RESEARCH ARTICLE

Wheat yield gaps across smallholder farming systems in Ethiopia

João Vasco Silva[1,2] · Pytrik Reidsma[2] · Frédéric Baudron[3] · Moti Jaleta[4] · Kindie Tesfaye[4] · Martin K. van Ittersum[2]

© The Author(s) 2021

Abstract

Wheat yields in Ethiopia need to increase considerably to reduce import dependency and keep up with the expected increase in population and dietary changes. Despite the yield progress observed in recent years, wheat yield gaps remain large. Here, we decompose wheat yield gaps in Ethiopia into efficiency, resource, and technology yield gaps and relate those yield gaps to broader farm(ing) systems aspects. To do so, stochastic frontier analysis was applied to a nationally representative panel dataset covering the Meher seasons of 2009 and 2013 and crop modelling was used to simulate the water-limited yield (Yw) in the same years. Farming systems analysis was conducted to describe crop area shares and the availability of land, labour, and capital in contrasting administrative zones. Wheat yield in farmers' fields averaged 1.9 t ha^{-1} corresponding to ca. 20% of Yw. Most of the yield gap was attributed to the technology yield gap (> 50% of Yw) but narrowing efficiency (ca. 10% of Yw) and resource yield gaps (ca. 15% of Yw) with current technologies can nearly double actual yields and contribute to achieve wheat self-sufficiency in Ethiopia. There were small differences in the relative contribution of the intermediate yield gaps to the overall yield gap across agro-ecological zones, administrative zones, and farming systems. At farm level, oxen ownership was positively associated with the wheat cultivated area in zones with relatively large cultivated areas per household (West Arsi and North Showa) while no relationship was found between oxen ownership and the amount of inputs used per hectare of wheat in the zones studied. This is the first thorough yield gap decomposition for wheat in Ethiopia and our results suggest government policies aiming to increase wheat production should prioritise accessibility and affordability of inputs and dissemination of technologies that allow for precise use of these inputs.

Keywords Wheat self-sufficiency · Sustainable intensification · Farmer field data · Stochastic frontier analysis · Crop modelling

1 Introduction

Ethiopia is the largest wheat producer in sub-Saharan Africa with a record harvest of 4.6 million metric tons registered in 2017 (CSA 2019). However, during that same year, the country imported 1.5 million tons of wheat, corresponding to a value of around US$600 million (CSA 2019). Further

increases in demand for wheat (and other cereals), due to population growth and dietary changes (van Ittersum et al. 2016), are projected to put additional pressure on the national treasury, making the national economy vulnerable to cereal price volatility in world markets. These drivers have put wheat self-sufficiency high on the agenda, with a new initiative of the Ethiopian government targeting self-sufficiency in wheat in the coming few years (https://www.press.et/english/?p=815).

Increasing wheat yield in Ethiopia, through narrowing yield gaps, is important to reduce the import dependency for this crop while avoiding area expansion. This needs to occur in a smallholder agriculture setting as wheat is cultivated by approximately 4.2 million smallholders on ca. 1.7 million ha (Fig. 1, CSA 2019). Currently, wheat is produced mostly under rainfed conditions and with relatively low inputs (Anteneh and Asrat 2020). Despite the yield growth observed during the past 15 years (ca. 63 kg ha^{-1} year^{-1}), with wheat yield doubling to values reaching ca. 2.7 t/ha

✉ João Vasco Silva
joao.silva@wur.nl; jvasco323@posteo.net

1 Centre for Crop Systems Analysis, Wageningen University, Wageningen, The Netherlands

2 Plant Production Systems, Wageningen University, Wageningen, The Netherlands

3 Sustainable Intensification Program, CIMMYT-Zimbabwe, Harare, Zimbabwe

4 Socioeconomics program, CIMMYT-Ethiopia, Addis Ababa, Ethiopia

INRAE ⌂ Springer

Fig. 1 Wheat fields and landscapes in Ethiopia. The crop is cultivated in Ethiopia by ca. 4.2 million smallholders and covers an area of ca. 1.7 million ha. Photo credits: F. Baudron (left) and M. Jaleta (right)

(CSA, 2018), the current wheat yield is only ca. 20% of its water-limited potential (Silva et al. 2019a; van Ittersum et al. 2016). Understanding the key drivers behind this large yield gap is thus important to help prioritise policies and interventions towards wheat self-sufficiency in Ethiopia.

Yield gap analysis at regional level is useful to investigate the key factors limiting and reducing crop production in farmers' fields (Hochman and Horan 2018; Rattalino-Edreira et al. 2017; van Ittersum et al. 2013a). Previous research concluded wheat yield gaps in West Arsi, one of the wheat belts in Ethiopia, can be largely attributed to technology yield gaps (Silva et al. 2019a). This means that technologies currently used by farmers do not reach agronomic best practices and that considerably more and better use of inputs is needed if wheat yield gaps are to be narrowed (Habte et al. 2014; Tanner et al. 1993). Competition for labour during sowing, weeding, and harvesting was also observed in this administrative zone, as labour peaks for other cereal and legume crops overlap with labour peaks for wheat. This results in potential trade-offs on resource allocation at farm level and reflects the importance of contextualizing yield gaps within broader farming system aspects (Silva and Ramisch 2019b).

The objective of this manuscript is twofold: (1) to decompose wheat yield gaps across smallholder farming systems in Ethiopia in order to identify relevant management, technological, and policy interventions corresponding to efficiency, resource, and technology yield gaps, and (2) to relate those yield gaps to broader farm(ing) systems contexts. For this purpose, we applied a framework combining frontier analysis and crop modelling to a large and nationally representative database of individual farm field data. This is the first agronomic assessment of the causes of wheat yield gaps across the main wheat producing administrative zones,

agro-ecological zones, and farming systems of Ethiopia and we suggest our findings are useful for policy makers aiming to improve food security and to ensure wheat self-sufficiency in the country.

2 Materials and methods

2.1 Concepts and definitions

The yield gap analysis presented here builds upon the frameworks of Silva et al. (2017) and van Dijk et al. (2017), which consider six yield levels to decompose the yield gap (Table 1). The water-limited yield (Yw) refers to the maximum yield that can be obtained under rainfed conditions in a well-defined biophysical environment (van Ittersum et al. 2013a). Yw can be simulated with crop growth models or derived from field trials with high levels of nutrients applied and pests, diseases, and weeds fully controlled. The highest farmers' yields (Y_{HF}) refer to the maximum yields (e.g. average above the 90th percentile of actual farmers' yields) observed in a sample of farmers sharing similar biophysical conditions (weather and soils) and technologies adopted (e.g. varieties). Differently from Y_{HF}, van Dijk et al. (2017) consider economic yields (Ye) and feasible yields (Yf): the former refers to the yield level in which marginal costs are equal to marginal revenue and the latter refers to the maximum yield that can be reached with available technology and best-practice management but without economic constraints. The technically efficient yields (Y_{TEx}) comprise the maximum yield that can be achieved for a given input level and can be computed using methods of frontier analysis in combination with concepts of production ecology (Silva et al. 2017). The latter provide

Table 1 Concepts and definitions of the yield levels and yield gaps estimated in this manuscript

Notation	Concept	Definition and reference
Yield levels		
Y_a	Actual yield	Yield observed in farmers' fields as a result of growth-limiting and -reducing factors (van Ittersum et al. 2013a);
Y_{TEx}	Technical efficient yield	Maximum yield that can be achieved for a given input level (Silva et al. 2017);
Y_{HF}	Highest farmers' yield	Average yield across top 10th percentile of actual yields (van Ittersum et al. 2013a);
Y_e	Economic yield	Maximum yield with available technology and best-practice management in which marginal costs equal marginal revenue (van Dijk et al. 2017);
Y_f	Feasible yield	Maximum yield with available technology and best-practice management but no economic constraints (van Dijk et al. 2017);
Y_w	Water-limited yield	Maximum yield of a given genotype when crop growth is limited by water supply (van Ittersum et al. 2013a).
Yield gaps		
Yg	Total yield gap	Difference between water-limited yield (Y_w) and actual yield (Y_a);
Eff. Yg	Efficiency yield gap	Difference between technical efficient yield (Y_{TEx}) and actual yield (Y_a);
Res. Yg	Resource yield gap	Difference between highest farmers' (Y_{HF}) or feasible yield (Y_f) and technical efficient yield (Y_{TEx});
Alloc. Yg	Allocative yield gap	Difference between economic yield (Y_e) and technical efficient yield (Y_{TEx});
Econ. Yg	Economic yield gap	Difference between feasible yield (Y_f) and economic yield (Y_e);
Tech. Yg	Technology yield gap	Difference between water-limited yield (Y_w) and highest farmers' (Y_{HF}) or feasible yield (Y_f).

We note these yield levels and yield gaps are only meaningful, and therefore also defined, for a given genotype × environment combination

a framework to assess the relative importance of different growth factors and inputs to actual yields and resource-use efficiencies (van Ittersum and Rabbinge 1997). Finally, the actual yield (Y_a) refers to the yield in farmers' fields as recorded in farm surveys.

Six yield gaps can be distinguished based on these yield levels (Table 1). The total yield gap is defined as the difference between Y_w and Y_a. The efficiency yield gap is defined as the difference between Y_{TEx} and Y_a and it is explained by crop management imperfections related to time, form, and/or space of the inputs applied. The resource yield gap is defined as the difference between Y_{HF} and Y_{TEx} and captures the yield penalty due to a sub-optimal amount of inputs applied. According to van Dijk et al. (2017), the resource yield gap can be further decomposed into an allocative yield gap ($Y_e - Y_{TEx}$) and into an economic yield gap ($Y_f - Y_e$), which allows exploring how profit-maximisation affects the amount of inputs used. These allocative and economic yield gaps are particularly important in developing countries due to market failures and lower amounts of inputs used in highest-yielding fields than those required to reach Y_w (van Ittersum et al. 2013a). The technology yield gap is defined as the difference between Y_w and Y_{HF} (Silva et al. 2017) or between Y_w and Y_f (van Dijk et al. 2017), which can be caused by resource yield gaps of specific inputs and/or the use of technologies in farmers' fields where Y_w is not achieved.

