insights into yield gaps at farm(ing) systems level in order to identify opportunities for sustainable intensification at local level.

Three contrasting case studies representing a gradient of intensification were selected for this purpose, namely mixed crop-livestock systems in Southern Ethiopia, rice based-farming systems in Central Luzon (Philippines) and arable farming systems in the Netherlands. Methods of frontier analysis were used in combination with concepts of production ecology to decompose yield gaps into efficiency, resource and technology yield gaps using individual farm data. The same methodological protocol was applied for the main crops across the three case studies. Moreover, different analyses were conducted at the farm level to understand how farmers' objectives, farm(er) characteristics, rotational effects and resource constraints interact with yield gaps and management practices at crop level.

A literature review of 50 peer-reviewed articles on yield gap analysis was conducted to summarize the yield gap explaining factors identified so far and to assess opportunities for bottom-up data collection approaches (**Chapter 2**). Yield gap explaining factors vary among regions and crops and were largely determined by the objective of the study, prior knowledge about possible explaining factors and method applied. However, edaphic (e.g. soil fertility) and management factors (e.g. nutrient management) were more often considered to explain the yield gap than factors capturing farm(er) characteristics or socio-economic conditions. In terms of nutrient management, yield gap studies focused mostly on the quantity of fertiliser used rather than on the timing of application, but the latter, when considered, explained the yield gap more often than the former. Crop models were the preferred method used for yield

gap analysis, even though they were limited in the type and the number of factors considered.

A theoretical framework combining methods of frontier analysis and concepts of production ecology was developed to decompose crop yield gaps (**Chapter 3**). A proof of concept was performed for rice farming in Central Luzon, the Philippines, using a combination of stochastic frontier analysis and crop modelling applied to a historical household survey. Between 1979 - 2012, the rice yield gap was 3.2 and 4.8 t ha^{-1} (or ca. 55 and 49% of the simulated climatic potential yield, Yp) in the wet and dry season, respectively. The contribution of the efficiency, resource and technology yield gaps was equally important in the wet season (i.e. 18, 13 and 15% of Yp, respectively) while in the dry season the contribution of the technology yield gap (24% of Yp) was larger than that of the efficiency and resource yield gaps (each ca. 13% of Yp). The efficiency yield gap was partly explained by untimely application of mineral fertiliser and plant protection agents; the resource yield gap by lower N, P and K rates in lowest yielding fields in the dry season; and the technology yield gap by a lack of use of the proper rice varieties (wet season) and sub-optimal application of water and nutrients compared to the requirements to achieve Yp (dry season).

An integrated assessment of rice yield gaps in Central Luzon was conducted at field, farm and regional levels to link management and edaphic factors explaining yield gaps according to Chapter 3 to farm(er) characteristics and socio-economic conditions (**Chapter 4**). The two hypotheses tested were: 1) there are trade-offs between closing rice yield gaps and maximising labour productivity, gross margin or N use efficiency and 2) farm(er) characteristics and socio-economic conditions affect the management practices used by farmers. The Central Luzon Loop Survey and another, more recent and spatially representative, household survey were used to test these hypotheses. Trends over time indicate rice yields, labour productivity and the proportions of hired labour and rice sold increased over the past half-century. Closing rice yield gaps incurred trade-offs with labour productivity and N use efficiency, but not as much with gross margin. Yield levels maximising labour productivity or N use efficiency were 25 - 35% lower than Yp while yield levels maximising gross margin were ca. 20% lower than Yp in both seasons. The results regarding the second hypothesis were not always conclusive which suggests that management practices used by farmers were clearly not constrained by factors at farm or regional levels.

Yield gaps of the major arable crops in the Netherlands were decomposed into efficiency, resource and technology yield gaps, and explained at crop rotation level using information on crop area shares and farmers' objectives (**Chapter 5**). The methodological approach used in Chapter 3 was applied in combination with regression analysis to data of specialized Dutch arable farms covering the period 2008 - 2012. The yield gap ranged between ca. 25 and 40% of Yp for respectively winter wheat and starch potato. The yield gap was explained mostly by the 1) efficiency yield gap for ware potato and spring onion; 2) efficiency and technology yield gaps for sugar beet and cereals; 3) resource yield gap for seed potato (a crop for which tuber size is particularly important), and 4) technology yield gap for starch potato. Rotational effects on crop yields yielded inconclusive results, which may suggest that agronomic principles became obscured at ’systems level’ due to confounding with e.g. rented land or short-term economic performance. Significant differences in farm performance were observed between farms maximising N or energy production, gross margin or N use efficiency or minimizing labour use. As an example, gross margin maximising, labour minimising and N use efficiency maximising farms can, respectively, increase N production by ca. 17, 34 and 21%. Differences in farm performance were explained mostly by differences in crop area shares rather than by yields of individual crops.

A similar approach was followed to decompose, and explain, the yield gap of maize and wheat across smallholder farms in Southern Ethiopia (**Chapter 6**). The analyses built upon a household survey requesting detailed information on labour use at crop and farm levels in the farming systems around Hawassa (maize) and Asella (wheat). The actual yield of maize and wheat was 1.6 and 2.8 t ha^{-1}, which is much smaller than the water-limited yield of 7.0 and 10.0 t ha^{-1}, respectively. The technology yield gap explained most of the yield gap of maize and wheat, ca. 45 and 52% of Yw, respectively. The efficiency yield gap was ca. 20% of Yw for both crops and associated with sowing date for maize and hired labour for wheat. In Hawassa, households with more oxen applied more N to maize, and achieved higher maize yields, than households with fewer oxen. In contrast, in Asella households with more oxen cultivated more land than households with fewer oxen and there were no differences in input use and wheat yields between the two groups. The crops cultivated in Hawassa (maize, bean and enset) exhibited a complementary use of labour while the dominant crops in Asella (small grains and pulses) ’competed’ for labour in specific months, which indicates there may be still trade-offs in labour allocation for the different crops in this site.

Sustainable intensification and yield gaps are contentious topics in agronomy, particularly because they have been used to justify interventions in developing countries without thorough consideration of the context in which smallholder farmers operate

(**Chapter 7**). Yield gap analysis for smallholder farmers in the tropics needs to acknowledge both the riskiness and variability of improvement options to narrow the yield gap and the role that such yield gains might play within local livelihoods. An example, and reflection, about how to address the aforementioned aspects is provided using empirical data from maize-based farming systems in Western Kenya and rice-based farming systems in Central Luzon, Philippines. Participatory research conducted in Western Kenya questions the appropriateness of yield measurements taken from on-station trials to estimate actual and water-limited yields at regional level and shows the challenges in assessing the real magnitude of yield gaps in farmers' fields in case maize yields and yield gaps are expressed as cumulative probabilities. Large rice yield gaps persist in Central Luzon while at the same time there appear to be few incentives to close them given the marginal economic returns to labour for farm work and the off-farm opportunities available to farmers.

