



# Narrowing maize yield gaps across smallholder farming systems in Zambia: what interventions, where, and for whom?

João Vasco Silva<sup>1,2</sup> · Frédéric Baudron<sup>1</sup> · Hambulo Ngoma<sup>1</sup> · Isaiah Nyagumbo<sup>1</sup> · Esau Simutowe<sup>3</sup> · Kelvin Kalala<sup>3</sup> · Mukwemba Habeenzu<sup>3</sup> · Mtendere Mphatso<sup>4</sup> · Christian Thierfelder<sup>1</sup>

Accepted: 5 January 2023  
© The Author(s) 2023

## Abstract

Maize production in Zambia must increase with a view towards improved food security and reduced food imports whilst avoiding cropland expansion. To achieve this, it is important to understand the causes behind the large maize yield gaps observed in smallholder farming systems across the country. This is the first study providing a yield gap decomposition for maize in Zambia, and combining it with farm typology delineation, to identify the key limiting factors to maize yield gaps across the diversity of farms in the country. The analysis builds upon a nationally representative household survey covering three growing seasons and crop model simulations to benchmark on-farm maize yields and N application rates. Three farm types were delineated, including households for which maize is a marginal crop, households who are net buyers of maize, and households who are market-oriented maize producers. Yield gap closure was about 20% of the water-limited yield, corresponding to an actual yield of 2.4 t ha<sup>-1</sup>. Market-oriented maize farms yielded slightly more than the other farm types, yet the drivers of yield variability were largely consistent across farm types. The large yield gap was mostly attributed to the technology yield gap indicating that more efficient production methods are needed to raise maize yields beyond the levels observed in highest yielding fields. Yet, narrowing efficiency and resource yield gaps through improved crop management (i.e., sowing time, plant population, fertilizer inputs, and weed control) could more than double current yields. Creating a conducive environment to increase maize production should focus on the dissemination of technologies that conserve soil moisture in semi-arid areas and improve soil health in humid areas. Recommendations of sustainable intensification practices need to consider profitability, risk, and other non-information constraints to improved crop management and must be geographically targeted to the diversity of farming systems across the country.

**Keywords** Food security · Sustainable intensification · Farm typology · Global Yield Gap Atlas · Fertilizer input subsidy program

## 1 Introduction

Economic development in Zambia is strongly linked to productivity growth in agriculture and sustainable management

✉ João Vasco Silva  
j.silva@cgiar.org; jvasco323@posteo.net

- <sup>1</sup> Sustainable Agrifood Systems Program, CIMMYT-Zimbabwe, Harare, Zimbabwe
- <sup>2</sup> Plant Production Systems Group, Wageningen University & Research, Wageningen, The Netherlands
- <sup>3</sup> Sustainable Agrifood Systems Program, CIMMYT-Zambia, Lusaka, Zambia
- <sup>4</sup> Food and Agriculture Organization of the United Nations, FAO-Zambia, Lusaka, Zambia

of farming systems (IAPRI 2020). Approximately 75% of the population rely on smallholder farming for their livelihoods (MoA/CSO 2019). Maize (*Zea mays* L.) is the main staple food crop in the country, as in other Southern African countries (Smale 1995), with a harvested area of approximately 1 Mha and providing 50–90% of the caloric intake of the national population. Maize production in Zambia is associated with low use of mineral fertilizers and low adoption of other sustainable intensification practices (e.g., conservation agriculture and improved maize legume cropping systems; Arslan et al. 2014). Poor soil fertility and adverse effects of increased climate variability reduce farmers' financial resource base (Komarek et al. 2019) and contribute to low adaptive capacity of maize-based farming systems in the country (Cairns et al. 2013).

Smallholder farming systems in sub-Saharan Africa are highly diverse and farm typologies have proven useful to identify farms with different levels of resource endowments and livelihood strategies (Tittonell et al. 2010). The same is true in Zambia where approximately 1.6 million farmers are considered small scale with 70% having farm sizes below 2 ha, 25% having farm sizes between 2 and 5 ha, and 5% having farm sizes between 5 and 20 ha (Ngoma et al. 2019), and where poor subsistence farming co-exists with more market-oriented emerging commercial farming (Alvarez et al. 2018). Grain legumes are often produced alongside maize (Mwila et al. 2021) and livestock is kept in dry land areas of Southern and Western provinces characterized by low and erratic rainfall. Identifying different farm types is a means to consider farmers' socio-economic context and resource endowment when promoting agricultural technologies (e.g., Jayne et al. 2019) and an important first step to target technologies for different farm types (Berre et al. 2017).

Yield gaps of rain-fed crops are defined as the difference between the water-limited yield ( $Y_w$ ) and the actual yield ( $Y_a$ ) observed in farmers' fields (van Ittersum et al. 2013).  $Y_w$  is defined as the maximum yield that can be obtained under rain-fed conditions in a well-defined biophysical environment and without nutrient limitations or yield reductions due to pests, diseases, or weeds. Currently,  $Y_a$  for maize in Zambia ranges between 1.4 and 3.0 t ha<sup>-1</sup>, which is considerably lower than a  $Y_w$  of 8–15 t ha<sup>-1</sup> that could be achieved with best agronomic practices (Figure 1; van Ittersum et al. 2016). Yield gap decomposition is a means to unpack the causes behind yield gaps as it identifies the key crop management factors limiting or reducing  $Y_a$  (Silva et al. 2017). The resource yield gap indicates the scope to increase  $Y_a$  through higher amounts of inputs, whereas the efficiency yield gap indicates the scope to increase  $Y_a$  through fine tuning current management practices and technologies in terms of the time, space, and application form of these inputs. The technology yield gap indicates the possible yield increases beyond current best performing technologies on-farm. This decomposition is important to derive policy recommendations and prioritize research and development interventions towards increasing maize yields in existing cropland as food security and biodiversity conservation are dependent on such improvements.

This is the first study providing a yield gap decomposition for maize in Southern Africa and combining it with farm typology delineation to identify what interventions are needed, where, and for which farm types to narrow existing yield gaps. We hypothesized that the magnitude and the determinants of the yield gap differ across farm types with different production orientations and resource endowments. The main objective of this study was thus to characterize farm diversity across maize-based farming systems in



**Fig. 1** Maize yield gaps in Eastern Zambia. Maize plants on the left refer to an on-farm baby trial under good agronomic management (i.e., timely sowing, high plant population, hybrid maize variety, and proper fertilizer inputs). Maize plants on the right show crop performance under actual farm management. Credits: J.V. Silva, February 2022.

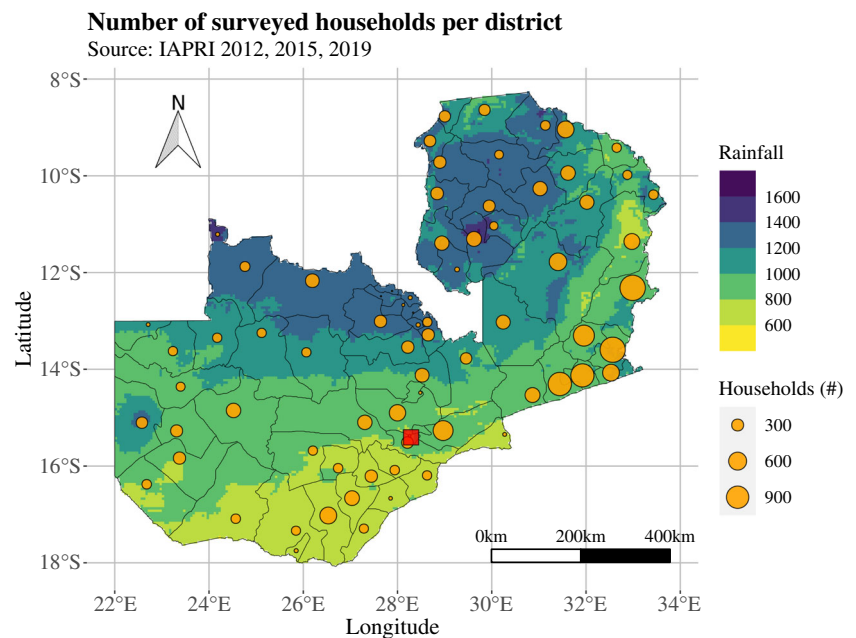
Zambia, and to identify the key limiting factors to maize yield gaps across the diversity of farms in the country. The analyses built upon a nationally representative household survey covering the 2011/12, 2014/15 and 2017/18 growing seasons (Figure 2; IAPRI 2012, 2015, 2019). Multivariate statistical techniques were used to construct the farm typology (Alvarez et al. 2018) and yield gaps were decomposed using a combination of frontier analysis and crop modeling (Silva et al. 2017). The latter was used to simulate  $Y_w$  and estimate the nitrogen (N) rates needed to reach it, which were then used to benchmark maize yields and N rates observed in farmers' fields.