2.2 Farm household survey

The Wheat Adoption and Impact Survey (WAIS) was conducted by the International Maize and Wheat Improvement Center (CIMMYT), in collaboration with the Ethiopian Institute of Agricultural Research (EIAR), for the purpose of tracking varietal change and assessing the impact of genetic improvement of wheat in Ethiopia (Jaleta et al. 2019; Tolemariam et al. 2018). The survey is a panel at household level covering the growing seasons of 2009 and 2013 and it is representative of Ethiopian wheat growing areas. The sampling frame comprised the selection of 148 major wheat growing districts, followed by a random selection of 120 farmers' associations (communities) within these districts and by a random selection of 15 to 18 households within each farmers' association (Tolemariam et al. 2018; Abro et al. 2017). This resulted in a sample of 2096 representative farmers (Fig. 2).

The survey included a wide range of farm and household characteristics as well as detailed information on the types and quantities of inputs used and crop yields obtained in all fields of each farm (Figs. S1 and S2), which makes it suitable for yield gap analysis (Beza et al. 2017). Descriptive statistics of the data are provided in Table 2 for the administrative zones with more than 100 wheat plots in both survey rounds. The large sample size and national coverage make this survey suitable for analysis
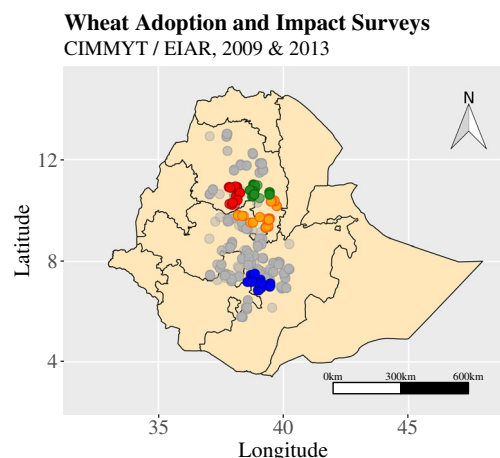
INRAE Springer

Wheat Adoption and Impact Surveys
CIMMYT / EIAR, 2009 & 2013

Fig. 2 Location of the households surveyed during the Meher seasons of 2009 and 2013. Regional analyses were conducted for the administrative zones of East Gojam (red), South Wollo (green), North Showa (orange), and West Arsi (blue). Grey dots show the location of the additional households included in the national level analysis

at national level and allows for more focused analyses across different administrative zones. The latter was done for West Arsi, North Showa, East Gojam, and South Wollo administrative zones given the large sample size for each zone and the similar agro-ecological conditions and farming systems across zones (Table 2). Yet, these zones differ in a number of socio-economic conditions such as other crops cultivated and land availability.

All data in the WAIS were self-reported by the farmer at the end of the growing seasons of 2009 and 2013. Actual yields were calculated based on farmers' report on wheat production after threshing and plot area. Wheat production after threshing was assumed to have a dry-matter content of 86.5%. The amount of urea applied was not properly recorded in Arsi, West Arsi, and Jimma administrative zones. For this reason, these data were imputed based on the amount of di-ammonium phosphate (DAP) applied, for which reliable data were available, while assuming a linear relationship between both. We also note that plots with 0 kg N ha^{-1} applied ($n = 1043$) were not considered in the yield gap analysis as we were not sure whether this value refers to no application or to missing data. We excluded wheat plots belonging to households with more than 20 pairs of oxen ($n = 2$) and plots with wheat yields greater than 10 t DM ha^{-1} ($n = 11$) and unspecified soil fertility status ($n = 1$). This resulted in 3818 (bread) wheat plots during the Meher growing seasons of 2009 ($n = 1751$) and 2013 ($n = 2067$).

Wheat yields and yield gaps were averaged per administrative zone, agro-ecological zone, or farming system (as per Table 2). The administrative zones were retrieved from the household survey while the other classifications were obtained from secondary sources based on the GPS coordinates of the individual households.

The agro-ecological classification combines temperature, rainfall, elevation, and the length of the growing season for the main crops and was retrieved from the Ministry of Agriculture of Ethiopia (MoA 1998). The farming system classification combines agro-ecological information with expert knowledge of the main farming systems in Ethiopia as documented by Amede et al. (2017).

2.3 Yield gap analysis at field level

2.3.1 Stochastic frontier analysis

Stochastic frontier analysis was used to estimate the production frontier, Y_{TEx}, and the efficiency yield gap for wheat production in Ethiopia. The estimated models assumed a Cobb-Douglas functional form (i.e. only first-order terms included) to describe the relationship between wheat yield and a vector of biophysical and agronomic variables. Models with a translog functional form were also fitted to assess interactions between variables but these are presented as Supplementary Material only (Table S1). We note parameter estimates were rather similar between both functional forms and the Cobb-Douglas has a considerable lower number of parameters and is easier to interpret. The formulation of the stochastic frontier model with a Cobb-Douglas functional form and the calculation of Y_{TEx} and the efficiency yield gap (Eff. Yg) were as follows (Silva et al. 2017; Battese and Coelli 1995):

$$ln\ y_{it} = \alpha_0 + \sum_{k}^{K} \beta_k\ ln\ x_{kit} + v_{it} - u_{it} \qquad (1)$$

$$v_{it} \sim N(0, \sigma_v^2) \qquad (2)$$

$$u_{it} \sim N^+(\mu, \sigma_u^2) \qquad (3)$$

$$\text{Eff. Yg}_{it} = 1 - exp(-u_{it}) \qquad (4)$$

$$Y_{TEx_{it}} = y_{it} \times exp(-u_{it})^{-1} \qquad (5)$$

where y_{it} represents the wheat dry-matter (DM) grain yield reported in plot i and in year t, x_{kit} is a vector of agronomic inputs k used on plot i and year t, and α_0 and β_k are parameters to be estimated. The stochastic frontier accounts for two random errors, v_{it} (random noise) and u_{it} (technical inefficiency), which are assumed to be independently distributed from each other and to follow a normal (2) and half-normal distribution truncated at 0 (3), respectively (Battese and Coelli 1995). The Cobb-Douglas model fitted to the pooled data was used to estimate the feasible yield (Yf) for different input levels as explained in Section 2.3.3.

Table 2 Number of households per administrative zone, agro-ecological zone, and farming system and mean values of the variables related to biophysical conditions, yield, crop management, and farm characteristics for wheat production in the Meher seasons of 2009 and 2013 across Ethiopia

	West Arsi	Arsi	North Showa	West Showa	East Showa	Jimma	East Gojam	North Gonder	South Gonder	North Wollo	South Wollo	Hadiya	Gurage	Others
Agro-ecology														
Humid highlands (H2)	158	174	0	0	1	59	0	0	0	0	0	0	0	0
Humid sub-afroalpine (H3)	11	55	0	0	0	0	0	0	0	0	0	0	0	0
Sub-humid lowlands (SH1)	0	0	0	0	0	0	0	0	0	0	0	97	0	0
Sub-humid highlands (SH2)	252	26	1	0	89	12	0	0	0	0	0	48	94	67
Moist highlands (M2)	4	176	214	176	0	0	228	22	21	0	107	0	23	184
Moist sub-afroalpine (M3)	0	0	18	0	0	0	23	3	72	0	0	0	0	8
Sub-moist highlands (SM2)	83	67	294	0	127	0	0	0	0	40	367	0	0	0
Sub-moist sub-afroalpine (SM3)	0	0	0	0	0	0	0	0	0	0	95	0	0	0
Other	44	151	0	0	1	0	0	0	0	8	0	0	0	64
Farming system														
Highland mixed	424	580	527	176	133	3	251	25	93	48	569	145	117	323
Highland perennial	31	0	0	0	0	68	0	0	0	0	0	0	0	0
Maize mixed	97	69	0	0	85	0	0	0	0	0	0	0	0	0
Biophysical conditions														
Altitude ($\times$ 1000 m)	2.38	2.53	2.73	2.62	1.96	2.3	2.54	2.62	2.9	2.99	2.7	2.1	2.24	2.37
Growing degrees days ($\times$100 °C)	55.0	55.7	55.9	59.6	67.3	61.1	57.6	67.0	53.2	57.4	56.6	59.5	61.9	59.8
Aridity index ($\times$100 mm mm^{-1})	62.9	72.5	69.5	71.2	54.6	101.3	86.3	67.2	90.7	61.9	74.9	71.5	67.0	79.4
Wheat yields														
Ya (t DM ha^{-1})	2.3	2.1	1.7	1.5	1.9	1.6	1.8	1.2	1.5	1.5	1.2	2.0	1.8	1.5
Y_{HF} (t DM ha^{-1})	4.5	4.0	4.2	3.2	3.9	3.6	3.1	2.4	2.7	2.8	3.4	4.4	3.0	3.7
Y_w (t DM ha^{-1})	10.3	9.2	9.3	8.7	8.3	9.9	8.5	9.3	8.3	8.7	8.4	9.1	10.6	8.9
Crop management														
Seed rate (kg ha^{-1})	211.3	216.3	198.2	168.5	222.9	156.1	195.4	166.8	198.3	171.0	163.8	195.6	160.4	188.4
Nitrogen use (kg N ha^{-1})	41.3	39.6	63.0	49.5	53.7	46.4	62.5	60.6	30.6	43.6	48.8	65.3	42.4	40.8
Herbicide use (L ha^{-1})	0.9	0.8	0.3	0.8	0.5	0.9	0.2	0.0	0.3	0.3	0.0	1.0	0.9	0.8
Hand-weeding (persday ha^{-1})	13.0	12.5	31.7	18.9	16.9	22.9	31.6	54.9	51.1	80.9	57.1	9.6	11.8	13.7
Labour use (persday ha^{-1})	41.2	48.7	80.6	60.7	59.1	79.4	79.1	96.4	90.0	120.6	94.2	59.8	50.6	59.5
Farm characteristics														
Field size (ha)	0.6	0.4	0.4	0.5	0.5	0.3	0.3	0.3	0.3	0.2	0.3	0.4	0.4	0.5
Pairs of oxen (#)	2.4	2.5	1.9	3.7	2.4	1.9	1.6	0.7	1.1	0.8	1.2	1.4	1.9	2.1