Embedding yield gaps within the broader farm level opportunities and constraints is a knowledge and data intensive exercise. Multiple concepts are required for this purpose and different methodologies need to be combined with local knowledge about the farming system to thoroughly explain and understand the causes behind yield gaps in farmers' fields (**Chapter 8**). In short, the efficiency yield gap explained most of the yield gap in Dutch arable farming, the technology yield gap was most important across smallholder farms in Southern Ethiopia and the three intermediate yield gaps were equally relevant in rice-based farming systems of Central Luzon. Little trade-offs between yield gap closure and gross margin per ha were observed in the three case studies, while yields maximising labour productivity (gross margin per labour unit) and N use efficiency were 30 - 40% lower than the potential yield. Thus, it remains a challenge for farmers to reconcile intensification through yield gap closure with high resource use efficiency and the use of labour-saving technologies.

Three 'directions of change' reveal themselves regarding the future of agricultural systems: sustainable intensification ('more output with less input'), extensification ('same output with less input') and intensification ('more output with more input'). There is a large scope to improve N use efficiency through 1) sustainable intensification and intensification in Hawassa and Central Luzon, 2) intensification to avoid mining soil N in Asella and 3) sustainable intensification, or even extensification, in the Netherlands. This supports the conventional wisdom that intensification of agriculture needs to occur in the 'developing South' while a focus on improving sustainability is more appropriate in the 'developed North', based on sustainable intensification or even extensification. This will need to be achieved by individual farms who are far too often required to prioritise short-term needs over long-term aspirations. Yield gap analysis will remain an important exercise to understand how growth-defining, -limiting and -reducing factors affect actual yields in farmers' fields but it needs to consider explicitly the broader livelihood context within which farming takes place as well as possible explanations at 'higher systems levels'.

Samenvatting

Analyses die het verschil tussen de potentiële en actuale opbrengsten (de zogenaamde opbrengstverschillen of 'yield gaps') verklaren zijn traditioneel uitgevoerd met mechanistische gewasmodellen. Deze helpen om de relatieve bijdrage van groei-definiërende, beperkende en reducerende factoren aan de actuele opbrengsten te begrijpen. Deze analyses worden meestal op veldniveau uitgevoerd en direct opgeschaald naar het regionale niveau, zonder factoren op het bedrijfsniveau in ogenschouw te nemen. Boeren nemen echter beslissingen over de toewijzing van hulpbronnen en prioritering van het gewasmanagement over het hele bedrijf. Deze beslissingen bepalen de opbrengst van landbouwgewassen en landbouwbedrijven, evenals de balans tussen opbrengsten en kosten. Het doel van dit proefschrift was om inzicht te krijgen in het verschil tussen de potentiële en actuele opbrengsten op het bedrijfsniveau om kansen voor duurzame intensivering op lokaal niveau te identificeren.

Voor dit doel zijn drie contrasterende case studies onderzocht die een gradiënt in intensivering representeren, namelijk gemengde systemen in Zuid-Ethiopië, rijstsystemen in de Filipijnen (Central Luzon) en akkerbouwsystemen in Nederland. Econometrische methodes zijn gebruikt in combinatie met concepten van productie ecologie om het verschil tussen potentiële en actuele opbrengsten te ontleden in verschillen in efficiëntie, gebruik van inputs en gebruik technologie, met behulp van individuele bedrijfsdata. Dezelfde methode is toegepast op de belangrijkste gewassen in de drie case studies. Bovendien zijn op het bedrijfsniveau verschillende analyses uitgevoerd om te begrijpen hoe de doelstellingen van de boeren, hun karakteristieken, rotaties en beperkingen in hulpbronnen invloed hebben op de opbrengstverschillen en gewasmanagement.

Een literatuuroverzicht van 50 wetenschappelijke artikelen over het verschil tussen potentiële en actuele opbrengsten is gemaakt om factoren die het verschil verklaren te beoordelen en om kansen te identificeren voor innovatieve dataverzameling (**Hoofdstuk 2**). De verklarende factoren varieerden tussen regio's en gewassen en werden grotendeels bepaald door de doelstelling van de studie, voorafgaande kennis over mogelijke verklarende factoren en de gebruikte methode. Echter, fac-

toren gerelateerd aan bodem (bijvoorbeeld bodemvruchtbaarheid) en gewasmanagement (bijvoorbeeld nutriëntenbeheer) werden vaker als mogelijke verklarende factoren beschouwd dan bedrijfskarakteristieken of sociaal-economische omstandigheden. Wat het nutriëntbeheer betreft, richtten de studies zich vooral op de hoeveelheid gebruikte meststoffen en minder op het tijdstip van toepassing; maar als het tijdstip in overweging werd genomen, verklaarde dit vaker het opbrengstverschil dan de hoeveelheid. Gewasmodellen waren de meest gebruikte methode voor de analyse van opbrengstverschillen, hoewel ze beperkt waren in het type en aantal mogelijke verklarende factoren.

Een theoretisch kader dat econometrische methoden en concepten van productie ecologie combineert, is ontwikkeld om de verschillen tussen potentiële en actuele opbrengsten te ontleden (**Hoofdstuk 3**). Dit kader werd allereerst toegepast voor rijstsystemen in Central Luzon, de Filipijnen. Een combinatie van 'stochastic frontier analysis' en gewasmodellering is toegepast op een historische dataset met individuele bedrijven. Tussen 1979 en 2012 was de yield gap 3.2 en 4.8 ton per ha (of ongeveer 55 en 49% van de gesimuleerde potentiële opbrengst, Yp), respectievelijk in het natte en droge seizoen. De invloed van de verschillen in efficiëntie, inputs en technologie was even belangrijk in het natte seizoen (respectievelijk 18, 13 en 15% van Yp), terwijl in het droge seizoen de bijdrage van beperkte technologie (24% Yp) groter was dan die van beperkte efficiëntie en inputs (elk ongeveer 13% van Yp). Het opbrengstverschil door beperkte efficiëntie werd deels verklaard door de vroegtijdige toepassing van kunstmest en gewasbeschermingsmiddelen; het input-opbrengstverschil door lagere N, P en K giften in velden met de laagste opbrengsten in het droge seizoen; en het technologie-opbrengstverschil door een gebrek aan de juiste rijstvariëteiten (nat seizoen) en suboptimaal gebruik van water en nutriënten in vergelijking met het gebruik dat nodig is om Yp te halen (droog seizoen).