## 2 Materials and methods

### 2.1 Rural Agricultural Livelihoods Survey (RALS)

Data from the Rural Agricultural Livelihoods Survey (RALS) was used to identify the main farm types engaged in maize production and to determine the drivers of maize yield variability in Zambia. The RALS comprises a panel of households interviewed over three different periods and is statistically representative of the rural population at the province and national levels. The surveys were conducted by the Indaba Agricultural Policy Research Institute (IAPRI) in collaboration with the Ministry of Agriculture and the Zambia Statistics Agency. The first round of RALS was conducted in May/June 2012, the second in June/July 2015, and the third in June/July 2019. The months when the RALS were conducted coincide with the harvesting period of the previous agricultural production season and with the agricultural marketing season. A total of 8839, 7934, and 7241 households were surveyed in 2012,

**Fig. 2** Spatial distribution of the households included in the Rural Agricultural Livelihoods Survey (RALS) across Zambia. Background layer displays the total annual rainfall (in mm) average over the period 2000–2019. *Source:* Climate Hazards Group Infra-Red Precipitation with Station data (CHIRPS; Funk et al. 2015).



2015, and 2019, respectively, with 6531 panel households interviewed in all three waves.

The spatial distribution of households included in the RALS is provided in Figure 2. The survey requested information on farm(er) characteristics and on field-specific crop management practices, thus meeting the requirements for yield gap decomposition (Beza et al. 2017). A unimodal rainfall regime with one wet season lasting from November to April in each year was observed across the country (Herrmann and Mohr 2011). Yet, annual rainfall was lowest in the Southern and Western regions of Zambia, with an average between 600 and 800 mm per year, intermediate in the central regions, with an average between 800 and 1200 mm per year, and highest in the Northern regions, with an average above 1200 mm per year (Figure 2).

Secondary data were retrieved from spatial products using the GPS coordinates of the individual households. Climatic data were retrieved from the climate zone scheme of the Global Yield Gap Atlas (GYGA) and comprised three variables: growing degrees days, temperature seasonality, and aridity index (Van Wart et al. 2013). Soil data on clay, silt and sand contents, pH in water and exchangeable acidity were retrieved from SoilGrids at 250m resolution (Hengl et al. 2017) and on rooting depth and soil available water from AfsIS-GYGA (Leenaars et al. 2015). Simulated water-limited yields for maize were retrieved from GYGA. Rainfall data were obtained from Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS, Funk et al. 2015) and used to determine the dekad corresponding to the onset of the rains for each of the growing seasons surveyed. The onset of the rains was defined as the first dekad with a cumulative rainfall equal to or greater than

25mm between the months of September and December (Hachigonta et al. 2008).

## 2.2 Farm typology delineation

The farm typology was constructed using principal component analysis (PCA) followed by hierarchical clustering (HC; Alvarez et al. 2018) on the pooled data. PCA is a technique used to reduce the number of dimensions in a dataset to a few synthetic and uncorrelated variables called principal components. The principal components are linear combinations of the original variables, which can be conceptualized as the directions of high-dimensional data that capture the maximum amount of variance and project it onto a smaller dimensional subspace. The principal components retained for analysis were those with an eigenvalue greater than one. PCA was conducted in R using the *dudi.pca()* function of the *ade4* package (Dray and Dufour 2007). HC refers to the hierarchical decomposition of the data based on group similarities and was then applied to a distance matrix calculated for the principal components selected following the PCA. Similarities between clusters were calculated using the Ward method. The final number of clusters was identified through visual inspection of the resulting dendrogram aiming to reach not less than three and not more than five clusters. HC was conducted with the *hclust()* function of the *R stats* package (R Core Team 2013).

Thirteen variables aggregated at the farm level were used to construct the farm typology, seven of which were structural variables (i.e., describing the structure of the household, variables that tend to remain constant from one season to the next) and six of which were

functional variables (i.e., describing the performance of the household). The farm(er) characteristics included in the typology were the age of the household head (years), household size (#), and area of owned cultivated land (ha) at the time of the surveys. Resource endowments were captured with variables referring to the cash available to each household (ZMW), farm assets calculated as the sum of the assets owned by each household multiplied by their respective economic value (in Zambian Kwacha, ZMW), total cultivated land in ha, and livestock ownership in tropical livestock units (TLU; Jahnke 1982) for each survey year. The total amount of maize produced, sold and bought per farm (all in kg) and the area cultivated with maize and legumes (both in ha) were included to assess the level of engagement of each farm in maize and legume production, whereas the total fertilizer use at farm level (in kg) was included to assess the level of agricultural intensification of each farm. Variables were screened for outliers and standardized using the *scale()* function in R to avoid the influence of different levels of variation due to the unit of measurement of each variable. The mean value of each variable was compared for each farm type and the number of households per farm type were summarized per province and per year.

## 2.3 Yield gap decomposition

### 2.3.1 Concepts and definitions

Yield gap decomposition (Silva et al. 2017) relies on four yield levels to diagnose agronomic constraints in cropping systems at regional level (Doré et al. 1997). In addition to  $Y_w$  and  $Y_a$  (van Ittersum et al. 2013), the highest farmers' yield ( $Y_{HF}$ ) is defined as the average top 10th percentile of farmers' yields whereas the technically efficient yield ( $Y_{TEX}$ ) is defined as the maximum yield that can be achieved for a given input level in a well-defined biophysical environment. The efficiency yield gap refers to the difference between  $Y_{TEX}$  and  $Y_a$  and is explained by suboptimal crop management in relation to time, space and form of inputs applied. The resource yield gap refers to the difference between  $Y_{HF}$  and  $Y_{TEX}$  and is explained by suboptimal amounts of inputs applied. The technology yield gap refers to the difference between  $Y_w$  and  $Y_{HF}$  and is explained by low input use and the lack of use of specific technologies. The feasible yield ( $Y_f$ ) was also considered to unpack the contribution of suboptimal input use (i.e., resource yield gaps) and variety choice to the technology yield gap.  $Y_f$  is defined as the maximum yield with available technology and best-practice management but with no economic constraints (van Dijk et al. 2017).

### 2.3.2 Stochastic frontier analysis

Stochastic frontiers account for two random errors,  $v_{it}$  (random noise) and  $u_{it}$  (technical inefficiency), assumed to be independently distributed from each other when estimating production functions (Kumbhakar and Lovell 2000). A Cobb-Douglas functional form (Equation 1), comprising only first-order terms in the production frontier, was used to describe the relationship between maize yield and a vector of agronomic relevant variables defined according to principles of production ecology (van Ittersum and Rabbinge 1997). A translog functional form was also fitted to test the effect of second-order terms (i.e., squared and interactions) on maize yield. The results of the translog functional form are presented in Supplementary Material given the large number of estimated parameters (Supplementary Table 3). Inefficiency effects, i.e., the drivers of the efficiency yield gap, were also estimated through a one-step estimation of the production frontier and the second-stage regression (Equation 2; Battese and Coelli 1995), as follows:

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

$$u_{it} = \sum_j^J \delta_j \ln z_{jit} + \epsilon_{it} \quad (2)$$

$$v_{it} \sim N(0, \sigma_v^2) \quad (3)$$

$$u_{it} \sim N^+ \left( \sum_j^J \delta_j \ln z_{jit}, \sigma_u^2 \right) \quad (4)$$

$$\text{Eff. } Y_{g_{it}} = 1 - \exp(-u_{it}) \quad (5)$$

$$Y_{\text{TEX}_{it}} = y_{it} \times \exp(-u_{it})^{-1} \quad (6)$$

where  $y_{it}$  represents the maize yield in field  $i$  and in year  $t$ ,  $x_{kit}$  is a vector of agronomic inputs  $k$  used on field  $i$  and year  $t$  and,  $\alpha_0$  and  $\beta_k$  are parameters to be estimated. The vector  $z_{jit}$  comprises the  $j$  crop management drivers of the efficiency yield gap in field  $i$  and in year  $t$ .  $Y_{\text{TEX}}$  and  $Y_f$  were estimated for each field using the Cobb-Douglas model described earlier (Equations 1 and 6), but without considering inefficiency effects. Model parameters were estimated for the pooled data and for each farm type with maximum likelihood using the *sfa()* function of the R package *frontier* (Coelli and Henningsen 2013). Continuous variables were  $\ln$ -transformed prior to the analysis and data were used as a cross-section rather than as a panel, hence technological change and time-(in)variant technical efficiency were not assessed.