Data are presented for administrative zones with more than 100 observations in both survey rounds

The vector of inputs x_{kit} was defined according to principles of production ecology (van Ittersum and Rabbinge 1997). The growth-defining factors included in the analysis were temperature seasonality and growing degrees days (both obtained from the climate zonation of Van Wart et al. 2013b), year of the survey (Meher seasons of 2009 and 2013), seed rate, and type of variety (improved or unknown landrace). Temperature seasonality refers to the standard deviation of monthly average temperatures and growing degrees days consider a base temperature of 0 °C. The growth-limiting factors related to water included in the analysis were aridity index (i.e. the ratio between annual total precipitation and annual total potential evapotranspiration, also from Van Wart et al. 2013b), soil available water (obtained from the Africa Soil Information Service, AfSIS) and farmer reported information on soil depth (deep, medium, shallow), occurrence of water logging (yes/no), occurrence of drought (yes/no), use of water conservation techniques (yes/no), and ploughing frequency (less than three times, three times, four times, and five times or more). The growth-limiting factors related to nutrients included in the analysis were the farmer reported soil fertility status of the plot (rich, medium, poor), the use of manure (yes/no), incorporation of crop residues (yes/no), previous crop type (cereal, legume, other), and N applied (kg N ha^{-1}). P applied was not included due to strong collinearity with N applied. Finally, herbicide use (L ha^{-1}), hand-weeding (person-day ha^{-1}), a dummy variable to distinguish weeded from non-weeded plots, pesticide use (yes/no), and occurrence of pests or diseases (yes/no) were included to capture or control for growth-reducing factors. Missing data on seed rate ($n = 828$ plots) were filled with the mean value of the pooled sample and fields with no N applied were excluded from the analysis. All continuous input-output variables were ln-transformed prior to the analysis, so that parameter estimates can be interpreted as elasticities.

The stochastic frontier model (1)–(3) was fitted to the pooled sample (national analysis) and to subsets of the data for selected administrative zones (namely West Arsi, North Showa, East Gojam, and South Wollo) using maximum likelihood, as implemented in the *sfa()* function of the R package *frontier* (Coelli and Henningsen 2013). Efficiency yield gaps (4) and Y_{TEx} (5) were derived from the stochastic frontier model fitted to the pooled sample. We tried to assess the determinants of the efficiency yield gap with a second-stage regression (Battese and Coelli 1995) but refrain from showing these because most models did not converge and results were inconclusive. Data were used as a cross-section rather than as a panel of households in all models estimated meaning that the effects of technological change and time-(in)variant technical inefficiencies were not tested.

2.3.2 Input use across actual yield percentiles

Farmers' fields within a unique year × climate zone × soil fertility combination were categorised into highest-, average- and lowest-yielding based on their actual yields. Year, climate zone, and soil fertility were considered in this analysis to ensure yield differences between the three field types were only explained by differences in crop management and not in biophysical conditions. Year refers to the Meher seasons of 2009 and 2013, the climate zones were obtained from the Global Yield Gap Atlas (Van Wart et al. 2013b), and the soil fertility was based on farmer's own assessment. Varieties were not considered because we found no significant yield differences between variety types (Table 3).

Highest-yielding fields were identified as the observations above the 90th percentile of Ya and the highest-farmers' yields (Y_{HF}) were computed as the mean Ya for these fields. Similarly, the lowest-yielding fields were identified as the observations below the 10th percentile of Ya (Y_{LF}), and the average-yielding fields as the observations between the 10th and the 90th percentile of Ya (Y_{AF}). Significant differences across the different field categories were tested with analysis of variance (ANOVA) followed by a Tukey HSD post hoc test (considering a 5% significance level) for wheat yield, seed rate, N applied, total labour use (for land preparation, sowing, hand-weeding, and harvesting), labour use for hand-weeding, and herbicide application. This was implemented for selected administrative zones using the *scipy* and *statsmodels* libraries in Python (Virtanen et al. 2020; Seabold and Perktold 2010).

2.3.3 Crop modelling and variety trials

Water-limited yields of wheat across Ethiopia were simulated with the WOFOST crop model (Boogaard et al. 2014) for the Meher seasons of 2009 and 2013 following the protocols of the Global Yield Gap Atlas (GYGA; Grassini et al. 2015; van Bussel et al. 2015). These provide a bottom-up approach to estimate Yw (with a dry-matter content of 86.5%) within a spatial framework using climate zones based on local weather, soil, and agronomic data. Daily weather data on minimum and maximum temperatures and precipitation for 12 weather stations across the country were acquired from the National Meteorology Agency of Ethiopia. Gridded soil data on rootable depth and soil water availability were obtained from AfSIS, and crop management information was obtained through expert knowledge and literature review. Further details about model calibration and validation can be found in Tesfaye (2016). The simulated Yw was linked to the household survey based on the GPS coordinates of each farm.

The technology yield gap was calculated as the difference between Yw and Y_{HF} for unique year × climate zone × soil fertility combinations (Section 2.3.2.). The simulated Yw was further compared with wheat yields observed in variety trials conducted in 2016 and 2017 in Debre Zeit, Kulumsa, Bekoji, and Dawa Busa (Bezabih et al. 2018). This comparison was done for Arsi (Kulumsa) only with the purpose to cross-validate the simulated Yw and assess the contribution of varietal differences in yield potential to the technology yield gap. Moreover, the feasible yield was estimated with the Cobb-Douglas stochastic frontier model fitted to the pooled data (1)–(3) for non-limiting amounts of inputs applied as proposed by van Dijk et al. (2017). In other words, the parameter estimates of the fitted stochastic frontier model were used to predict wheat yields at high input levels. This was done for three different levels of N applied (150, 250, and 350 kg N ha^{-1}) in combination with a seed rate of 200 kg ha^{-1}, 1.5 L ha^{-1} of herbicide use, and 25 person-day ha^{-1} for hand-weeding labour. These seed and N rates are in line with the amounts needed to reach Yw (www.yieldgap.org) while weeding requirements reflect current management in highest-yielding fields of West Arsi. The contribution of sub-optimal amounts of inputs to the technology yield gap was further assessed by estimating an additional resource yield gap, i.e. the difference between the estimated feasible yields and Y_{HF}. These resource yield gaps are thus part of the technology yield gap defined as the difference between Yw and Y_{HF}.

2.4 Resource allocation at farm level

Wheat cultivation by smallholders in Ethiopia occurs alongside the cultivation of other crops. This has important implications for the allocation of resources at the farm level and may lead to trade-offs depending on the level of resource constraints. Crop area shares of wheat, other cereals (e.g. barley and tef), pulses (e.g. faba bean, field peas and chickpeas), oilcrops, and vegetables were computed to assess the level of specialisation in wheat production of individual households. Resource allocation at farm level was further studied by comparing the number of ploughing days, labour use for weeding, and total labour use (incl. land preparation, sowing, hand-weeding, and harvesting) for wheat and for other crops within each household. We found no evident substitution or competition effects and results are thus presented as Supplementary Material (Fig. S6). It was not possible to relate the amount of labour used with the timing of the different management operations as the dataset lacked information on the latter.

The pairs of oxen owned by each household were used to investigate whether the farming system was limited by land (intensification pathway) or by labour (extensification pathway) in relative terms (Silva et al. 2019a). Four

different groups were identified based on this information: households with no oxen, households with one pair of oxen, households with two pairs of oxen, and households with three or more pairs of oxen. Differences in wheat yield, resource availability, and input use between households owning different pairs of oxen were tested for significance in selected administrative zones with ANOVA followed by a Tukey HSD post hoc test, as explained in Section 2.3.2. Other resource variables analysed included wheat cultivated area, total labour use for land preparation, sowing, hand-weeding and harvesting, farm assets as reported by each household (excluding livestock), seed rate, N fertiliser rate, herbicide use, and labour use for hand-weeding (all referring to wheat production).

3 Results and discussion

3.1 Magnitude of wheat yield gaps in Ethiopia

Wheat yields in farmers' fields were on average 1.9 t ha^{-1} for the pooled sample, which corresponds to a yield gap closure of 21% of Yw (Fig. 3). The lowest Ya (less than 1.7 t ha^{-1}) was recorded in the moist and sub-moist agro-ecological zones (M2, M3, and SM3) while the highest Ya (1.8–2.2 t ha^{-1}) was recorded in the humid (H2 and H3) and sub-humid agro-ecological zones (SH1 and SH2; Fig. 3A and D). The latter are indeed amongst the most suitable agro-ecological zones for wheat production in Eastern Africa (Negassa et al. 2013; Hodson and White 2007). Ya also varied across administrative zones with a minimum of 1.2 t ha^{-1} in South Wollo and North Gonder and a maximum of 2.9 t ha^{-1} in West Arsi (Fig. 3B and E). No major differences in Ya were observed between highland mixed and highland perennial farming systems (Fig. 3C and F) while wheat yields were considerably greater in maize mixed farming systems, ca. 3.2 t ha^{-1}, than in the other farming systems. However, we note the bulk of the sample is classified as highland mixed farming system (Table 2). Finally, Yw varied between 8.3 t ha^{-1} in South Gonder (moist sub-afroalpine areas) and in East Showa (sub-moist and sub-humid highlands) up to 10.5 t ha^{-1} in West Arsi (humid and sub-humid highlands) and Gurage (sub-humid highlands; Fig. 3A and B).