Om factoren gerelateerd aan gewasmanagement en bodem die opbrengstverschillen verklaren (Hoofdstuk 3), te koppelen aan bedrijfskarakteristieken en sociaaleconomische omstandigheden, is een geïntegreerde evaluatie van de rijstopbrengsten in Central Luzon uitgevoerd op veld, bedrijfs- en regionaal niveau (**Hoofdstuk 4**). De twee geteste hypothesen waren: 1) er zijn afwegingen tussen de verhoging van rijstopbrengsten en het maximaliseren van de arbeidsproductiviteit, de winstmarge of de stikstofefficiëntie en 2) de bedrijfskarakteristieken en sociaaleconomische omstandigheden beïnvloeden het gewasmanagement op bedrijfsniveau. De 'Central Luzon Loop Survey' en een andere, recentere en ruimtelijk representatieve enquête op bedrijfsniveau zijn gebruikt om deze hypothesen te testen. De trend is dat in de afgelopen halve eeuw de rijstopbrengsten, arbeidsproductiviteit, het aandeel ingehuurde arbeid en het aandeel verkochte rijst zijn toegenomen. Het verhogen van rijstopbrengsten had een lagere arbeidsproductiviteit en stikstofefficiëntie tot gevolg, maar had geen negatieve invloed op de winstmarge. De opbrengstniveaus die de arbeidsproductiviteit of stikstofefficiëntie maximaliseren waren 25 tot 35% lager dan Yp, terwijl de opbrengstniveaus die de winstmarge maximaliseren 20% lager waren dan Yp, in zowel het natte als droge seizoen. De resultaten betreffende de tweede hypothese waren niet altijd duidelijk, wat suggereert dat gewasmanagement niet aantoonbaar beperkt werd door factoren op bedrijfs- of regionaal niveau.

Ook de verschillen tussen de potentiële en actuele opbrengsten van de grote akkerbouwgewassen in Nederland werden ontleed in verschillen door efficiëntie, gebruik van hulpbronnen (inputs) en technologie. Tevens werden deze verschillen verklaard op gewasrotatieniveau door gebruik te maken van informatie over de arealen van gewassen in het bouwplan en de doelstellingen van de boeren (**Hoofdstuk 5**). De methodologie die in Hoofdstuk 3 is gebruikt, werd in combinatie met regressieanalyse toegepast op data van gespecialiseerde Nederlandse akkerbouwbedrijven uit de periode 2008-2012. Het verschil tussen de potentiële en actuele opbrengsten varieerde tussen ca. 25 en 40% van Yp voor, respectievelijk, wintertarwe en zetmeelaardappellen. Het opbrengstverschil werd voornamelijk verklaard door 1) efficiëntie voor consumptie-aardappellen en zaai-ui; 2) efficiëntie en technologie voor suikerbieten en granen; 3) gebruik van inputs voor pootaardappellen (een gewas waarvoor de knolgrootte bijzonder belangrijk is), en 4) technologie voor zetmeelaardappellen. Het effect van rotaties op gewasopbrengsten was niet duidelijk, wat er op kan duiden dat agronomische principes op ’systeemniveau’ minder zichtbaar zijn door verstrengeling met bijv. gehuurde grond of economische prestaties op korte termijn. Er zijn significante verschillen in indicatoren gevonden tussen bedrijven die de stikstofproductie, energieproductie, winstmarge of stikstofefficiëntie maximaliseerden of het gebruik van arbeid minimaliseerden. Het maximaliseren van de winstmarge zorgde voor een 17% lagere stikstofopbrengst dan voor bedrijven die de stikstofproductie maximaliseerden, het minimaliseren van arbeid voor 34% lagere stikstofopbrengst, en het maximaliseren van de stikstofefficiëntie voor 21% lagere stikstofopbrengst. Verschillen in de prestaties op het gebied van deze indicatoren werden voornamelijk verklaard door verschillen in gewasaandelen in het bouwplan, en minder door de opbrengsten van individuele gewassen.

Een soortgelijke aanpak werd gevolgd om de opbrengstverschillen van maïs en tarwe op kleine bedrijven in Zuid-Ethiopië te ontleden en te verklaren (**Hoofdstuk 6**).

De analyses maakten gebruik van individuele bedrijfsdata waarin gedetailleerde informatie over arbeidskrachten op veld- en bedrijfsniveau in de landbouwsystemen rondom Hawassa (maïs) en Asella (tarwe) beschikbaar was. De gemiddelde geobserveerde opbrengsten van maïs en tarwe waren 1,6 en 2,8 ton per ha, wat veel lager is dan de waterbeperkte potentiele opbrengsten (Yw) van 7,0 en 10,0 ton per ha, respectievelijk. Het technologie-opbrengstverschil was het grootst, met ca. 45% van Yw voor maïs en 52% voor tarwe. Het efficiëntie-opbrengstverschil was ca. 20% van Yw voor beide gewassen, en was gerelateerd aan de zaaidatum voor maïs en ingehuurde arbeid voor tarwe. In Hawassa hebben huishoudens met meer ossen meer stikstof toegepast op maïs, en deze behaalden hogere maïsopbrengsten dan huishoudens met minder ossen. Daarentegen konden in Asella huishoudens met meer ossen meer areaal telen dan huishoudens met minder ossen, en er waren geen verschillen in het gebruik van inputs en tarweopbrengsten tussen de twee groepen. In Hawassa was het gebruik van arbeid voor de verschillende gewassen (maïs, boon en enset) complementair, terwijl er voor de belangrijkste gewassen in Asella (kleine granen en peulvruchten) in bepaalde maanden concurrentie was om arbeid. Dit laatste betekent dat huishoudens afwegingen moeten maken bij de toewijzing van arbeid voor de verschillende gewassen.

Duurzame intensivering en opbrengstverschillen tussen potentiële en actuele opbrengsten zijn omstreden onderwerpen in de agronomie, vooral omdat ze gebruikt zijn om interventies in ontwikkelingslanden te rechtvaardigen zonder grondig te kijken naar de context waarin kleine boeren actief zijn (**Hoofdstuk 7**). Een analyse van de opbrengstverschillen bij kleine boeren in de tropen moet zowel de risico's als de variabiliteit van de innovaties erkennen om de opbrengst te verhogen, alsmede de rol die dergelijke opbrengstvoordelen kunnen spelen in het lokale levensonderhoud. Een voorbeeld en reflectie omtrent hoe de bovengenoemde aspecten kunnen worden aangepakt, wordt gegeven op basis van empirische gegevens uit maïsgebaseerde landbouwsystemen in West-Kenia en rijstgebaseerde landbouwsystemen in Central Luzon, de Filipijnen. Onderzoek met boeren in West-Kenia toont aan dat opbrengstmetingen op proefstations om de gemiddelde actuele en waterbeperkte potentiele opbrengsten op regionaal niveau te schatten van beperkte waarde zijn. Maar ook bij het beoordelen van opbrengstverschillen op basis van velden van boeren zijn er uitdagingen. In Central Luzon blijven de opbrengsten relatief laag, en tegelijkertijd lijken er weinig prikkels te zijn om deze te verhogen, vanwege de marginale winstmarge per arbeidskracht en de mogelijkheden voor werk buiten de landbouw.