The vector of inputs  $x_{kit}$  was designed to capture the effect of growth-defining, growth-limiting, and growth-reducing factors on maize yield (Silva et al. 2017). Growth-defining factors were controlled for with the

following variables: growing degrees day considering a base temperature of 0 °C (Van Wart et al. 2013), temperature seasonality defined as the standard deviation of average monthly temperatures (Van Wart et al. 2013), seed rates ( $\text{kg ha}^{-1}$ ), replanting (yes or no), and variety type (open-pollinated, hybrid, or unknown). Growth-limiting factors related to water included variety classification according to drought tolerance (yes, no, or unknown), aridity index defined as the ratio between total annual precipitation and annual total potential evapotranspiration (Van Wart et al. 2013), soil rooting depth and soil available water (Leenaars et al. 2015), soil texture class constructed based on spatial predictions of clay, silt, and sand contents (Hengl et al. 2017), location of the field in a wetland (yes or no), and presence of erosion or flood control practices (yes or no). Growth-limiting factors related to nutrients included the rate of N applied ( $\text{kg N ha}^{-1}$ ), pH in water, and exchangeable acidity (Hengl et al. 2017). Finally, growth-reducing factors were captured with the number of weeding operations (none or one, two, and three or more), herbicide use (yes or no), and insecticide use (yes or no). Sowing date, expressed in weeks after the onset of the rains, and date of the first weeding operation, expressed in weeks after sowing, were included in the model as inefficiency effects. The variance inflation factors indicated no multicollinearity between the considered variables.

The Cobb-Douglas frontier model without inefficiency effects was used to predict Yf for specific values of some of the input variables. To do so, seed rate was set at  $25 \text{ kg ha}^{-1}$ , which is the recommended seed rate for maize in Zambia. N application rate was set at  $350 \text{ kg N ha}^{-1}$ , which is the minimum N requirement for a target of 80% of Yw in the high rainfall areas of Zambia ([www.yieldgap.org](http://www.yieldgap.org)). It was further assumed that drought tolerant hybrid maize varieties were used in combination with replanting of maize seedlings, herbicides, and insecticides. The estimation of Yf further assumed that fields with a pH in water below 6.5 were corrected to a pH in water of 6.5 and that fields with exchangeable acidity above  $0.2 \text{ cmol+ kg}^{-1}$  were corrected to that level in fields with pH below 6.5.

### 2.3.3 Distribution of actual yields

Farmers' fields were categorized as highest, average, and lowest yielding fields based on the distribution of Ya observed for a given variety type and climate zone x soil type combination. Highest yielding fields were identified as those with Ya above the 90th percentile. Average yielding fields were identified as those with Ya between the 10th and the 90th percentiles and lowest yielding fields as those with Ya below the 10th percentile. Highest ( $Y_{HF}$ ), average ( $Y_{AF}$ ) and lowest farmers' yields ( $Y_{LF}$ ) were calculated as the average Ya for the fields in each respective group. The

field classification was specific to each of three variety types and to each unique climate zone (Van Wart et al. 2013) and soil type (Hengl et al. 2015), so genotype and biophysical factors were controlled for when comparing maize yields and management practices across the different fields.

### 2.3.4 Global Yield Gap Atlas (GYGA)

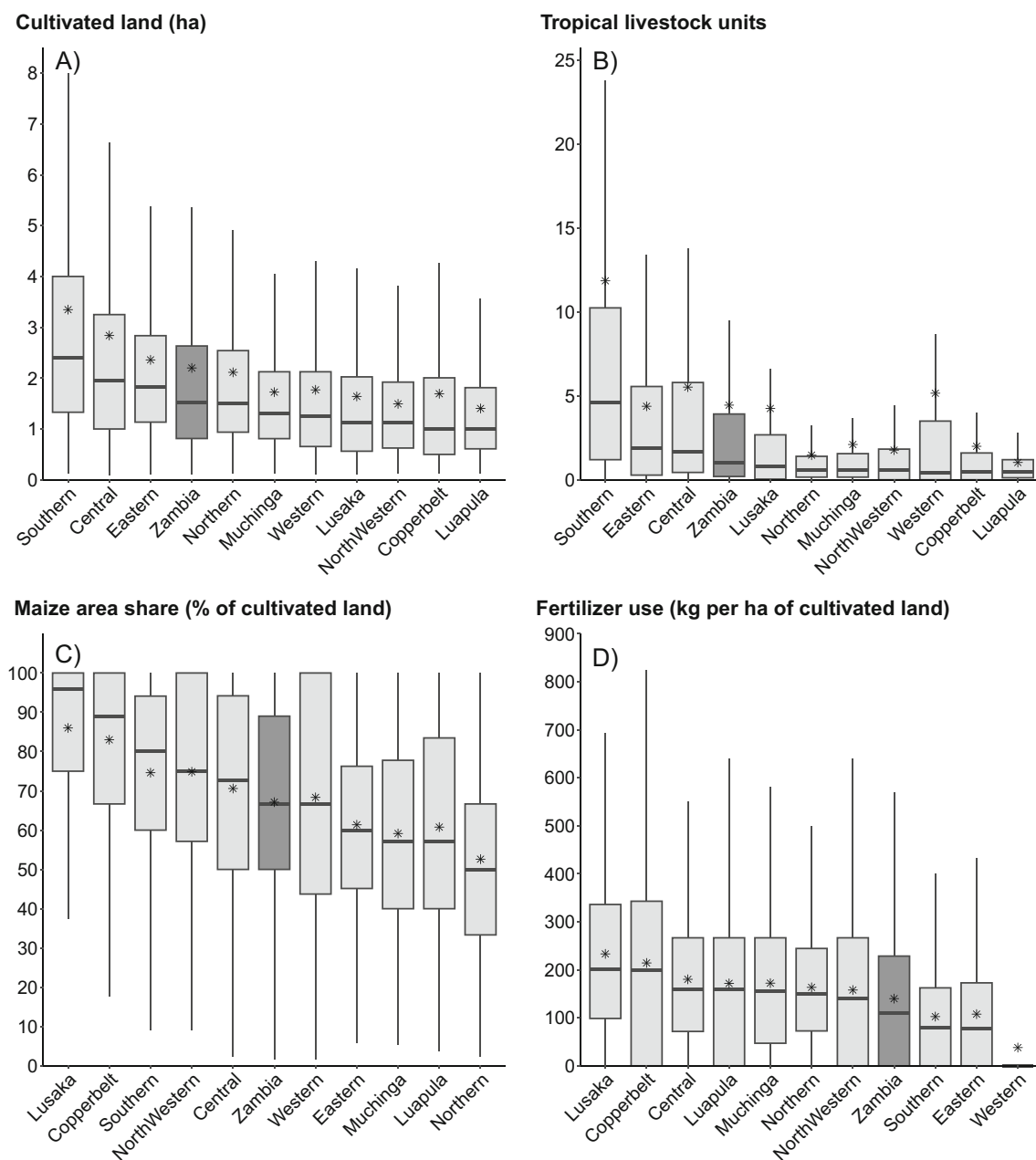
Yw for rain-fed maize across Zambia was obtained from GYGA. Maize Yw in Zambia was simulated with the HybridMaize crop model (Yang et al. 2004) for the period 2001–2010 (see [www.yieldgap.org/Zambia](http://www.yieldgap.org/Zambia) for further details). The average Yw data over the period 2001–2010 for a given climate zone was used here to benchmark Ya in farmers' fields and the technology yield gap was then calculated as the difference between Yw and  $Y_{HF}$  for unique climate zone x soil type x variety combinations. It was not possible to make use of year-specific Yw data for the same growing seasons in which the surveys were conducted due to lack of Yw data for the growing seasons surveyed, which introduces uncertainties in the magnitude of the overall yield gap estimated, particularly in regions with erratic rainfall. Therefore, coefficients of variation of maize Yw were computed to better characterize inter-annual yield variability across Zambia. The N rates needed to reach 80% of Yw were also retrieved from GYGA (ten Berge et al. 2019) to benchmark N used in farmers' fields.

## 3 Results

### 3.1 Maize-based farming systems in Zambia

Rural agricultural households across Zambia cultivate on average 2.2 ha of land and own 4.5 tropical livestock units (TLU; Figure 3A and B). Yet, the median values were considerably lower with 50% of the surveyed households cultivating less than 1.6 ha and owning less than 1.1 TLU. Maize was cultivated throughout the country with an average and median maize area share of 67% of the total cultivated (Figure 3C). This corresponds to an average maize area per farm of about 1.4 ha. Fertilizer use across the country was on average  $140 \text{ kg ha}^{-1}$  of cultivated land, with 50% of the surveyed farms using less than 110 kg of fertilizer per ha of cultivated land across the three survey periods (Figure 3D).