The actual yield reported in the household survey at national level was 2.1 and 1.7 t ha^{-1} in 2009 and 2013, respectively. We acknowledge yield progress may have occurred since then but our analysis is still relevant as it focuses on comparisons between farms and regions. Yet, the mean actual yield in 2009 was slightly greater than that reported in the official statistics in the same year while the opposite was true in 2013 (CSA, 2018). Official statistics indicate actual yields of wheat in Ethiopia of 1.8 and 2.5 t
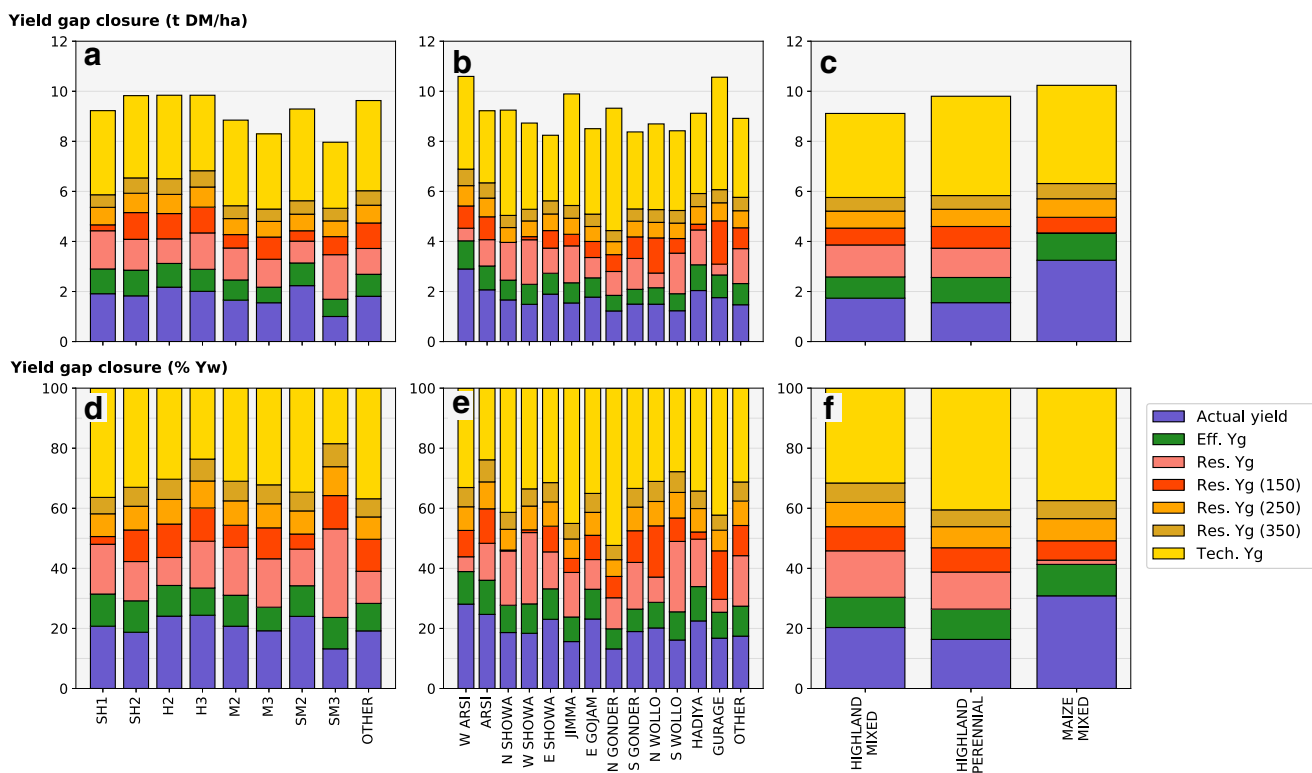
Yield gap closure (t DM/ha)

Fig. 3 Wheat yields and yield gaps in Ethiopia disaggregated by (A, D) agro-ecological zone, (B, E) administrative zone, and (C, F) farming system. Panels in the top row show data in absolute terms (t DM ha^{-1}) and panels in the bottom row show data in relative terms (% of Yw). Codes: SH1 = 'sub-humid lowlands', SH2 = 'sub-humid highlands', H2 = 'humid highlands', H3 = 'humid sub-afroalpine', M2 = 'moist highlands', M3 = 'moist sub-afroalpine', SM2 = 'sub-moist highlands', SM3 = 'sub-moist sub-afroalpine', 'Eff. Yg' = efficiency yield gap, 'Res. Yg' = resource yield gap, 'Res. Yg (150, 250, 350)' = resource yield gap assuming 150, 250, or 350 kg N ha^{-1}, 'Tech. Yg' = technology yield gap

ha^{-1} for the years 2009 and 2013, respectively, and clearly show wheat yield progress since the early 2000s (FAOSTAT, 2019). Actual yields for West Arsi in the household survey analysed here (i.e. WIAS) were on average 3.4 and 2.2 t ha^{-1} in 2009 and 2013, respectively, values which are in line with the 2.7 t ha^{-1} reported in Silva et al. (2019a) for a different household survey conducted in the same zone in 2012.

Wheat yield gaps were mostly attributed to the technology yield gap (> 50% of Yw) but narrowing efficiency and resource yield gaps can still double Ya (Fig. 3). This was true for most agro-ecological zones, administrative zones, or farming systems. The efficiency yield gap was on average 10% of Yw and did not differ much between agro-ecological zones (7.8–10.7% of Yw, Fig. 3D), administrative zones (7.5–11.4% of Yw, Fig. 3E), or farming systems (10.1–10.5% of Yw, Fig 3F). The resource yield gap was on average 15% of Yw and was smallest in the highland agro-ecological zones (SH2, H2, and SM2, 9.3–13.1% of Yw) and greatest in the SM3, M2, M3, and SH1 agro-ecological zones (> 15% of Yw, Fig. 3D). In terms of administrative zones, the resource yield gap was smaller than 10% of Yw in West Arsi, North Gonder, and Gurage

and above 20% of Yw in West Showa, South Wollo, and North Wollo (Fig. 3E). The resource yield gap was negligible for maize mixed farming systems and ca. 15% of Yw for highland mixed and highland perennial farming systems (Fig. 3F). These efficiency and resource yield gaps seem to be small when expressed in relation to Yw but we note they are far from insignificant when compared to Ya (Fig. 3). Finally, high seed rates and weed control together with N application rates of 150, 250, and 350 kg N ha^{-1} resulted in an average yield gap closure to 50, 60, and 70% of Yw, respectively. For instance, an application of 350 kg N ha^{-1} increased wheat yields to ca. 80% of Yw in sub-afroalpine agro-ecological zones (H3 and SM3; Fig. 3D) and to 60–75% of Yw in a number of administrative zones (Fig. 3E) and highland mixed farming systems (Fig. 3F).

In summary, fine-tuning current crop management practices can deliver the additional production needed to reach wheat self-sufficiency without expanding wheat area. However, further narrowing yield gaps towards Yw requires inputs and technologies currently lacking in highest-yielding fields. Technologies currently not used by many farmers include for instance mechanisation of land preparation, planting and harvesting operations, effective

control of pests, diseases, and weeds or other nutrients beyond N and P. Efficiency and resource yield gaps as large as current Ya and technology yield gaps as large as 50% of Yw have also been reported in other studies on wheat yield gaps in Ethiopia (Silva et al. 2019a) and Rwanda (Baudron et al. 2019), and on maize yield gaps in Ethiopia (Assefa et al. 2020) and Tanzania (van Dijk et al. 2017).

3.2 Drivers of yield variability and gaps at field level

3.2.1 Production frontier and yield variability

The magnitude, sign, and significance of the first-order terms of growth-defining, -limiting, and -reducing factors on wheat yields were consistent between the Cobb-Douglas and translog stochastic models fitted to the pooled sample (Table S1). Regarding the second-order terms, the translog model revealed positive quadratic effects of seed rate, N rate and herbicide use, a negative interaction between seed and N rates, and positive interactions between temperature seasonality and seed rate and hand-weeding, and between available water and herbicide use (Table S1).

Wheat yields decreased with increased growing degrees days and temperature seasonality, after controlling for other factors, and there were no significant differences across varieties and years (Table 3). There was a negative effect of aridity index on wheat yields, a result also found for maize yields in Ethiopia (Assefa et al. 2020). Seed rates had a significant positive effect on wheat yields and increasing the former by 1% resulted in ca. 0.10% increase of the latter. This positive association between plant population and wheat yields was also documented for wheat in Rwanda (Baudron et al. 2019). Crop establishment remains a challenge in smallholder conditions due to manual sowing which leads to large variation in sowing depths and heterogeneous plant populations across the field. Plots where water logging or drought were reported by the farmer yielded 35–45% less than plots where these were not reported, and plots with deeper soils yielded ca. 8% more than plots with medium or shallow soil depths. Frequent ploughing was found to increase wheat yields using the same household survey data (Abro et al. 2018) but in our analysis, this did not translate into significantly greater wheat yields (Table 3). We note the analysis of Abro et al. (2018) focused exclusively on investigating the effect of ploughing frequency on wheat yields while our analysis investigates a broader range of biophysical and management drivers, which overtake the level of yield variation explained by ploughing frequency.