Het verklaren van opbrengstverschillen op bedrijfsniveau is kennis- en data-intensief. Voor dit doel zijn meerdere concepten nodig, en verschillende methoden moeten worden gecombineerd met lokale kennis over het landbouwsysteem om de oorzaken van de opbrengstverschillen in boerenvelden grondig te verklaren en te begrijpen (**Hoofdstuk 8**). Samenvattend, het efficiëntie-opbrengstverschil verklaarde het grootste deel van het verschil tussen potentiële en actuele opbrengsten in de Nederlandse akkerbouw. Het technologie-opbrengstverschil was het belangrijkst bij kleine boerderijen in Zuid-Ethiopië en voor rijstsystemen in Central Luzon waren efficiëntie, gebruik van inputs en technologie even relevant. In de drie case studies werden slechts kleine negatieve relaties tussen de verhoging van de opbrengst en de winstmarge per ha waargenomen, terwijl de opbrengsten bij bedrijven die de arbeidsproductiviteit (winstmarge per arbeidseenheid) of de stikstofefficiëntie maximaliseerden 30 tot 40% lager lagen dan de potentiële opbrengst. Duurzame intensivering blijft dus een uitdaging voor boeren, omdat het lastig is om tegelijkertijd de opbrengst te verhogen, de efficiëntie van het gebruik van hulpbronnen en het gebruik van arbeidsbesparende technologieën.

Er zijn drie mogelijke paden voor de toekomst van de landbouw: duurzame intensivering ('meer output met minder input'), extensivering ('zelfde output met minder input') en intensivering ('meer output met meer input'). Er is een grote ruimte om stikstofefficiëntie te verbeteren door 1) duurzame intensivering en intensivering in Hawassa en Central Luzon, 2) intensivering om bodemuitputting van stikstof te vermijden in Asella en 3) duurzame intensivering, of zelfs extensivering, in Nederland. Dit ondersteunt de conventionele wijsheid dat intensivering van de landbouw moet plaatsvinden in het 'ontwikkelende zuiden', terwijl focus op duurzaamheid belangrijk is in het 'ontwikkelde noorden', gebaseerd op duurzame intensivering of zelfs extensivering. Dit moet worden bereikt door individuele bedrijven, die vaak niet anders kunnen dan het prioriteren van korte termijnbehoeften boven de lange termijn. De analyse van verschillen tussen potentiële en actuele opbrengsten blijft belangrijk om te begrijpen hoe de groei-definiërende, beperkende en reducerende factoren de actuele opbrengsten in velden van boeren beïnvloeden, maar die analyse moet expliciet rekening houden met de bredere context waarin landbouw plaatsvindt, evenals mogelijke verklaringen op 'hogere systeemniveaus'.

Resumo

As análises de *yield gap*, definido como a diferença entre a produtividade potencial e a produtividade actual, são úteis para quantificar a contribuição relativa dos factores de produção que influem no desenvolvimento, crescimento e produtividade de uma cultura num determinado ambiente. A produtividade potencial de uma cultura, adiante representada por Yp, é geralmente estimada com recurso a modelos de simulação do seu desenvolvimento e crescimento sob determinadas condições climáticas e de cultivo. A comparação desta com a produtividade actual é por norma realizada ao nível da parcela e directamente extrapolada para a escala regional, sem levar em consideração dinâmicas e interações ao nível da exploração agrícola e do respectivo sistema de agricultura. Este último aspecto é crucial uma vez que a gestão de recursos visando a optimização de processos é por norma realizada ao nível da exploração, integrando a produtividade de cada cultural / parcela com o desempenho de toda a exploração. Esta tese teve como objectivo aprofundar o conhecimento das principais causas de quebras de produtividade (ou *yield gap*) num vasto número de explorações agrícolas, de modo a identificar oportunidades e restrições para a intensificação sustentável dos sistemas de agricultura actuais.

Três casos de estudo foram selecionados para este objectivo, tais como sistemas agro-pecuários de pequena escala no Sul da Etiópia; sistemas de produção de arroz em Central Luzon, nas Filipinas, e sistemas de produção de culturas arvenses na Holanda. A metodologia desenvolvida incorporou conceitos agronómicos, e modelos de simulação do desenvolvimento e crescimento de culturas, em métodos econométricos para a análise de fronteiras estocásticas. A aplicação destes métodos em cada caso de estudo permitiu quantificar a magnitude das quebras de produtividade, observadas para uma determinada cultura, devido a um ineficiente uso de inputs relativamente ao momento, distribuição espacial e forma dos inputs aplicados (*efficiency yield gap*), a uma quantidade insuficiente de inputs aplicados (*resource yield gap*) e, à falta de uso de tecnologias que permitam obter Yp (*technology yield gap*). Uma análise detalhada foi também conduzida ao nível da exploração agrícola com recurso a métodos estatísticos para estudar a relação entre, por um lado, os diferentes objectivos prosseguidos pelos agricultores, a disponibilidade limitada de recursos, as características estruturais da exploração agrícola e do perfil do agricultor e efeitos da

frequência de cultivo de diferentes culturas e, por outro, as quebras de produtividade observadas e as técnicas culturais adoptadas para cada cultura.

As principais causas das quebras de produtividade observadas para diferentes culturas e em diferentes regiões geográficas foram identificadas através de uma revisão bibliográfica de 50 artigos publicados (**Capítulo 2**). De um modo geral, as causas por detrás do *yield gap* variam entre culturas e entre regiões e são influenciadas pelo objectivo do estudo, conhecimento prévio acerca de possíveis causas e método de análise utilizado. Além disso, factores edáficos (por exemplo, a fertilidade do solo) e factores associados com a técnica cultural utilizada (e.g., fertilização) tendem a ser considerados com maior frequência para explicar quebras de produtividade, em detrimento de factores associados às características da exploração agrícola ou com as condições sócio-económicas envolventes. No que respeita à aplicação de nutrientes, as análises de *yield gap* tendem a incidir na quantidade de fertilizantes utilizados e não tanto no momento de aplicação, o qual quando considerado, explica o *yield gap* num grande número de casos. Do ponto de vista metodológico, esta revisão da literatura comprova que modelos de simulação do desenvolvimento de crescimento das culturas têm sido o método mais utilizado para estudar quebras de produtividade, apesar das suas limitações no tipo e diversidade de factores considerados.