There were wide variations in total cultivated land, livestock ownership, maize share of cultivated cropland, and total fertilizer use across the different provinces (Figure 3 and Supplementary Table 1). The average total cultivated land was larger than the national average in the Southern (3.4 ha), Central (2.8 ha), and Eastern provinces (2.4 ha),



**Fig. 3** Main characteristics of farming systems in Zambia and their variability at national level and per province: (A) cultivated land in ha, (B) livestock ownership in tropical livestock units, (C) proportion of the cultivated land occupied by maize in %, and (D) fertilizer used

per ha of cultivated land. Data for the entire country are highlighted in dark gray. Asterisks show the mean value across the farm-year combinations of each province.

and lower in all other provinces (1.4–2.1 ha; Figure 3A). The same was true for livestock ownership which was on average 11.9, 5.5, and 4.4 TLU in the Southern, Central, and Eastern provinces, respectively, and much lower in all other provinces, notably those in the Northern part of the country (Figure 3B). Maize represented more than 50% of the cultivated land for at least 50% the surveyed farms in all provinces (Figure 3C). The average maize share of cultivated cropland was above 80% in the provinces

of Lusaka and Copperbelt, between 70 and 75% in the Southern, Northwestern, and Central provinces, and about 60% in the Eastern, Muchinga, and Luapula provinces. The Northern province was where the maize share of cultivated cropland was lowest, ca. 55% of the total cultivated land. Finally, fertilizer use was below the national average in the Southern, Eastern, and Western provinces (50–100 kg ha<sup>-1</sup>), and slightly above the national average in the other provinces (Figure 3D).

**Table 1** Maize yield potential (Yp) and water-limited yield (Yw) for eight weather stations located across Zambia. Means and coefficients of variation (CV) are provided for the years 2000–2010. ‘N requirements’ refer to the minimum N rates needed to reach 80% of Yw in

each location, averaged over the same period as the yield ceilings. Source: [www.yieldgap.org](http://www.yieldgap.org); van Ittersum et al. (2016); ten Berge et al. (2019).

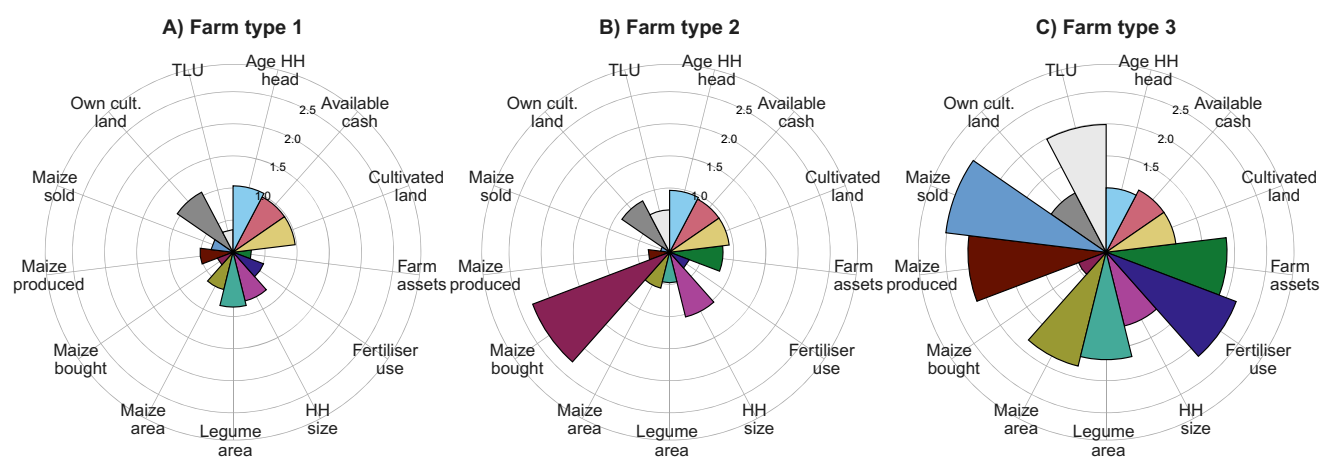
Province	Weather station	Mean Yp (t ha <sup>-1</sup> )	CV Yp (%)	Mean Yw (t ha <sup>-1</sup> )	CV Yw (%)	N requirements (kg N ha <sup>-1</sup> )
Northern	Kasama	18.71	0.05	18.59	0.05	325.1
Luapula	Mansa	18.34	0.07	17.36	0.09	303.6
Muchinga	Mpika	16.78	0.05	14.48	0.21	253.1
Eastern	Chipata	16.56	0.06	13.29	0.31	232.3
Central	Kabwe	16.79	0.05	12.63	0.35	220.7
Central	Mumbwa	16.25	0.08	10.03	0.57	175.4
Western	Mongu	16.80	0.07	9.79	0.47	171.2
Southern	Choma	13.28	0.06	9.38	0.40	164.0

### 3.2 Farm types and importance of maize

The farm typology was constructed using principal component analysis (PCA) followed by hierarchical clustering (HC). Four principal components had an eigenvalue greater than one and were retained for further analysis. These four principal components explained approximately 60% of the cumulative variance in the data. Three clusters were identified in the dissimilarity dendrogram of the HC analysis, corresponding to three distinct farm types. In short, Farm Type 1 (FT1) exhibited a low dependency on maize production and consumption, Farm Type 2 (FT2) were net buyers of maize and exhibited low levels of maize area and production, and Farm Type 3 (FT3) were market-oriented maize producers engaged in agricultural activities, as indicated by

the large number of livestock kept and large amount of fertilizer used (Figure 4 and Supplementary Table 2).

The age of the household head did not vary significantly across farm types (Figure 4) whereas household size was lower for FT1 (5.5 individuals), intermediate for FT2 (7.2 individuals), and higher for FT3 (8.2 individuals). FT1 owned 1.5 TLU and cultivated a total of 1.4 ha, 0.8 ha of which were allocated to maize and 0.3 ha to legumes, and used 140 kg of fertilizer per farm per year. FT1 produced an average of 1500 kg of maize, sold 600 kg of maize, and bought 50 kg of maize per farm per year. FT2 had access to 2.7 TLU and cultivated a total of 1.3 ha, of which 0.8 and 0.1 ha were cultivated with maize and legumes, respectively. Fertilizer use was lower in FT2 than in FT1 (Figure 4) with a rate of 90 kg fertilizer per farm per year, and so was maize



**Fig. 4** Radar charts represent all studied quantitative variables on individual axes starting from the same central point for each farm type. The variables displayed were used in the principal component analysis followed by hierarchical clustering to delineate the farm typology for the pooled data. Data are scaled with the average value of each

variable for all farm types (cf. Supplementary Table 2). The spatial and temporal distribution of the farm types is provided in Supplementary Figures 1 and 2, respectively. Abbreviations: ‘HH’ = household, ‘TLU’ = tropical livestock units.

production and maize sold (Figure 4), with an average of 1000 kg and 250 kg per farm per year, respectively. FT3 used 600 kg of fertilizer, produced 6500 kg of maize, sold 1600 kg of maize, and purchased 80 kg of maize per farm per year.

There were slight differences in the spatial distribution of the three farm types (Supplementary Figure 1). In Western province, nearly 70% of the farms were classified as FT2 and only 10% of the farms were classified as FT3. By contrast, in Southern and Central provinces as much as 50% of the farms were classified as FT3 whereas 20% and 30% were classified as FT1 and FT2, respectively. In Luapula, Muchinga, Northern, and Northwestern provinces, 35–40% of the farms were classified as either FT1 or FT3. Farms were evenly distributed amongst farm types (ca. 30% per farm type), in the Eastern and Copperbelt provinces. There were no major changes in farm type classification for single farms over time (Supplementary Figure 2): out of 5238 farm-year combinations, 715 were classified as FT3, 412 as FT2, and 209 as FT1 in the three rounds of the survey. Other changes in farm type classification were not consistent and were likely to reflect fluctuations in farm performance over time.

### 3.3 Yields and yield gaps of rain-fed maize

Maize  $Y_a$  across all farm-year combinations analyzed ranged between nil and  $9.0 \text{ t ha}^{-1}$  (Figure 5).  $Y_a$  was smaller and more variable in 2019 than in 2012 and 2015 harvest years (Figure 5A), with average values of 2.6, 2.4, and  $2.2 \text{ t ha}^{-1}$  and a coefficient of variation (CV) of 67, 67, and 77% during the 2012, 2015 and 2019 harvest years, respectively (Figure 5A). There were also clear differences in the distribution of  $Y_a$  across agro-ecological zones, farm types, and variety types.  $Y_a$  was smallest and most variable in agro-ecology IIb (mean =  $1.3 \text{ t ha}^{-1}$ , CV = 82%) and greatest and least variable in agro-ecology III ( $2.7 \text{ t ha}^{-1}$ , 61%), with intermediate values observed in agro-ecology IIa and I (Figure 5B).  $Y_a$  was also smallest and most variable for FT2 ( $1.8 \text{ t ha}^{-1}$ , 76%), intermediate for FT1 ( $2.4 \text{ t ha}^{-1}$ , 66%), and greatest and least variable for FT3 ( $2.9 \text{ t ha}^{-1}$ , 61%; Figure 5C). Finally,  $Y_a$  was on average 1.9 and  $2.9 \text{ t ha}^{-1}$ , with a CV of 61 and 73%, for open-pollinated and hybrid maize varieties, respectively (Figure 5D).