There was a clear yield response to N across models estimated for the pooled sample and for specific administrative

regions: on average, wheat yields increased by ca. 0.27% with 1% increase in N applied (Table 3). Earlier studies also identified N fertilisation as a key determinant of wheat yields in Ethiopia (Habte et al. 2014; Tanner et al. 1993) but further research should investigate whether increasing N rates is economically viable for smallholders (cf. van Dijk et al. 2017). High fertiliser prices were identified as an important constraint to increase fertiliser access and use by wheat smallholders in Ethiopia (Anteneh and Asrat 2020). In addition to profitability, smallholders' decisions to apply fertiliser also depend on the area share of each crop, household wealth, access to rental land, and the level of land fragmentation (Yu and Nin-Pratt 2014). At regional level, the distance from the input distribution to the farmer was found to increase the price of mineral fertilisers as a result of greater transaction and transportation costs (Minten et al. 2013). Overcoming these constraints at farm and regional levels thus remains important to increase access to and use of fertilisers in the country.

Fertile plots yielded 6% and 16% more than medium and poor fertile plots, respectively. No significant yield differences were observed between plots with and without legumes as preceding crop, which is not in agreement with earlier empirical findings (Taa et al. 2004), nor between plots with and without manure application. This may be due to the heterogeneity in manure management and legume productivity and residue management between farms and to the relatively low number of fields with manure use reported ($n = 534$) and legumes recorded as previous crop ($n = 840$). Finally, herbicide use was positively associated with wheat yields (but the effects were small), pesticide use translated into 12% greater wheat yields, and disease occurrence reduced wheat yields by ca. 30%. There were no positive significant effects of hand-weeding on wheat yields as this operation might be done after the crop suffers from severe competition from weed as a result of labour shortages or inconvenient working days during more critical periods of the growing season.

The stochastic frontier model with a Cobb-Douglas functional form was fitted to a subset of the data for the administrative zones West Arsi, North Showa, East Gojam, and South Wollo (Table 3). The results obtained for these administrative zones were largely consistent with the results of the national analysis reported above, particularly for seed rate (only non-significant in West Arsi), N application rate (strongly positive in all zones), occurrence of drought (strongly negative in all zones), and occurrence of diseases (strongly negative in all zones). The most notable difference between both national and regional analyses was that the significance of biophysical variables (e.g. growing degrees day, temperature seasonality and aridity index) observed

Table 3 Parameter estimates of the stochastic frontier model estimated for wheat-based farming systems in Ethiopia (Meher seasons of 2009 and 2013; Eq. 1). The same model was fitted to the pooled sample (Ethiopia) and to selected administrative zones (West Arsi, North Showa, East Gojam, and South Wollo). Significance is indicated by the codes: '***' 0.1%, '**' 1%, '*' 5%, '#' 10%. *n.a.* not applicable

	Ethiopia	West Arsi	North Showa	East Gojam	South Wollo
Intercept	9.771***	0.144	−9.289	21.686	−8.166
Meher_2013	−0.054*	−0.152*	0.123	−0.064	−0.051
Defining factors					
Growing degrees day	−0.584***	−0.158	0.871	−1.756#	−0.771
Temperature seasonality	−0.325***	−0.089	0.195	0.168	1.861***
Seed rate (kg/ha)	0.098***	−0.040	0.359***	0.193***	0.247***
Variety_improved	0.002	0.015	−0.001	−0.078	0.030
Limiting factors (water)					
Aridity index	−0.331***	0.096	−0.372	−0.987	0.171
Soil available water	−0.008	0.350	0.294	0.307	0.277
Soil depth_medium	−0.074***	−0.103#	−0.116#	−0.057	−0.180***
Soil depth_shallow	−0.081**	−0.216*	−0.024	0.058	−0.099
Water logging_yes	−0.347***	−0.277#	−0.547***	0.012	−0.241***
Drought_yes	−0.447***	−0.660***	−0.605#	−0.455***	−0.338***
Water conservation_yes	0.059*	0.091	0.119	0.029	0.064
Plough frequency_three	−0.102#	0.246	−0.023	−0.015	−0.155*
Plough frequency_four	−0.013	0.158	0.019	0.032	−0.031
Plough frequency ≥five	0.053	0.211	−0.020	0.156	0.201
Limiting factors (nutrients)					
Soil fertility_poor	−0.163***	−0.219*	−0.065	−0.306***	−0.170*
Soil fertility_medium	−0.057**	−0.008	0.020	−0.177**	−0.043
N applied (kg N/ha)	0.272***	0.302***	0.419***	0.225***	0.262***
Manure use_yes	0.037	−0.103	0.148#	−0.015	0.043
Crop residues_yes	0.032	−0.068	0.033	0.094	−0.193*
Previous crop_legume	0.022	−0.112	0.096	0.143#	0.050
Previous crop_other	0.123***	0.107#	0.112	0.133	0.103
Reducing factors					
Herbicide use (L/ha)	0.013***	−0.003	−0.017	0.009	0.008
Hand-weeding (person-day/ha)	−0.004#	−0.007#	−0.010	0.005	0.162***
Weeding_yes	0.037	0.386#	0.128	−0.317	−2.252***
Pesticide use_yes	0.121*	0.209**	−0.272	0.025	n.a.
Disease occurrence_yes	−0.316***	−0.456***	−0.309**	−0.393***	−0.335***
Pest occurrence_yes	−0.086	−0.106	0.115	−0.009	−0.459*
Model evaluation					
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.598***	0.449***	0.605***	0.208***	0.604***
$\gamma = \sigma_u^2 / \sigma^2$	0.742***	0.831***	0.677***	0.571*	0.812***
Sample size (n)	3694	537	520	243	562

in the former tend to disappear in the latter and wheat yield responses to N were largest in North Showa. This is expected as the pooled sample used in the national analysis exhibits greater variation in biophysical conditions between households compared to the subset used for the regional analyses.

3.2.2 Resource yield gap and yield response to inputs

Y_{HF} were 3.4 t ha^{-1} in South Wollo, 3.1 t ha^{-1} in East Gojam, 4.2 t ha^{-1} in North Showa, and 4.5 t ha^{-1} in West Arsi (Fig. 4A and Table 2). Y_{AF} and Y_{LF} were higher in West Arsi (2.2 and 0.8 t ha^{-1}), intermediate in East Gojam
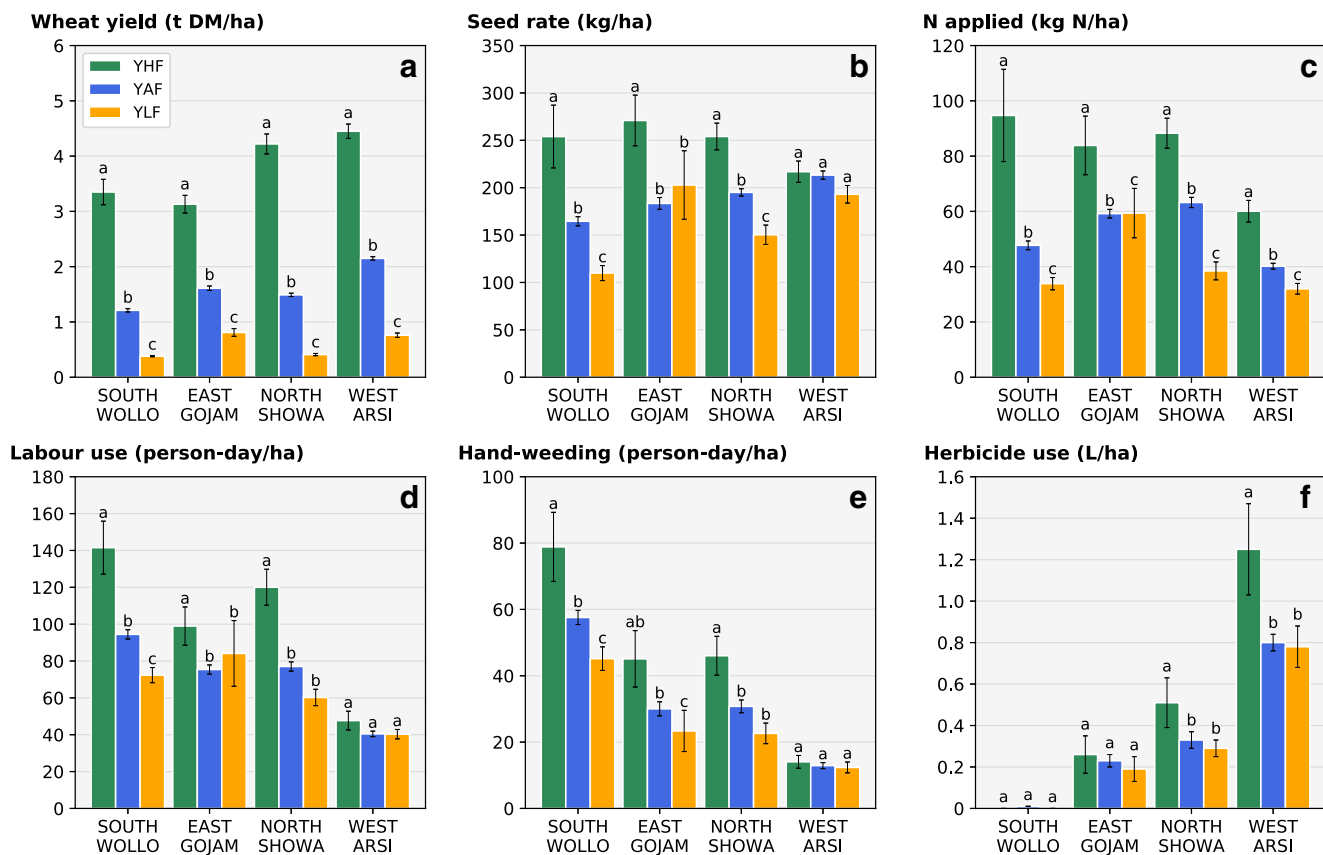
Fig. 4 Wheat yield and input use across highest- (Y_{HF}), average- (Y_{AF}), and lowest-yielding fields (Y_{LF}) in selected administrative zones of Ethiopia. Bars show the average value across the Meher seasons of 2009 and 2013. Error bars show the standard error of the mean and different letters show differences between groups in each zone at 5% significant level

(1.6 and 0.8 t ha^{-1}), and North Showa (1.5 and 0.4 t ha^{-1}) and lower in South Wollo (1.2 and 0.4 t ha^{-1}).