Uma abordagem teórica foi desenvolvida, com base em métodos de análise de fronteira estocástica e em conceitos agronómicos de produção de plantas, para identificar as causas responsáveis por quebras de produtividade em diferentes explorações agrícolas (**Capítulo 3**). Dados históricos de diferentes produtores de arroz na região de Central Luzon (Filipinas) foram utilizados para testar esta abordagem. Entre 1979 e 2012, o *yield gap* da cultura de arroz foi estimado em 3.2 e 4.8 t ha^{-1} (ou cerca de 55% e 49% Yp) na estação das chuvas e na estação seca, respectivamente. Na estação das chuvas, quebras de produtividade na ordem dos 18% Yp deveram-se a um ineficiente uso de inputs, 13% Yp a uma quantidade de inputs sub-óptima e 15% Yp à falta de uso de tecnologias capazes de atingir os valores de Yp. Na estação seca, as quebras de produtividade deveram-se sobretudo à falta de uso de tecnologias capazes de atingir Yp (24% Yp). O uso ineficiente de inputs e uma quantidade insuficiente dos inputs usados contribuiram para quebras de produtividade na ordem dos 13% Yp. De um modo geral, observou-se uma relação entre o *efficiency yield gap* e o momento de aplicação de fertilizantes minerais e produtos fitofarmacêuticos; entre o *resource yield gap* e um défice de N, P e K aplicados em parcelas com baixas produtividades na estação seca, e; entre o *technology yield gap* e uma falta de uso de variedades com alto rendimento (estação húmida) assim como o uso insuficiente de água e nutrientes relativamente ao necessário para obter Yp (estação seca).

As causas responsáveis por quebras de produtividade na cultura de arroz em Central Luzon (Filipinas) identificadas no Capítulo 3 foram sujeitas a uma análise integrada ao nível da parcela, da exploração e do sistema de agricultura (**Capítulo 4**). Para este objectivo foram testadas as seguintes hipóteses: 1) existência de um compromisso entre reduzir o *yield gap* e maximizar a produtividade por unidade mão-de-obra, margem bruta e eficiência de uso do nutriente N e 2) as características estruturais da exploração agrícola e o perfil do agricultor, bem como outros factores socio-económicos, determinam as técnicas culturais adoptadas pelos agricultores nas suas parcelas. Duas bases de dados contendo informação detalhada para um elevado número de agricultores foram utilizadas para testar estas duas hipóteses. Durante os últimos cinquenta anos registou-se um aumento da produtividade de arroz quer por unidade de área (ha) quer por unidade de mão-de-obra (em horas, hr) bem como um aumento da importância da mão-de-obra contratada e da proporção de arroz vendido relativamente à quantidade de arroz produzida. Reduzir o *yield gap* neste sistema de produção compromete a maximização da produtividade por hora e a eficiência do uso de N, mas não a margem bruta por unidade de área (ha). As produtividades de arroz por hectare que maximizam a produtividade por hora e a eficiência do uso de N são, cerca de 25 - 35% inferiores a Yp, enquanto que produtividades de arroz por hectare que maximizam a margem bruta por hectare são cerca de 20% inferiores a Yp, quer na estação das chuvas quer na estação seca. Os resultados referentes à segunda hipótese foram inconclusivos o que sugere que as técnicas de cultivo utilizadas pelos agricultores parecem não ser constrangidas por factores ao nível da exploração agrícola e ao nível do sistema de agricultura.

A metodologia introduzida no Capítulo 3 foi utilizada para identificar as causas das quebras de produtividade nas principais culturas arvenses cultivadas na Holanda (**Capítulo 5**). Para além disso, informação àcerca da frequência de cultivo de cada cultura e dos diferentes objectivos prosseguidos pelos agricultores foi utilizada para estudar a sua relação com as quebras de produtividade observadas. A análise incidiu sobre uma base de dados contendo informação detalhada ao nível da cultura e da exploração agrícola para um elevado número de agricultores durante o período 2008 - 2012. Quebras de produtividade na ordem dos 25% e 40% Yp foram observadas para a cultura do trigo de inverno e da batata de indústria, respectivamente. De maneira geral, estas quebras de produtividade atribuem-se a um uso ineficiente de inputs no caso da batata de consumo e da cebola, a um uso ineficiente de inputs e à falta de uso de tecnologias capazes de atingir Yp no caso da beterraba e dos cereais, a uma quantidade insuficiente de inputs no caso da batata para semente e, a uma falta de uso de tecnologias capazes de atingir Yp no caso da batata de indústria. Os efeitos da rotação cultural na produtividade das diferentes culturas foram analisados e revelaram-se inconclusivos. Contudo, foram observadas diferenças significativas na produção de N ou energia, margem bruta, eficiência do uso de N ou mão-de-obra entre explorações agrícolas que procuram optimizar o seu desempenho nestes indicadores. Estas diferenças entre explorações com diferentes objectivos devem-se sobretudo a diferenças na área destinada a cada cultura, e não a diferenças na produtividade de cada cultura.

As causas das quebras de produtividade em sistemas de produção de milho e de trigo no Sul da Etiópia foram analisadas no **Capítulo 6**. Para tal, a metodologia introduzida no Capítulo 3 foi aplicada a uma base dados contendo informação detalhada sobre o uso de mão-de-obra ao nível da cultura e ao nível da exploração agrícola nos sistemas de agricultura de pequena escala em redor de Hawassa (milho) e Asella (trigo). A produtividade actual de milho e de trigo é cerca de 1.6 e 2.8 t ha^{-1}, valores muito inferiores aos da produtividade limitada por água (7.0 e 10.0 t ha^{-1}, adiante representada por Yw) simulada com recurso a modelos, respectivamente. Quebras de produtividade nestes sistemas de produção devem-se sobretudo à falta de uso de tecnologias capazes de atingir Yw, o que explica cerca de 45% Yw para o milho e 52% Yw para o trigo. O uso ineficiente de inputs explica cerca de 20% Yw para ambas as culturas e, foi associado à data de sementeira no caso da cultura do milho e à quantidade de mão-de-obra contratada no caso da cultura do trigo. Em Hawassa, as explorações com maior efectivo animal (bovinos de tração) aplicaram uma maior quantidade de N mineral à cultura do milho, com consequentes produtividades de milho mais elevadas, comparativamente às explorações com menor efectivo animal. Em Asella, as explorações com maior efectivo animal disponível para tração cultivaram uma área superior comparativamente a explorações com menor efectivo animal, não tendo sido encontradas diferenças entre estes dois grupos quer no uso de inputs quer na produtividade de trigo. Por fim, observou-se que as culturas cultivadas em Hawassa (milho, feijão e ensete) exibem um uso complementar de mão-de-obra ao longo do ano enquanto que as culturas predominantes em Asella (cereais e leguminosas) competem por mão-de-obra em determinados períodos da campanha agrícola. Tal sugere que deverão ter de ser estabelecidas prioridades para determinadas operações culturais para as diferentes culturas neste sistema de agricultura devido a limitações de recursos (nomeadamente mão-de-obra) ao nível da exploração.