Simulated yield potential ( $Y_p$ ) ranged between 13 and  $19 \text{ t ha}^{-1}$  in the Southern and Northern provinces, respectively, without a clear spatial distribution across the country (Table 1). Conversely,  $Y_w$  was greatest and least variable in the Northern, Luapula, and Muchinga provinces, intermediate in the Eastern and Central provinces, and smallest and most variable in the Southern and Western provinces (Table 1).  $Y_w$  was on average  $18 \text{ t ha}^{-1}$  in the Northern province,  $13 \text{ t ha}^{-1}$  in the Eastern province,

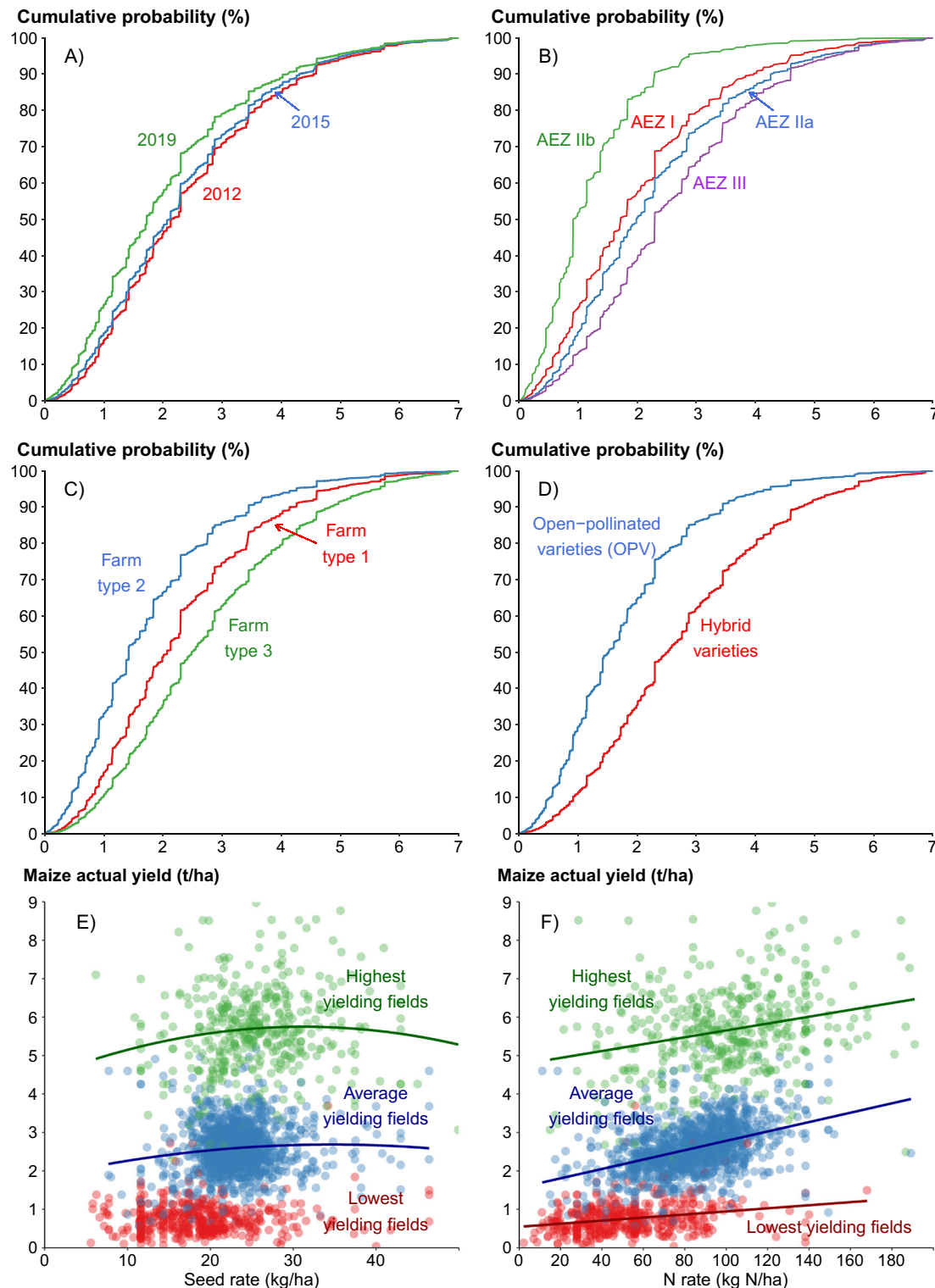
and  $9.5 \text{ t ha}^{-1}$  in the Western and Southern provinces. The respective CV for  $Y_w$  was 5, 30, and 45% for the Northern, Eastern, and Western and Southern provinces, respectively (Table 1). The difference between  $Y_p$  and  $Y_w$  indicates the yield gap due to water limitations, whose magnitude increased along a North-South gradient (Table 1) characterized by lower and more erratic rainfall (Figure 2). N rates needed to reach 80% of  $Y_w$  were greater than  $250 \text{ kg N ha}^{-1}$  in the Northern, Luapula, and Muchinga provinces, ca.  $230 \text{ kg N ha}^{-1}$  in the Eastern province, and about  $170 \text{ kg N ha}^{-1}$  in the Western and Southern provinces (Table 1).

Yield gap closure (i.e., the ratio between  $Y_a$  and  $Y_w$ ) was on average 21% of  $Y_w$  and varied with agro-ecological zone, province, and farm type (Figure 6). Yield gap closure was greatest in agro-ecology I (35% of  $Y_w$ ), intermediate in agro-ecology IIa (23% of  $Y_w$ ), and smallest in agro-ecologies IIb and III (15% of  $Y_w$ ; Figure 6A and B). Yield gap closure per province was similar to that per agro-ecology (Figure 6B and E) because most of the Southern province is in agro-ecology I, the Central and Eastern provinces are in agro-ecology IIa, the Western province is in agro-ecology IIb, and the Northern, Northwestern, Luapula, Muchinga and Copperbelt provinces are in agro-ecology III. Finally, yield gap closure was on average 30% of  $Y_w$  for FT3, 20% of  $Y_w$  for FT1, and only 15% of  $Y_w$  for FT2 (Figure 6C and F).

Most of the yield gap was attributed to the technology yield gap, which accounted for  $7.2 \text{ t ha}^{-1}$  (50% of  $Y_w$ ) on average, yet narrowing efficiency and resource yield gaps could more than double  $Y_a$  for maize in Zambia (Figure 6). The efficiency yield gap was on average  $1.6 \text{ t ha}^{-1}$  (14% of  $Y_w$ ) and the resource yield gap was on average  $1.7 \text{ t ha}^{-1}$  (16% of  $Y_w$ ), which means that fine tuning current crop management practices and increasing input use to the level of highest yielding fields can increase yields from the current  $2.4 \text{ t ha}^{-1}$  to  $5.7 \text{ t ha}^{-1}$ . The resource yield gap considering the feasible yield (i.e., maximum yield with available technology and best-practice management but with no economic constraints) as ceiling was small with an average of  $1.0 \text{ t ha}^{-1}$  (7% of  $Y_w$ ). This means that resource-use efficiency in farmers' fields is low and must be improved to realize the yield gains associated with increased input use and better technology. The large technology yield gap is thus a result of suboptimal input use compared to what is needed to reach  $Y_w$  and of low resource-use efficiency of current farm practices.