Y_{HF} were associated with significantly greater seed and N application rates compared to Y_{AF} and/or Y_{LF} across all four administrative zones (Fig. 4B and C). Seed rates in highest-yielding fields were ca. 250 kg ha^{-1} in North Showa, East Gojam, and South Wollo, which was significantly greater than the average 180 kg ha^{-1} used in average- and lowest-yielding fields. The variation in seed rates between field classes was smaller (and not significant) in West Arsi compared to other administrative zones: ca. 220 and 190 kg ha^{-1} in highest- and lowest-yielding fields, respectively. N application rates in highest-yielding fields were ca. 90 kg N ha^{-1} in North Showa, East Gojam, and South Wollo, which was significantly greater than the ca. 60 kg N ha^{-1} used in average-yielding fields in North Showa and East Gojam, the ca. 45 kg N ha^{-1} used in South Wollo, and the ca. 30 kg N ha^{-1} (60 kg N ha^{-1}) observed across the lowest-yielding fields in North Showa and South Wollo (East Gojam).

Labour use for land preparation, sowing, hand-weeding and harvesting was significantly greater for Y_{HF} than for Y_{AF} and Y_{LF} in all administrative zones except West Arsi (Fig. 4D and E). Significant differences in herbicide use across groups were only observed in West Arsi and North Showa (Fig. 4F). As an example, highest-yielding fields were associated with a total labour use of ca. 140, 120, and 100 person-day ha^{-1} in South Wollo, North Showa, and East Gojam, respectively, while labour use ranged between 60–80 person-day ha^{-1} in the lowest-yielding fields of these zones. Considerably more labour was used in North Showa, East Gojam, and South Wollo than in West Arsi and there was an inverse relationship between labour use for hand-weeding and herbicide use (Fig. 4E and F). This is best seen in West Arsi where herbicide use was greatest (ca. 0.8 L ha^{-1}) and labour for hand-weeding was lowest (ca. 12 person-day ha^{-1}), data which validate those of an independent household survey conducted in 2012 in the same region and analysed by Silva et al. (2019a).

3.2.3 Technology yield gap and increased amounts of inputs

The simulated Yw in Arsi administrative zone was 8.6 t ha^{-1} in 2009 and 9.7 t ha^{-1} in 2013, which was

considerably greater than the values observed for Y_{HF} during the same years (Figs. S5A and S5B). The variety trials described by Bezabih et al. (2018) were conducted at Kulumsa Agricultural Research Center (KARC), Arsi administrative zone, in 2016 and 2017 and simulated Yw was 9.3 and 9.1 t ha^{-1} in these years, respectively. The yields observed in these trials ranged between 4.9 and 7.7 t ha^{-1} in 2016 and between 4.4 and 7.0^{-1} in 2017 (Figs. S5C and S5D). Despite differences in varieties used in the highest-yielding fields and the variety trials (data not shown), most of the varieties cultivated in highest-yielding fields were improved genotypes bred at KARC with parental material from CIMMYT (Fig. S4). This means farmers use improved varieties that can reach up to ca. 80% of Yw on-station and, hence, that the technology yield gap is likely caused by other factors than lack of improved varieties.

The low amount of inputs (particularly seeds and fertilisers) used in highest-yielding fields compared to what is needed to reach Yw (Fig. 4C and Section 2.3.3) and the lack of certain inputs and technologies in these fields are the most likely drivers of the technology yield gap of wheat across Ethiopia. The former is reflected by the difference between the feasible yield (Yf) and Y_{HF}, and it is depicted in Fig. 3 as the additional resource yield gap for increasing amounts of N applied (150, 250, and 350 kg N ha^{-1}). For instance, combining high seed rates with intensive weeding practices and 150 kg N ha^{-1} can increase wheat yields up to an average Yf of 4.6 t ha^{-1}. This corresponds to a technology yield gap of ca. 47% of Yw. Applying N rates of 250 and 350 kg N ha^{-1} result in average Yf of 5.3 and 5.8 t ha^{-1} which reduces the technology yield gap to about 39 and 33% of Yw, respectively. We note the aforementioned N application rates are way above those currently observed in highest-yielding fields (Fig. 4C) but these high N rates could be reduced if efficient N management practices are adopted cf. (Assefa et al. 2020; ten Berge et al. 2019). Another point of concern is that most wheat in Ethiopia is currently cultivated in acid soils where yield responses to N are not always clear and hence, increasing the N rates further aggravates soil acidity and lowers wheat yield (Regassa and Agegnehu 2011). This clearly indicates that the scope to increase fertiliser rates is context specific and needs to be integrated with other soil management practices and excellent agronomy. Other factors explaining this yield gap may include poor crop establishment and poor weed control, which currently rely heavily on draught power and manual labour, and poor pest and disease control (as partly shown in Table 3). It is also important to consider that row planting improves radiation interception under high seed rates (high plant populations) as compared to the current farmer practice of broadcasting (Alemu et al. 2014).

## 3.3 Farming systems analysis across different zones

3.3.1 Crop diversity at farm level

The total cultivated land area per farm was on average 2 ha in West Arsi and North Showa and 1.6 and 1.4 ha in East Gojam and South Wollo, respectively (Figs. 5 and S7). This land was allocated differently to different crops in different administrative zones. The share of wheat in the total cultivated land was high in South Wollo and West Arsi, on average ca. 45%, and low in North Showa (35%) and East Gojam (25%). In West Arsi, households allocated 47% and 6% of their cultivated land to other cereals (mostly barley) and to legumes (mostly faba bean), respectively (Fig. 5A). In North Showa, the share of other cereals (mostly barley and red tef) and legumes (mostly faba bean) of the total cultivated land was ca. 37% and 27%, respectively (Fig. 5B). In East Gojam, around 60% of the cultivated land was allocated to other cereals (mostly red and white tef) and only ca. 10% was cultivated with legumes (Fig. 5C). In South Wollo, both other cereals and legumes were cultivated on ca. 25% of the total cultivated land (Fig. 5D). This indicates farms in North Showa and East Gojam are more diverse regarding the crop types cultivated than farms in West Arsi and South Wollo (Fig. 4D and E). The different crop types are known to compete for labour during key periods of the growing season (Silva et al. 2019a), but we were not able to find clear substitution or competition for land and labour between wheat and other crops possibly due to a lack of information on the timing of different operations (Fig. S6). Further research is thus needed to clarify the importance of wheat as a source of income, and priority for investment, in the more diversified farming systems.

3.3.2 Availability of land, labour, and capital

Oxen ownership was associated with slightly greater wheat yields in West Arsi, North Showa, and East Gojam but the effects were only significant in East Gojam (Fig. 6A). In addition, households with more oxen pairs tended to cultivate larger wheat areas than households with few oxen pairs (Fig. 6B). This was particularly true in West Arsi and North Showa, where land is more 'abundant' (Fig. 5A and B), and not as much and significantly in East Gojam and South Wollo, where land is constrained (Fig. 5C and D). No significant differences in total labour use for wheat were observed for different levels of oxen ownership in either zone (Fig. 6C); hence, oxen ownership did not translate into labour savings per unit land and into substitution of manual labour by draught power. The economic value of farm assets increased on average with increasing oxen ownership,

Fig. 5 Cultivated land per household during the 2013 Meher seasons in selected administrative zones of Ethiopia: (A) West Arsi, (B) North Showa, (C) East Gojam, and (D) South Wollo. Different colours depict different types of crops. Data for 2009 are provided in Fig. S7
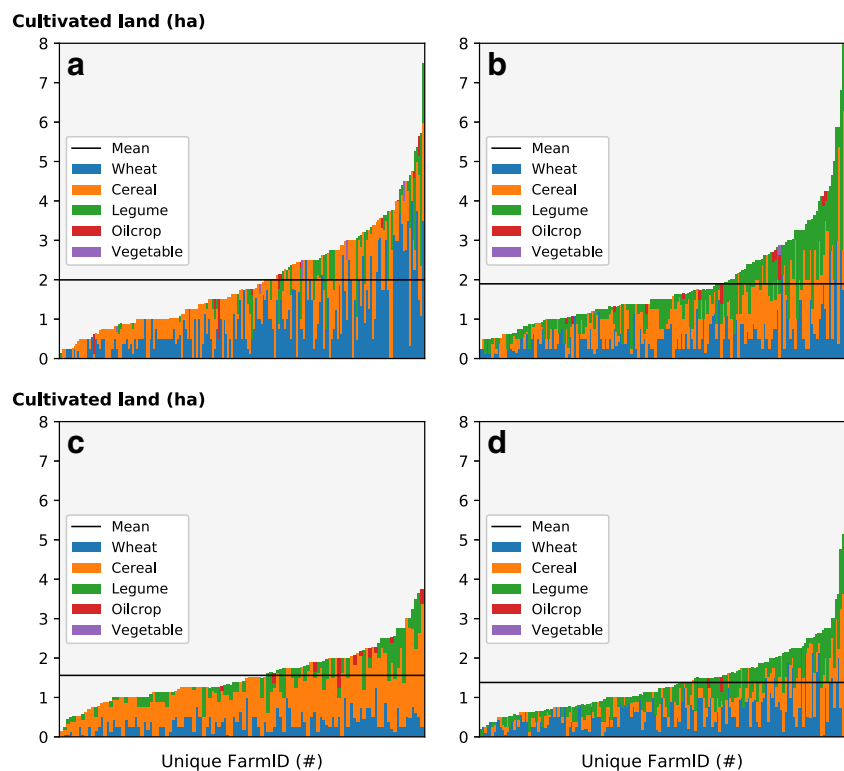
which was particularly clear in West Arsi and North Showa (Fig. 6D). No major significant differences in input use for wheat were observed across different levels of oxen ownership in either zone (Fig. S8).