A intensificação sustentável e o conceito de *yield gap* revelam alguma controvérsia em agronomia uma vez que nem sempre seguem definições consistentes e raramente consideram o contexto no qual os agricultores operam (**Capítulo 7**). Nos sistemas de agricultura de pequena escala nos trópicos, não basta compreender as causas das quebras de produtividade sendo imperativo conciliar essa análise com o estudo do risco e da variabilidade inerente à redução do *yield gap* e do impacto de possíveis ganhos de produtividade nos meios de subsistência locais. Um exemplo e reflexão sobre estes aspectos foi desenvolvido com base em dados empíricos de sistemas de produção da cultura do milho no Oeste do Quénia e da cultura do arroz em Central Luzon (Filipinas). Os resultados obtidos no Oeste do Quénia questionam a relevância de medições de produtividade em estações experimentais para estimar a produtividade actual e a produtividade limitada por água a nível regional. Para além disso, revelam as dificuldades em quantificar a magnitude exacta das quebras de produtividade quando estas são analisadas em termos de probabilidades acumuladas. Apesar das quebras de produtividade estimadas no Capítulo 3 para a cultura de arroz em Central Luzon (Filipinas) indicarem que a produção de arroz nesta região pode aumentar de modo considerável, parece haver poucos incentivos para a sua redução nas parcelas dos agricultores. Isto deve-se ao baixo retorno económico associado ao cultivo de arroz e à existência de outras oportunidades de emprego economicamente mais atractivas, e estáveis, comparativamente ao sector agrícola.

Contextualizar quebras de produtividade relativamente à realidade da exploração agrícola requer não só a aplicação de diferentes metodologias e a análise de diferentes tipos de dados, mas também um conhecimento do sistema de agricultura a ser estudado. Estes aspectos são fundamentais para compreender e explicar as quebras de produtividade observadas em diversos sistemas de produção actuais (**Capítulo 8**). Em resumo, as quebras de produtividade observadas para culturas arvenses na Holanda devem-se sobretudo a um ineficiente uso de inputs; para a cultura do milho e do trigo no Sul da Etiópia à falta de uso de tecnologias capazes de atingir a productividade limitada por água e, para a cultura do arroz em Central Luzon (Filipinas), devido aos factores previamente descritos assim como a uma quantidade sub-óptima de inputs aplicados (e.g., fertilizantes). Mitigar estas quebras de produtividade parece não compremeter a maximização da margem bruta por ha nos três casos de estudo mas requer perdas de produtividade por hora (ou margem bruta por hora) e de eficiência de uso de N. Por exemplo, as produtividades por hectare que maximizam a produtividade por hora ou a eficiência de uso de N são 30 - 40% inferiores à produtividade potential (ou a produtividade limitada por água no Sul da Etiópia). Esta diferença indica que continua a ser um desafio para os agricultores conciliar um aumento da produtividade por hectare com uma eficiência de uso de recursos elevada e com a adopção de tecnologias que reduzem o uso de mão-de-obra.

Três trajectórias podem ser indentificadas com base nos resultados desta tese a respeito do futuro dos sistemas de agricultura actuais: 1) intensificação sustentável (e.g., aumentos da produtividade actual face a reduções no uso de inputs), 2) extensificação (e.g., manutenção da produtividade actual face a reduções no uso de inputs) e 3) intensificação (e.g., aumentos da produtividade actual face a aumentos no uso de inputs). Neste contexto, uma intensificação nos sistemas de produção de milho em Hawassa e de trigo em Asella (Sul da Etiópia), e uma intensificação sustentável no sistema de produção de arroz em Central Luzon (Filipinas), é necessária para diminuir as quebras de produtividade e aumentar a eficiência do uso de recursos, enquanto que uma extensificação para ser mais adequada para os sistemas de produção de culturas arvenses na Holanda de modo a reduzir o uso de inputs sem compromoter a produtividade actual. Estas observações suportam o paradigma de que a agricultura deve ser intensificada em 'regiões em desenvolvimento' enquanto que um foco na melhoria da sustentabilidade é mais apropriado em 'regiões desenvolvidas'. Em ambos os casos, as reduções das quebras de produtividade e/ou melhorias na eficiência de uso dos recursos terão de ser alcançadas por agricultores individuais que são frequentemente forçados a sobrepôr necessidades a curto prazo sobre aspirações pessoais a longo prazo.

Acknowledgements

Five years have passed since I started working on my doctoral thesis back in December 2012. This has been a long journey and I am very grateful to have received the support and guidance from many colleagues, family and friends - you have made this journey not only possible but also much more enjoyable!

This thesis would not have been finalized without proper supervision and I could not have been more lucky to have Prof. Martin van Ittersum, Prof. Ken Giller and dr. Pytrik Reidsma as supervisors. Martin, you always pushed me to the 'next' level and I have no doubts the quality of this thesis would have been lower without your structured thinking and in-depth knowledge. I appreciate very much the guidance you have been giving me over the past years and your constant support and commitment to my work. Ken, your passion for farming systems research is contagious and it has been a source of inspiration for me throughout this thesis. One of the things I take from you is the importance of 'the context' in which (smallholder) agriculture takes place, and that we need to experience it in order to understand it. Pytrik, it has been a real pleasure to have you as a supervisor! You were always able to guide me, and keep me motivated, while giving me freedom to learn at my own pace and to explore my own ideas. I always felt your firm support, and open mindset, and deeply enjoyed the stimulating theoretical discussions we had along the way, and that laid the foundation of the framework(s) used in this thesis. Thank you very much!!

Another crucial aspect for the completion of this thesis was the interest and collaboration of numerous colleagues in the Philippines, the Netherlands and Ethiopia. I benefited enormously from you not only through the availability and provision of raw data but also through your commitment and interest in my work.

I was able to develop a case study on rice farming in the Philippines thanks to dr. Alice Laborte, who welcomed me as a scholar at the International Rice Research Institute (IRRI, Los Baños, the Philippines) for a period of six months. Alice, thank you for giving me this opportunity and supporting me over the past years as well as for the freedom you gave me to link up with other colleagues at IRRI. Special thanks to: 1) Ferdinand Corcuera for the help and great time we had together collecting soil samples across Central Luzon, 2) Ma Lourdes Velasco for preparing the MISTIG data used in Chapter 4 and 3) Piedade Moya (and colleagues) for making available the valuable Central Luzon Loop Survey. I would also like to thank dr. Tao Li and dr. Roland Buresh, and their respective teams, for the support provided on crop modelling and data sharing. A word of thanks also to dr. Alexander Stuart for the stimulating discussions we had, which resulted in a joint review paper on rice yield gaps. Finally, my gratitude to all the colleagues at the GIS lab for the welcoming atmosphere and to dr. Valerien Pede for making our football team nearly champion! Roman and Nikos: thanks for the great diving we had together in Mindoro and for the many dinners around Los Baños.