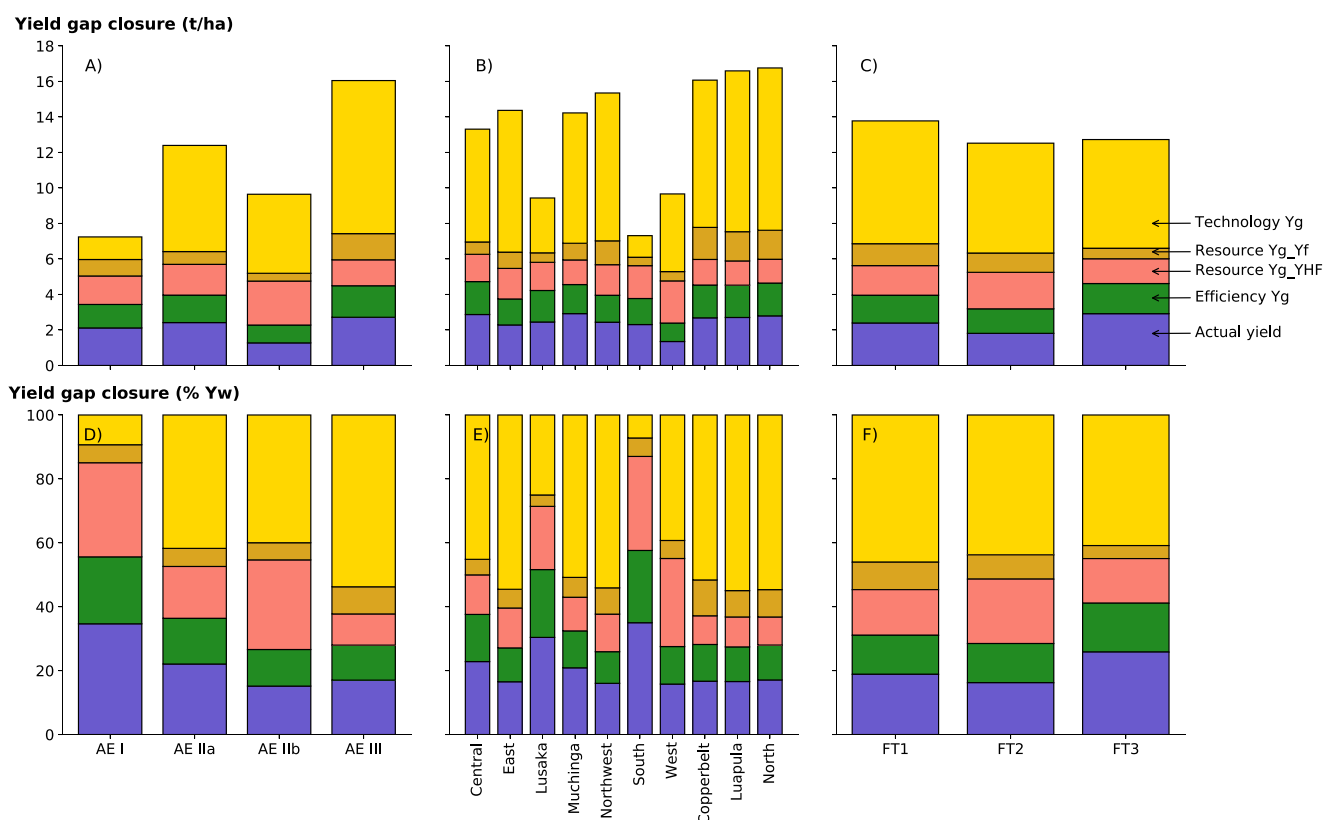
There were slight differences between agro-ecological zones and provinces in the relative contribution of each yield gap to the overall yield gap (Figure 6). For instance, the relative contribution of the technology yield gap to the total yield gap was less than 10% of  $Y_w$  in the Southern province (which is part of agro-ecological zone I; Figure 6D and





**Fig. 5** Maize actual yield variability across years (A), agro-ecology zones (AEZ, B), farm types (C), and variety types (D), and maize yield response to seed rate (E) and N applied (F). Lines in (A)–(D) display empirical cumulative distribution functions. Mean values (and coefficients of variation) are as follows: 2.6 t ha<sup>-1</sup> (67.0%) for year 2012; 2.4 t ha<sup>-1</sup> (66.6%) for year 2015; 2.2 t ha<sup>-1</sup> (76.8%) for year 2019; 2.1 t ha<sup>-1</sup> (74.0%) for AEZ I; 2.4 t ha<sup>-1</sup> (70.3%) for AEZ IIa; 1.1 t ha<sup>-1</sup> (82.0%) for AEZ IIb; 2.7 t ha<sup>-1</sup> (61.4%) for AEZ III; 2.4 t ha<sup>-1</sup>

(66.2%) for farm type 1; 1.8 t ha<sup>-1</sup> (76.2%) for farm type 2; 2.9 t ha<sup>-1</sup> (61.3%) for farm type 3; 2.9 t ha<sup>-1</sup> (60.9%) for hybrid varieties; 1.9 t ha<sup>-1</sup> (73.3%) for open-pollinated varieties. Data in (E) and (F) are aggregated per household × field type, and lines display statistically significant ordinary-least square regressions fitted to highest (Y<sub>HF</sub>), average (Y<sub>AF</sub>), and lowest yielding fields (Y<sub>LF</sub>, quadratic for seed rate and linear for N).



**Fig. 6** Maize yields and yield gaps in Zambia disaggregated by agro-ecological zones (A-D), provinces (B-E), and farm types (C-F). Panels in the top row display data in absolute terms ( $\text{t ha}^{-1}$ ) and panels in the bottom row display data in relative terms (% of  $Y_w$ ). Codes: 'AE' = agro-ecological zone, 'FT' = farm type, 'Efficiency Yg' = efficiency

yield gap, 'Resource  $Yg_{YHF}$ ' = resource yield gap considering the highest farmers' yields ( $Y_{HF}$ ) as benchmark, 'Resource  $Yg_{Yf}$ ' = resource yield gap considering the feasible yield ( $Y_f$ ) as benchmark, 'Technology Yg' = technology yield gap.

E), whereas the relative contribution of the efficiency and resource yield gaps were ca. 20% and 30% of  $Y_w$ . In Lusaka province (with areas also part of agro-ecological zone I), each of the three intermediate yield gaps accounted for ca. 20% of the total yield gap. The differences in the relative of contribution of the efficiency, resource, and technology yield gaps to the overall yield gap between these two provinces (Southern and Lusaka) and the other provinces is likely attributed to the low water-limited yield simulated, and hence small technology yield gap in absolute terms, for the Southern and Lusaka provinces (and respective agro-ecological zone, Figure 6A and B). There were also slightly differences in the causes of yield gaps for the different farm types (Figure 6C and F): the efficiency yield gap was slightly greater for FT3 (i.e., market-oriented maize farms) than for FT1 and FT2, whereas the opposite was true for the resource yield gap (Figure 6C and F).

### 3.4 Determinants of maize yield variability

The stochastic frontier model fitted to the pooled data revealed that seed rate, variety type, aridity index, soil

available water, and herbicide use were the key drivers of maize yield variability (Table 2). The seed rate had a significant positive effect on  $Y_a$  with a 1% increase in seed rate resulting in 0.33% increase in  $Y_a$ . There was also a significant effect of variety on  $Y_a$ , with hybrid varieties yielding ca. 13% more than open-pollinated varieties. The effects of temperature seasonality and replanting on  $Y_a$  were also statistically significant, but the effect was small. Aridity index and soil available water had a significant positive effect on  $Y_a$  with a 1% increase in these variables resulting into 0.50 and 0.20% increase in  $Y_a$ .  $Y_a$  in loamy sand soils were significantly lower (135%) than in clay soils and adoption of erosion and flood control practices increased  $Y_a$  by 5%. N applied had a significant positive effect on  $Y_a$  whereas exchangeable acidity had a significant negative effect on  $Y_a$ , but in both cases the effect was small. Herbicide use had a significant positive effect on  $Y_a$ , resulting in 12.5% greater  $Y_a$  compared to fields where herbicides were not used. Finally,  $Y_a$  was significantly lower in 2015 and in 2019 than in 2012 (cf. Figure 5A). The time of the first weeding, measured in number of days after sowing, had a significant negative effect on the efficiency

**Table 2** Parameter estimates of the stochastic frontier model fitted for maize yield in Zambia during the growing seasons of 2010/11, 2013/14, and 2017/18. The same model was fitted to the pooled sample (Zambia) and each of the farm types identified (Figure 4). Reference values: Year = '2012', Replant = 'No', Variety = 'OPV', Drought tolerant = 'No', Soil = 'Clay', Wetland = 'No', Erosion/Flood = 'No', Weeding = 'One or none', Herbicide use = 'No', Insecticide use = 'No'. Units: WFO = week from onset of rains; WAS = week after sowing. Significance is indicated by the codes: '\*\*\*' 0.1%, '\*\*' 1%, '\*' 5%, '#' 10%. n.a. = not applicable.

	Zambia	Farm type 1	Farm type 2	Farm type 3
<b>Production frontier</b>				
Intercept	2.196	-6.079 #	10.016*	4.120
Year_2015	-0.080***	-0.049	-0.102**	-0.087***
Year_2019	-0.271***	-0.193***	-0.280***	-0.265***
<b>Defining factors</b>				
Growing degrees day	-0.108	0.538 #	-0.647 #	-0.101
Temperature seasonality	0.093**	0.277***	-0.127 #	-0.045
Seed rate (kg ha <sup>-1</sup> )	0.333***	0.328***	0.307***	0.462***
Replant_Yes	-0.064***	-0.023	-0.009	-0.112***
Variety_Hybrid	0.128***	0.134**	0.084	0.105*
Variety_Unknown	0.065***	0.050	0.028	0.109**
<b>Limiting factors (water)</b>				
Drought tolerant_Yes	0.028	-0.014	0.043	0.045 #
Drought tolerant_Unknown	-0.120***	-0.058	-0.139**	-0.105**
Aridity index	0.502***	0.613***	0.242*	0.443***
Rooting depth	0.021	-0.003	0.011	0.019
Soil available water	0.214***	0.162*	0.261***	0.164***
Soil_Clay loam	-0.014	0.084	0.092	-0.117
Soil_Loam	0.234	0.032	1.265***	-0.510*
Soil_Loamy sand	-1.353*		-1.161 #	
Soil_Sandy clay	-0.026	-0.028	0.161	-0.099
Soil_Sandy clay loam	0.096	0.114	0.218 #	0.001
Soil_Sandy loam	0.110	0.150	0.164	0.037
Wetland_Yes	-0.035	-0.023	-0.073 #	0.013
Erosion/Flood_Yes	0.052**	-0.048	0.106**	0.043 #
<b>Limiting factors (nutrients)</b>				
N applied (kg N ha <sup>-1</sup> )	0.026***	0.021***	0.016***	0.024***
pH in H <sub>2</sub> O (unitless)	0.176	0.312	0.618	-0.134
Exch. acidity (cmol+ kg <sup>-1</sup> )	-0.018***	-0.010	-0.016 #	-0.003
<b>Reducing factors</b>				
Weeding_2	0.021	0.006	0.103***	-0.012
Weeding_3+	0.018	0.153**	0.057	-0.085*
Herbicide_Yes	0.126***	0.156 #	0.034	0.076*
Insecticide_Yes	0.087 #	0.181 #	0.055	0.018
<b>Inefficiency effects</b>				
Sowing date (WFO)	0.008 #	0.018*	0.016 #	0.000
Weeding timing (WAS)	-0.052***	-0.096**	-0.163***	-0.006
<b>Model evaluation</b>				
$\sigma^2 = \sigma_v^2 + \sigma_u^2$	0.964***	0.885***	1.363***	0.642***
$\gamma = \sigma_u^2 / \sigma^2$	0.820***	0.819***	0.869***	0.735***
<b>Sample size (n)</b>				
Field x year combinations (#)	30765	8245	10896	11335