In summary, oxen ownership was a proxy for draught power and capital availability and was associated with larger wheat area, particularly in the administrative zones with largest cultivated area per household (i.e. West Arsi and North Showa; Silva et al. 2019a). Hence, access to draught power and capital translates into increases in wheat production through expansion of cultivated land and not so much through intensification of wheat production via yield gap closure (Fig. 6A and B).

3.3.3 Comparisons across administrative zones

The four administrative zones analysed in greater depth in this study capture differences in the level of intensification of wheat production (Fig. 4) and in farming systems regarding the crop area shares and oxen ownership (Fig. S1). Wheat yields were greatest in West Arsi, intermediate in North Showa and East Gojam, and smallest in South Wollo, while the opposite was true for labour use for wheat (both total and hand-weeding; Fig. 4). West Arsi is distinct from the other zones mostly because herbicides are widely used, substituting labour for hand-weeding and possibly other inputs (e.g. N) in the short term (Fig. 4C and E). We also note that the cultivated area per farm is greatest

in West Arsi and smallest in South Wollo, which may explain the use of herbicides in the former and the heavy reliance on human labour in the latter (Fig. 5). Finally, there was a positive relationship between the number of pairs of oxen (a proxy for capital availability) and the wheat area cultivated per farm in West Arsi and North Showa, the zones where cultivated land per farm was greatest, while no relationship was observed between the number of pairs of oxen and the input use for wheat (Fig. 6). This means that increases in wheat production are mostly obtained through increases in cultivated areas rather than through yield gap closure and that households with more capital do not necessarily use more inputs for wheat. These results suggest that smallholders do not have proper access to inputs because these are either too expensive or unavailable when needed, which can be pointed as the main challenge for intensification of wheat production and the achievement of wheat self-sufficiency in Ethiopia without expansion of wheat area.

4 Conclusion

Wheat yields across farmers' fields in Ethiopia were only up to ca. 20% of the water-limited yield potential, the benchmark for what can be achieved with best agronomic practices under rainfed conditions. Most of the yield gap was attributed to the technology yield gap, meaning that
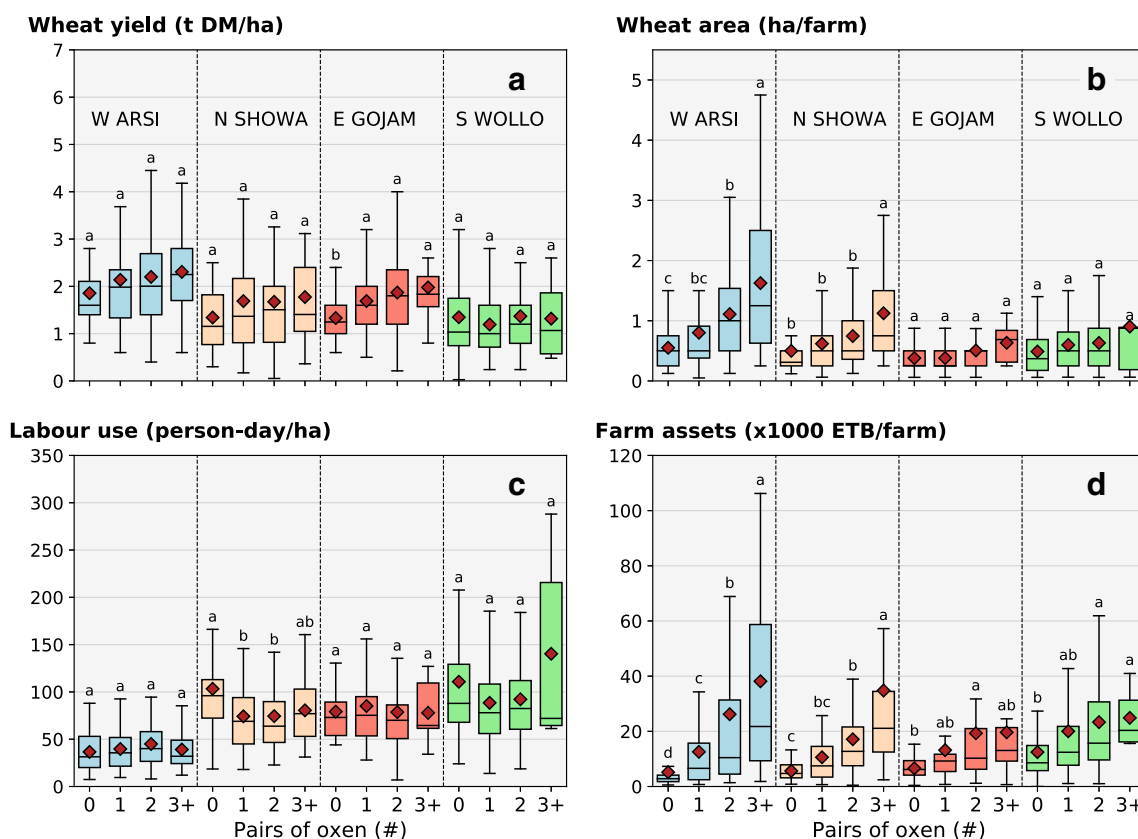
INRA℮ 🦋 Springer

Fig. 6 Relationship between number of pairs of oxen owned by households and (A) wheat yields, (B) wheat cultivated area, (C) labour use for land preparation, sowing, hand-weeding, and harvesting of wheat, and (D) farm assets owned by households for selected administrative zones in Ethiopia (West Arsi, North Showa, East Gojam, and South Wollo). For each zone, lower-case letters depict significant differences between groups at 5% significance level

certain inputs and technologies are entirely lacking in highest-yielding fields (such as technologies for optimal crop establishment and for control of pests, diseases, and weeds) and that the current input levels are not high enough to reach the water-limited yield. Despite their small share in explaining the wheat yield gap, narrowing the efficiency and resource yield gaps can nearly double actual yields and contribute to realise the yield progress needed to achieve wheat self-sufficiency in Ethiopia without having to expand the wheat area. However, achieving this requires increases in input use to the levels observed in highest-yielding fields and fine-tuning current crop management practices in relation to the time, space, and form of the inputs used.

Wheat is cultivated in Ethiopia alongside other cereal and legume crops, particularly in North Showa and East Gojam and to a lesser extent in West Arsi and South Wollo. This diversity is important for food and nutrition security at household level. We found no clear evidence of substitution of or competition for land and labour between wheat and other crops. Yet, this finding merits further research because smallholders are known to operate under resource constraints and the different crops compete for

labour in key periods of the growing season. Our results also indicate that households with more access to draught power and capital are not investing more in inputs per hectare of wheat but rather in allocating more land to wheat. This was particularly true in zones where the total cultivated land per household is relatively high, such as West Arsi and North Showa.

Here we show for the first time that it is possible to achieve wheat self-sufficiency in Ethiopia with current technologies (e.g. varieties) but that greater amounts, and more efficient use, of inputs are needed to do so. Narrowing technology yield gaps is also essential given the rapidly increasing demand for cereals due to population growth and dietary change (van Ittersum et al. 2016). Government policies aiming to increase wheat production should focus on fostering the accessibility and affordability of inputs, particularly fertilisers, and on promoting technologies that allow for a more precise management of these inputs (e.g. mechanisation and herbicides). This will also be essential to avoid environmental externalities of an intensification. Such policies also need to consider that wheat is one of the many crops cultivated by farmers, whose

livelihoods should not be forgotten. It is thus important to understand whether or not narrowing yield gaps towards 80% of Yw is desirable from, e.g. an economic, environmental, or labour productivity perspective under prevailing conditions.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Author contributions Conceptualisation: JVS, PR, MKvI; Software: JVS; Validation: FB, MJ, KT; Formal analysis: JVS; Investigation: JVS; Resources: PR, FB, MJ, MKvI; Data curation: JVS, MJ; Writing—original draft preparation: JVS; Writing—review and editing: JVS, PR, FB, MJ, KT, MKvI; Supervision: PR, MKvI; Project administration: MKvI; Funding acquisition: MKvI.

Data availability The data that support the findings of this study are available from the corresponding author, JVS, upon reasonable request.

Disclaimer Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of CRP WHEAT or the institutions the authors are affiliated to.

Supplementary Information The online version of this article (https://doi.org/10.1007/s13593-020-00654-z) contains supplementary material, which is available to authorised users.