The case study on arable farming in the Netherlands was formalized in collaboration with Ruud van der Meer and dr. Hennie van de Veen, who provided me access to individual farm level data collected by the Wageningen Economic Research Institute (WER, the Hague, the Netherlands). Ruud, thank you for timely sharing the endless amount of figures I produced and my apologies for having bothered you so much with this. A word of thanks also to dr. Jaap Schröder and dr. Hein ten Berge for helping me to identify an important mistake in my analysis prior to the publication of Chapter 5. I believe that, without your scepticism, my poor understanding of Dutch could have put most arable farms in the Netherlands in quite some trouble! Thanks also to ir. Bert Rijk and to Jarno Rietema and Anton Bartelen for sharing with me their profound, and practical, knowledge on Dutch arable farming.

The IMAGINE project brought me to Ethiopia to coordinate a household survey *in situ*, which allowed me to meet and interact with many farmers. This was done in collaboration with dr. Kindie Tesfaye Fantaye from the International Maize and Wheat Improvement Center (CIMMYT, Addis Ababa, Ethiopia), whom I thank for hosting me for nearly two months and for helping with the associated logistics. I would also like to thank Fite Getaneh, and all enumerators, for the great time we had in the field and for hanging around with a *farangi* at times of political unrest. However, the case study on mixed crop-livestock farming materialized in collaboration with dr. Frédéric Baudron (CIMMYT), who challenged me to study labour dynamics across small farms in Southern Ethiopia. Fred, I appreciate your provocative thinking and attitude and thank you for your persistence in convincing me to conduct this analysis! I would like to express my gratitude to Tesfaye Sida for the friendship and long discussions we had both in Addis and in Wageningen and to Beyene Teklu for sharing his knowledge about smallholder farming in Southern Ethiopia.

I believe this thesis would not look the same without the friendly, supportive and 'sceptical' atmosphere of the Plant Production Systems Group (PPS). To the 'staff', special thanks to Maja, Gerrie, Peter, Katrien, Tom, Joost W, Joost vH, Andy, René, Sander, Bert, Marcel and Mink for the eye-openers, methodological help and/or sharing your wisdom with me. Ria, Charlotte, Linda and Alex-Jan - thanks the administrative support and for always helping out so smoothly. To the fellow PhD's, thanks for the sharing 'the struggle' as well as your enthusiasm and rich experiences from your work. Special thanks to Aart, Alejandro, Banchayehu, Bouba, Esther, Gatien, Godfrey, Guillaume, Jairos, Jannike, Jiska, Joy, Lotte K, Mary, Pablo, Sakti, Tiemen, Victoria, Wilson, Wim, Wytze and Yodit. Renske and Lotte W: I have to add to what have been said before that it has also been a real pleasure, and a lot of fun, to share an office with you - thanks for your support, great atmosphere and nice discussions! Also thanks to Raymond, and to colleagues at Agrosystems Research, making this possible. Um cumprimento especial ao Paulo Martinho e ao Tomás Tenreiro, a quem tive a possibilidade de orientar a tese de mestrado de maneira mais ou menos formal. Finally, I would also like to show my gratitude to two of my co-authors: Eskender Beza and Joshua Ramisch. Eskender, I learnt a lot with you in the long hours we spent together discussing technical yield gap papers and appreciate your friendship along these years very much! Joshua, thanks for sharing your scepticism on yield gaps and for nearly putting David Beckham in lingerie in the graphical abstract of our chapter!

Many friends and housemates avoided I adopted R as a mother tongue. Stefano and Heike, you have been under the same roof as me nearly everyday of this PhD and your passion for cooking, eating and drinking did as good to my stomach as to my liver! This was very often shared with Xabierto, Ruby, Jan, Simon, Costas, Dimitris, Edo and Fabian, just to highlight a small fraction of the friends we had on a regular basis. I also want to mention Braschi and Fekas for the nice moments in 107 and beyond. Cheers to everyone for the intellectual discussions / night outs / soccer table tournaments / etc. and for making my time in Wageningen much more interesting! Special thanks to Christoph and Rossella for helping designing the cover of this thesis and, most important, their long-lasting friendship. Um abraço especial ao Henrique pelas jantaradas portuguesas em Wageningen e ao Sérgio e Manel pelas noitadas que tivémos no ISA: sem dúvida que esta tese beneficou do que aprendi e discuti com vocês, bem como da vossa maneira de viver a agronomia. A word of thanks also to Mareike, Martin and Cathrin for always welcoming me in Bonn and for helping me so much over the past year - very much appreciated!

Não tenho como agredecer aos meus pais toda a orientação e todo o apoio que me deram e continuam a dar. Como tal, quero simbolicamente dedicar-lhes esta tese como sinal da minha apreciação pela formação que me proporcionaram e por me suportarem quer nos bons quer nos maus momentos. Espero poder estar mais presente em casa daqui para frente para poder ajudar nas tarefas agrícolas e aproveitar a vossa companhia. Um beijinho especial também para a avó Lúcia por me ter sempre acompanhado de perto. Susanne, não tenho palavras para descrever todo o apoio que me tens dado ao longo destes anos. Não tenho dúvidas de que o teu espírito de sacrifício e suporte emocional foram absolutamente essenciais para eu poder terminar este trabalho; Vielen Dank!! Por fim, uma palavra especial para a Martinha por me alienar de vez em quando do mundo das pessoas grandes e me trazer de volta ao admirável mundo das crianças. Muito obrigado!

João Vasco Silva
November 14, 2017

Curriculum Vitae

João Vasco Silva was born on the 8th of January 1989 in Lisbon, Portugal. Although most of his childhood was spent in this city, his numerous and prolonged stays in the village of Vale de Cavalos (Ribatejo), exposed him to the rural environment and to the importance of agriculture and rural development. After finishing high school in 2007, he decided to study agricultural sciences as a way to merge his interest in different disciplines and to learn more about food production. In 2010, he obtained a BSc degree in Agricultural Engineering from Instituto Superior de Agronomia, Universidade Técnica de Lisboa. In that same year he obtained an Erasmus Mundus scholarship for an MSc exchange of six months in Wageningen University (Netherlands), which turned out to be a much longer stay. In 2012, he obtained a MSc degree in Plant Sciences with a specialization in Natural Resource Management in this University. As part of his MSc thesis, he spent six months in Minas Gerais, Brazil, to conducted field work in a collaborative project between the Plant Production Systems Group (WUR) and Universidade Federal de Viçosa (Brazil). Since then, he has been working as a PhD candidate at the Plant Production Systems Group (WUR) in a project aiming to explain yield gaps at farm(ing) systems level with case studies in Central Luzon (Philippines), Southern Ethiopia and the Netherlands. This encompassed field activities in the different regions and active collaboration with the International Rice Research Institute (IRRI), the International Maize and Wheat Improvement Center (CIMMYT) and the Wageningen Economic Research Institute (WER). Aside from academia, he has practical knowledge on viticulture and wine making acquired mostly in the smallholder hobby farm of his family. He is currently employed as a researcher in a joint position between the Centre for Crop Systems Analysis (WUR) and the Plant Production Systems Group (WUR) aiming to estimate and explain water and nutrient use efficiencies across Europe and Northern Africa.