yield gap, meaning that smaller efficiency yield gaps were observed when the first weeding was done at later dates, but again the effect was small.

The significance level and magnitude of the first-order terms derived from the survey data were comparable in both the Cobb-Douglas and translog stochastic frontier

models (Supplementary Table 3). Yet, variables derived from secondary sources (temperature seasonality, aridity index, rooting depth, soil available water, pH in water, and exchangeable acidity) showed contrasting signs and different effect sizes (Supplementary Table 3). Quadratic terms were statistically significant for all continuous variables, except soil available water (Supplementary Table 3), indicating a quadratic effect of seed rate on Ya and a quadratic positive effect of N applied on Ya (cf. Figure 5E and F). There were negative interactions between seed rate and growing degree days, aridity index and N applied, and positive interactions between seed rate and temperature seasonality and pH in water. N applied showed a negative interaction with growing degree days, seed rate, rooting depth and soil available water, meaning that maize yield response to N decreased with increases in these variables.

The effect of seed rate and N applied on maize yield was further investigated for highest, average, and lowest yielding fields. Maize yield ranged between 0 and 1.5 t ha<sup>-1</sup>, 1.5 and 4.0 t ha<sup>-1</sup>, and 4.0 and 9.0 t ha<sup>-1</sup> for lowest, average, and highest yielding fields (Figure 5E and F). Seed and N rates were lowest in lowest yielding fields (16 kg ha<sup>-1</sup> and 54 kg N ha<sup>-1</sup>), intermediate for average yielding fields (23 kg ha<sup>-1</sup> and 84 kg N ha<sup>-1</sup>), and greatest for highest yielding fields (25 kg ha<sup>-1</sup> and 100 kg N ha<sup>-1</sup>). There were no major differences in yield and input use for the different farm types across highest, average, and lowest yielding fields (data not shown). The quadratic effect of seed rate on yield was significant for highest and average yielding fields, but not for lowest yielding fields (Figure 5E), whereas the effect of N applied on yield was linear and positive for lowest, average, and highest yielding fields (Figure 5F). Yield response to N was greatest, intermediate, and smallest for average, highest, and lowest yielding fields, respectively.

The drivers of maize yield variability for each farm type were largely comparable to those observed for the pooled data (Table 2), as opposed to the results obtained for Northern, Eastern, and Southern provinces (Supplementary Table 4). For all farm types, seed rate, aridity index, soil available water, and N applied had a significant positive effect on Ya and Ya was significantly smaller in 2019 than in 2012. Variety type and herbicide use had a positive effect on Ya for FT1 and FT3, and fields weeded three or more times yielded 15% more for FT1, and 9% less for FT3, than fields weeded once or not weeded. Increasing temperature seasonality by 1% translated into increases in Ya of 28% for FT1, replanted fields yielded 11% less than non-replanted fields for FT3, and fields where erosion or flood control practices were adopted for FT2 had 11% greater Ya than fields where these practices were not adopted. Also for FT2, fields weeded twice yielded 10% more than fields with one or no weeding operations. The effects of soil type on Ya were not consistent across farm types. The seed rate

and N applied had a significant positive effect of maize, and a similar effect size, independently of the province (Supplementary Table 4) and the effect of biophysical variables (e.g., aridity index and soil available water) was not significant when the model was fitted per province (Supplementary Table 4).

## 4 Discussion

Agricultural productivity must increase in sub-Saharan Africa with a view towards improved food security and reduced food imports with minimum crop expansion in biodiversity and carbon-rich natural habitats (e.g., Giller et al. 2021a; Jayne and Sanchez 2021; Giller 2020; Keating et al. 2014). Zambia is no exception to this narrative (Figure 1), where narrowing yield gaps up to 80% of Yw is needed for the country to reach cereal self-sufficiency by 2050 with cropland expansion (van Ittersum et al. 2016). Yield gap closure for rain-fed maize across Zambia is only ca. 20% of Yw (Figure 6), which is similar for other crops in other countries across sub-Saharan Africa (van Ittersum et al. 2016; Tittonell and Giller 2013). The large yield gap of rain-fed maize in Zambia is mostly attributed to the technology yield gap (Figure 6) indicating that more efficient production methods are needed to narrow maize yield gaps. Yet, narrowing efficiency and resource yield gaps through fine tuning current farm practices could more than double current yields (Figure 6). The latter can be achieved through improved timeliness and precision of management operations and through increases in input use to levels observed in highest yielding fields (Figures 5E and 5F). Similar findings regarding the relative importance of efficiency, resource, and technology yield gaps were reported for cereal farming systems in Eastern Africa (Silva et al. 2019, 2021; Assefa et al. 2020; van Dijk et al. 2017), pointing to the need for making inputs available to farmers at the right amount, cost, and time, and of targeting and packaging technologies in ways that increase adoption at farm level.

Seed and N rates, variety, weed control, and sowing date were the most important management drivers of maize yield variability in Zambia (Table 2). All these are well-known drivers of maize yield variability in Eastern and Southern Africa (e.g., Burke et al. 2020; Assefa et al. 2020). First, seed rate and variety type had a large impact on maize yield, with a 1% increase in seed rate resulting ca. 0.35% increase in maize yield and hybrid varieties yielding 12% more than traditional OPVs (Table 2). Seed rate might well be a proxy for plant population, a key factor controlling maize productivity in Southern Africa (Nyagumbo et al. under review). Second, the timing of the first weeding operation was an important driver of the efficiency yield

gap (Table 2), reflecting the importance of timely weeding at the start of the growing season for maize productivity. Third, N fertilizer rate had a linear positive effect on maize yield (Figure 5F; Table 2), but the effect size was small due to the low amounts of N applied by farmers. In fact, the range of N application rates observed in farmers' fields was considerably lower than that needed to reach 80% of Yw (i.e., 170–320 kg N ha<sup>-1</sup>; Table 1). Such large N application rates are out of reach for most smallholders in the country, and may well not be profitable or desirable under prevailing conditions (e.g., input-output markets, infrastructure, and soil acidity). Lastly, the effect of timely sowing on maize productivity was very much related to the onset of the rains (Supplementary Figures 3 and 4), and appropriate-scale mechanization can contribute to timely and more precise sowing across the region (Baudron et al. 2015).

The drivers of maize yield variability were largely consistent across farm types (Table 2), but the importance of maize for rural livelihoods across Zambia was farm-type specific (Figure 4). This means that interventions aiming to narrow maize yield gaps will likely benefit the different farm types differently. For instance, boosting maize productivity can be a suitable 'stepping up' strategy for market-oriented maize farms (FT3), who achieve the highest maize yields in Zambia (Figure 6C). Targeting interventions to this type of farm might well be the most effective way to increase maize production at national level. Conversely, farms with low levels of assets (FT1 and FT2, Figure 4), for whom 'stepping out' of maize production through investments in new on-farm activities or off-farm activities is likely more suitable, do not seem to have the productive capacity to intensify maize production in the short-term. Yet, increasing maize yields would be more beneficial for FT2 than for FT1 given the large dependency on bought maize of the former (Figure 4). Clearly, strategies aiming to narrow maize yield gaps must thus be complemented with a suite of pro-poor policies and investments tailored to specific farm types. This will be crucial to stimulate and embed smallholder agriculture into a broader rural development program that can provide social safety nets in the absence of livelihood options off-farm (Giller et al. 2021a).