References

Abro ZA, Jaleta M, Qaim M (2017) Yield effects of rust-resistant wheat varieties in Ethiopia. Food Secur 9:1343–1357. https://doi.org/10.1007/s12571-017-0735-6

Abro ZA, Jaleta M, Teklewold H (2018) Does intensive tillage enhance productivity and reduce risk exposure? Panel data evidence from smallholders' agriculture in Ethiopia. J Agric Econ 69:756–776. https://doi.org/10.1111/1477-9552.12262

Alemu T, Emana B, Haji J, Legesse B (2014) Impact of wheat row planting on yield of smallholders in selected highland and lowland areas of Ethiopia. Int J Agric Forestry 4:386–393. https://doi.org/10.5923/j.ijaf.20140405.07

Amede T, Auricht C, Boffa JM, Dixon J, Mallawaarachchi T, Rukuni M, Teklewold-Deneke T (2017) A farming system framework for investment planning and priority setting in Ethiopia. ACIAR Technical Reports Series No. 90. Australian Centre for International Agricultural Research, Canberra

Anteneh A, Asrat D (2020) Wheat production and marketing in Ethiopia: review study. Cogent Food Agric 6:1778893. https://doi.org/10.1080/23311932.2020.1778893

Assefa BT, Chamberlin J, Reidsma P, Silva JV, van Ittersum MK (2020) Unravelling the variability and causes of smallholder maize yield gaps in Ethiopia. Food Secur 12:83–103. https://doi.org/10.1007/s12571-019-00998-9

Battese GE, Coelli TJ (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empir Econ 20:325–332. https://doi.org/10.1007/BF01205442

Baudron F, Ndoli A, Habarurema I, Silva JV (2019) How to increase the productivity and profitability of smallholder rainfed wheat in the Eastern African highlands? Northern Rwanda as a case study. Field Crops Res 236L121–131. https://doi.org/10.1016/j.fcr.2019.03.023

Beza E, Silva JV, Kooistra L, Reidsma P (2017) Review of yield gap explaining factors and opportunities for alternative data collection approaches. Eur J Agron 82, Part B:206–222. https://doi.org/10.1016/j.eja.2016.06.016

Bezabih M, Adie A, Ravi D, Prasad KVSV, Jones C, Abeyo B, Tadesse Z, Zegeye H, Solomon T, Blümmel M (2018) Variations in food-fodder traits of bread wheat cultivars released for the Ethiopian highlands. Field Crops Ress 229:1–7. https://doi.org/10.1016/j.fcr.2018.09.006

Boogaard HL, de Wit AJW, Te Roller JA, van Diepen CA (2014) WOFOST CONTROL CENTRE 2.1; User's guide for the WOFOST CONTROL CENTRE 2.1 and the crop growth simulation model WOFOST 7.1.7. Tech. rep, Wageningen University Research Centre, Alterra, Wageningen

Coelli TJ, Henningsen A (2013) Frontier: stochastic frontier analysis. R package version 1.1-0

CSA (2019) Agriculture sample survey 2018/2019: volume I report on area and production of major crops. 589 Statistical Bulletin. Addis Ababa, Ethiopia. Technical report

Grassini P, van Bussel LGJ, van Wart J, Wolf J, Claessens L, Yang H, Boogaard H, de Groot H, van Ittersum MK, Cassman KG (2015) How good is good enough? Data requirements for reliable crop yield simulations and yield gap analysis. Field Crops Res 177L49–63. https://doi.org/10.1016/j.fcr.2015.03.004

Habte D, Tadesse K, Admasu W, Desalegn T, 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. Afr J Agric Res 10:89–99. https://doi.org/10.5897/AJAR2014.9132

Hochman Z, Horan H (2018) Causes of wheat yield gaps and opportunities to advance the water-limited yield frontier in Australia. Field Crops Res 228:20–30. https://doi.org/10.1016/j.fcr.2018.08.023

Hodson DP, White JW (2007) Use of spatial analyses for global characterization of wheat-based production systems. J Agric Sci 145:115–125. https://doi.org/10.1017/S0021859607006855

Jaleta M, Hodson D, Abeyo B, Yirga C, Erenstein O (2019) Smallholders' coping mechanisms with wheat rust epidemics: lessons from Ethiopia. PloS one, 14:e0219327. https://doi.org/10.1371/journal.pone.0219327

Minten B, Koru B, Stifel D (2013) The last mile(s) in modern input distribution: pricing, profitability, and adoption. Agric Econ 44:629–646. https://doi.org/10.1111/agec.12078

MoA (1998) Agroecological zones of Ethiopia. Natural Resources Management and Regulatory Department and GIZ, Addis Ababa, Ethiopia

Negassa A, Shiferaw B, Koo J, Sonder K, Smale M, Braun H, Gbegbelegbe S, Guo Z, Hodson D, Wood S, Payne T, Abeyo B (2013) The potential for wheat production in Africa: analysis of biophysical suitability and economic profitability. Technical report, CIMMYT, Mexico

Rattalino-Edreira JI, Mourtzinis S, Conley SP, Roth AC, Ciampitti IA, Licht MA, Kandel H, Kyveryga PM, Lindsey LE, Mueller DS, Naeve SL, Nafziger E, Specht JE, Stanley J, Staton MJ, Grassini P (2017) Assessing causes of yield gaps in agricultural areas with diversity in climate and soils. Agric Forest Meteorol 247:170–180. https://doi.org/10.1016/j.agrformet.2017.07.010

Regassa H, Agegnehu G (2011) Potentials and limitations of acid soils in the highlands of Ethiopia: a review. In: Barley Research and Development in Ethiopia. ICARDA, Aleppo

Seabold S, Perktold J (2010) statsmodels: econometric and statistical modeling with python. In: 9th Python in Science Conference

Silva JV, Reidsma P, Laborte AG, van Ittersum MK (2017) Explaining rice yields and yield gaps in Central Luzon, Philippines: an application of stochastic frontier analysis and crop modelling. Eur J Agron 82, Part B:223–241. https://doi.org/10.1016/j.eja.2016.06.017

Silva JV, Baudron F, Reidsma P, Giller KE (2019a) Is labour a major determinant of yield gaps in sub-Saharan Africa? A study for cereal-based production systems in Southern Ethiopia. Agric Syst 175:39–51. https://doi.org/10.1016/j.agsy.2019.04.009

Silva JV, Ramisch JJ (2019b) Whose gap counts? The role of yield gap analysis within a development-oriented agronomy. Exper Agric 55:311–338. https://doi.org/10.1017/S0014479718000236

Taa A, Tanner DG, Bennie ATP (2004) Effects of stubble management, tillage and cropping sequence on wheat production in the south-eastern highlands of Ethiopia. Soil Tillage Res 76:69–82. https://doi.org/10.1016/j.still.2003.08.002

Tanner DG, Gorfu A, Taa A (1993) Fertiliser effects on sustainability in the wheat-based small-holder farming systems of southeastern Ethiopia. Field Crops Res 33:235–248. https://doi.org/10.1016/0378-4290(93)90082-X

ten Berge HFM, Hijbeek R, van Loon MP, Rurinda J, Tesfaye K, Zingore S, Craufurd P, van Heerwaarden J, Brentrup F, Schröder JJ, Boogaard HL, de Groot HLE, van Ittersum M. K (2019) Maize crop nutrient input requirements for food security in sub-Saharan Africa. Global Food Secur 23:9–21. https://doi.org/10.1016/j.gfs.2019.02.001

Tesfaye K (2016) Application of the GYGA approach to Ethiopia. Tech. rep, Global Yield Gap Atlas, http://www.yieldgap.org/ethiopia

Tolemariam A, Jaleta M, Hodson D, Alemayehu Y, Yirga C, Abeyo B (2018) Wheat varietal change and adoption of rust resistant wheat varieties in Ethiopia from 2009/10 to 2013/14. CIMMYT, Socioeconomics Working paper 12

van Dijk M, Morley T, Jongeneel R, van Ittersum MK, Reidsma P, Ruben R (2017) Disentangling agronomic and economic yield gaps: an integrated framework and application. Agric Syst 154:90–99. https://doi.org/10.1016/j.agsy.2017.03.004

van Ittersum MK, Rabbinge R (1997) Concepts in production ecology for analysis and quantification of agricultural input-output combinations. Field Crops Res 52:197–208. https://doi.org/10.1016/S0378-4290(97)00037-3

van Ittersum MK, Cassman KG, Grassini P, Wolf J, Tittonell P, Hochman Z (2013a) Yield gap analysis with local to global relevance - a review. Field Crops Res 143:4–17. https://doi.org/10.1016/j.fcr.2012.09.009

Van Wart J, van Bussel LGJ, Wolf J, Licker R, Grassini P, Nelson A, Boogaard H, Gerber J, Mueller ND, Claessens L, van Ittersum MK, Cassman KG (2013b) Use of agro-climatic zones to upscale simulated crop yield potential. Field Crops Res 143:44–55. https://doi.org/10.1016/j.fcr.2012.11.023

van Bussel LGJ, Grassini P, van Wart J, Wolf J, Claessens L, Yang H, Boogaard H, de Groot H, Saito K, Cassman KG, van Ittersum MK (2015) From field to atlas: upscaling of location-specific yield gap estimates. Field Crops Res 177L98–108. https://doi.org/10.1016/j.fcr.2015.03.005

van Ittersum MK, van Bussel LGJ, 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 PAJ, van Loon MP, Saito K, Adimo O, Adjei-Nsiah S, Agali A, Bala A, Chikowo R, Kaizzi K, Kouressy M, Makoi JHJR, Ouattara K, Tesfaye K, Cassman KG (2016) Can sub-Saharan Africa feed itself? Proc Natl Acad Sci 113:14964–14969. https://doi.org/10.1073/pnas.1610359113

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Jarrod Millman K, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey C, Polat İ, Feng Y, Moore EW, Vand erPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, Contributors S (2020) SciPy 1.0: fundamental algorithms for scientific computing in python. Nat Methods 17:261–272. https://doi.org/10.1038/s41592-019-0686-2

Yu B, Nin-Pratt A (2014) Fertilizer adoption in Ethiopia cereal production. J Dev Agric Econ 6:318–337