Peer-reviewed journal publications

- **Silva, J.V.**; Baudron, F.; Reidsma, P.; Giller, K.E. Is labour a major determinant of yield gaps in sub-Saharan Africa? A case study for cereals in Southern Ethiopia. To be submitted.
- **Silva, J.V.**; Reidsma, P.; Velasco, M.L.; Laborte, A.G.; van Ittersum, M.K. Intensification in rice-based farming systems in Central Luzon, Philippines: Constraints at field, farm and regional levels. *Agricultural Systems*. Under review.
- **Silva, J.V.**; Ramisch, J.J. Whose gap counts? The role of yield gap analysis within a development-oriented agronomy. *Experimental Agriculture*. Under review.
- **Silva, J.V.**; Reidsma, P.; van Ittersum, M.K. Yield gaps in Dutch arable farming: Analysis at crop and crop rotation levels. *Agricultural Systems*, 158, 78 – 92.
- **Silva, J.V.**; Reidsma, P.; Laborte, A.G.; van Ittersum, M.K. 2016. Explaining rice yield and yield gaps in Central Luzon, Philippines: An application of stochastic frontier analysis and crop modelling. *European Journal of Agronomy*, 82 Part B, 223 – 241.
- Beza, E.; **Silva, J.V.**; Kooistra, L.; Reidsma, P. 2016. Review of yield gap explaining factors and opportunities for alternative data collection approaches. *European Journal of Agronomy*, 82 Part B, 206 – 222.
- Stuart, A.M.; Pame, A.R.P.; **Silva, J.V.**; Dikitanan, R.C.; Rutsaert, P.; Malabayabas, A.J.G.; Lampayan, R.M.; Radanielson, A.M.; Singleton, G.R. 2016. Yield gaps in rice-based farming systems: Insights from local studies and prospects for future analysis. *Field Crops Research*, 194, 43 – 56.
- Leite, J.G.B.; **Silva, J.V.**; van Ittersum, M.K. 2014. Integrated assessment of biodiesel policies aimed at family farms in Brazil. *Agricultural Systems*, 131, 64 – 76.
- Leite, J.G.B.; Justino, F.B.; **Silva, J.V.**; Florin, M.J.; van Ittersum, M.K. 2014. Socioeconomic and environmental assessment of biodiesel crops on family farming systems in Brazil. *Agricultural Systems*, 133, 22 – 34.

- Leite, J.G.B.; **Silva, J.V.**; Justino, F.B.; van Ittersum, M.K. 2014. A crop model-based approach for sunflower yields. *Scientia Agricola*, 71, 345 – 355.

Conference proceedings

- Reidsma, P.; Rietema, J.; Yan, Y.; Kroes, J.; **Silva, J.V.** 2015. Potential yield and yield gap at farm level are different from the field level: A case study on a large Dutch potato farm. *5th International Symposium for Farming Systems Design*, Montpellier, France.

PE&RC PhD Education Certificate

With the training and education activities listed below the PhD candidate has complied with the requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)

The C.T. De Wit Graduate School PE&RC
PRODUCTION ECOLOGY & RESOURCE CONSERVATION

Review of literature (6 ECTS)

- Yield gaps: Magnitude and explanatory factors

Post-graduate courses (9 ECTS)

- Theory and practice of efficiency and productivity measurement; WUR, WASS (2013)
- Farming systems and rural livelihoods: vulnerability and adaptation; WUR, PERC (2013)
- Multivariate statistics; WUR, PERC (2013)
- Advanced applications with ORYZA; IRRI, CESD (2014)

Laboratory training and working visits (4 ECTS)

- Explaining rice yield gaps in Central Luzon, Philippines; International Rice Research Institute (IRRI), Los Baños, the Philippines (2014)
- Integrated assessment of maize yield gaps in Ethiopia; International Maize and Wheat Improvement Center (CIMMYT), Addis Ababa, Ethiopia (2016)

Invited review of (unpublished) journal manuscript (5 ECTS)

- Field Crops Research: closing maize yield gaps in SE Asia (2013)
- Environment, Development and Sustainability: sustainability of rice production systems (2014)

- Field Crops Research: yield gaps at cropping system level (2016)
- Agricultural Systems: rice yield development in China (2016)
- European Journal of Agronomy: effects of farm management on barley performance (2017)

Competence strengthening / skills courses (1.8 ECTS)

- PhD Competence assessment; WUR and Maas Organization & Career Development (2013)
- Ethics and philosophy of life sciences; WUR, WIAS (2014)

PE&RC Annual meetings, seminars and PE&RC weekend (2.5 ECTS)

- WPC Symposium; Wageningen, the Netherlands (2013)
- PERC Weekend, first year (2014)
- PERC Day: one's waste... another's treasure? (2015)

Discussion groups / local seminars / other scientific meetings (7.5 ECTS)

- Sustainable intensification of agricultural systems (2014-2017)
- R users' meeting (2014-2017)

International symposia, workshops and conferences (6.6 ECTS)

- International Rice Congress; Bangkok, Thailand (2014)
- Farming systems design; Montpellier, France (2015)
- Contested Agronomy; Brighton, UK (2016)

Lecturing / supervision of practical's / tutorials (1.2 ECTS)

- Integrated natural resource management in organic agriculture (2015)
- Global food security (2016)

Supervision of MSc students (6 ECTS)

- Explaining yield gaps on a Dutch potato farm: Case study at van den Borne Aardappelen
- Analysis of the impact of seed quality and seed management practices on the yield and quality of ware potatoes

This research was carried within 'Mapping for Sustainable Intensification', a strategic programme funded by Wageningen University & Research and co-funded by the C.T. de Wit Graduate School of Production Ecology & Resource Conservation (PE&RC).

Back cover photos by: Bert Rijk, João Vasco Silva and Frédéric Baudron

Cover design by Christoph Reimers (www.superservice-international.com)