Maize production in Zambia takes place across a gradient of agro-ecological conditions, which in turn have a considerable impact on yield gaps and their causes throughout the country (Figure 6; Supplementary Table 4). For instance, our analysis indicates that a 1% increase in soil available water translates into ca. 0.20% greater maize yield and that a 1% decrease in exchangeable acidity results into a 0.02% increase in maize yield across the pooled sample (Supplementary Table 4). Water is indeed a key limiting factor to production in the semi-arid areas of Southern and Western Zambia (Table 1, Figure 2;

Ngoma et al. 2021) whereas soil acidity is known to be a major constraint to agricultural production in the humid areas of Northern Zambia (Pelletier et al. 2020; Burke et al. 2017; Pauw 1994). These biophysical constraints may impact the adoption of mineral fertilizers to narrow resource yield gaps due to the risks involved in areas with low and erratic rainfall and the low nutrient-use efficiency in areas with acid soils, both with implications beyond maize farming in Zambia. Erratic rainfall is widespread across much of Eastern and Southern Africa (Muthoni et al. 2019) whereas soil acidity (defined here as low pH areas with high levels of exchangeable acidity) affects over half of all countries in sub-Saharan Africa (Silva et al., in preparation). These results support the revision of the subsidy program by the Government of Zambia (Morgan et al. 2019) to make it possible for farmers to access mechanized services and inputs (e.g., seeds, fertilizers, and lime) and to strengthen extension systems to deliver timely and site-specific agronomic recommendations (Jayne et al. 2018). This is crucial to improve soil health and sustainably intensify maize production in the country.

Further research is needed to understand how fertilizer use is influenced by climate variability and to identify profitable soil water conservation technologies for semi-arid areas. A range of new technologies building on previous conservation agriculture research (e.g., improved legume systems with strip-, double, relay and intercropping, green manure cover crops, and agroforestry species) are currently being tested on-farm in Zambia to address these challenges. For humid areas, it is crucial to revisit past research on soil acidity to assess the returns-on-investment associated with liming or acid soil management strategies (CIMMYT 2021; Burke et al. 2017). Simulated yield ceilings across the continent, and respective N rates needed to reach such yields (Table 1; van Ittersum et al., 2016), should also be thoroughly tested against empirical data as they are well above maximum yields reported in agronomic experiments under controlled conditions (see Masuka et al. 2017; Mupangwa et al. 2017 for examples in Zambia).

High rainfall variability makes rain-fed farming across Eastern and Southern Africa a risky activity for smallholders. Site-specific recommendations must thus consider year-to-year variation in profitability and smallholders' risk profile to cope with uncertain yield response to inputs (Descheemaeker et al. 2016), as these are known to constrain farmers' willingness to investment in technologies. More attention must be paid to incorporate the effects of rainfall variability and soil properties on yield response to inputs to better explain the adoption of technologies (Chamberlin et al. 2021; Burke et al. 2017), which appear to be profitable on average, but have high variance in outcomes over time. The role of non-information constraints, such as

alternative uses of labor at critical periods (Silva et al. 2019; Kamanga et al. 2014), to the adoption of improved crop management practices also needs to be explored as these can limit the timely management needed to narrow yield gaps. Small farm sizes are another important constraint to technology adoption and intensification of crop production in African smallholder farming systems (Harris and Orr 2014), as narrowing yield gaps on small farms is often not enough to ensure food self-sufficiency or a living income at household level (Giller et al. 2021b).

## 5 Conclusion

Maize is the dominant crop in Zambian farming systems, which range from mixed-crop livestock systems in semi-arid areas of the Southern and Western provinces to mixed maize systems in the rest of the country. This study combined for the first time a farm typology delineation with yield gap decomposition to gain insights on what interventions are needed, where, and for which farm types, to increase maize production in Zambia. Three farm types were identified, including households for which maize is a marginal crop, households which are net buyers of maize, and households which are market-oriented maize producers. Maize yield gap closure across the country was only 20% of the water-limited yield ( $Y_w$ ), corresponding to  $2.4 \text{ t ha}^{-1}$ , and was slightly larger for market-oriented maize farms. For nearly all agro-ecological regions, provinces, and farm types, about half of the yield gap was attributed to current technologies used by farmers not reaching their full agronomic potential. Yet, improving current technologies in terms of timeliness and precision of operations and increasing input use, particularly mineral fertilizers, could more than double current yields. Doing so requires targeted approaches for technology intervention, e.g., by focusing on market-oriented maize producers, accompanied by carefully designed policy interventions ensuring other households benefit from other value chains or off-farm opportunities. If profitable, adoption of practices that increase soil moisture in semi-arid areas, such as conservation agriculture, and management of soil acidity in humid areas are key to improve yield response to mineral fertilizers. Two avenues can facilitate the foregoing policy levers. First, the current national subsidy program needs to be flexible enough to make it possible for farmers to access mechanized services and inputs. Second, the extension systems need to be strengthened to help farmers cope with risk and uncertain crop yield response to inputs in areas with high rainfall variability. Further research is needed to better understand the profitability of maize production under rain-fed conditions and to disseminate technologies that can reduce the vulnerability of farmers

to inter-annual rainfall variability. Blanket, one-size-fits-all, recommendations should be avoided when promoting sustainable intensification practices aiming to increase yields in the country.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s13593-023-00872-1>.

**Acknowledgments** We thank the Indaba Agricultural Policy Research Institute (IAPRI) for having availed the three waves of panel data to be used to conduct this research and dr. Antony Chapoto for his constructive suggestions in an earlier version of the manuscript. The manuscript reflects the opinions of the authors but do not necessary represent the institutional policies, opinions and strategies of the European Union, FAO and CIMMYT.

**Authors' contributions** Conceptualization: JVS, FB, CT. Software: JVS, FB, HN. Validation: HN, IN, ES, KK. Formal analysis: JVS. Investigation: JVS. Resources: FB, HN, IN, MH, MM, CT. Data curation: JVS, HN. Writing—original draft preparation: JVS. Writing—review and editing: JVS, FB, HN, IN, CT. Supervision: FB, CT. Project administration: MH, ES, KK, MM, CT. Funding acquisition: MM, CT.

**Funding** This research was conducted under the European Union-funded “Sustainable Intensification of Smallholder Farming Systems in Zambia project” (SIFAZ, Grant No FED/2019/400-893), jointly implemented by the Ministry of Agriculture of Zambia, the Food and Agriculture Organization of the United Nation (FAO), and the International Maize and Wheat Improvement Center (CIMMYT). Financial and logistical support is gratefully acknowledged.

**Data availability** The datasets generated during and/or analyzed during the current study are not publicly available due privacy reasons but are available from the corresponding author on reasonable request.

**Code availability** The R and Python scripts used in the analysis are available upon request at the GitHub repository [https://github.com/jvasco323/SIFAZ\\_Yg\\_Decomposition](https://github.com/jvasco323/SIFAZ_Yg_Decomposition).

## Declarations

**Ethics approval** Not applicable

**Consent to participate** Not applicable

**Consent for publication** Not applicable

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alvarez S, Timler CJ, Michalscheck M, Paas W, Descheemaeker K, Tiftonell P, Andersson JA, Groot JCJ (2018) Capturing farm diversity with hypothesis based typologies: an innovative methodological framework for farming system typology development. *PLoS One* 13:e0194757. <https://doi.org/10.1371/journal.pone.0194757>
- Arslan A, McCarthy N, Lipper L, Asfaw S, Cattaneo A (2014) Adoption and intensity of adoption of conservation farming practices in Zambia. *Agric Ecosyst Environ* 187:72–86. <https://doi.org/10.1007/s00704-007-0306-4>
- 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* 20:621–332. <https://doi.org/10.1007/BF01205442>
- Baudron F, Sims B, Justice S, Kahan DG, Rose R, Mkomwa S, Kaumbutho P, Sariah J, Nazare R, Moges G, Gerard B (2015) Re-examining appropriate mechanization in Eastern and Southern Africa: two-wheel tractors, conservation agriculture, and private sector involvement. *Food Secur* 7:889–904. <https://doi.org/10.1007/s12571-015-0476-3>
- Berre D, Corbeels M, Rusinamhodzi L, Mutenje M, Thierfelder C, Lopez- Ridaura S (2017) Thinking beyond agronomic yield gap: smallholder farm efficiency under contrasted livelihood strategies in Malawi. *Field Crops Res* 214:113–122. <https://doi.org/10.1016/j.fcr.2017.08.026>
- 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>
- Burke WJ, Jayne TS, Black JR (2017) Factors explaining the low and variable profitability of fertilizer application to maize in Zambia. *Agric Econ* 48:115–126. <https://doi.org/10.1111/agec.12299>
- Burke WJ, Snapp SS, Jayne TS (2020) An in-depth examination of maize yield response to fertilizer in Central Malawi reveals low profits and too many weeds. *Agric Econ* 51:923–940. <https://doi.org/10.1111/agec.12601>
- Cairns JE, Hellin J, Sonder K, Araus JL, MacRobert JF, Thierfelder C, Prasanna BM (2013) Adapting maize production to climate change in sub-Saharan Africa. *Food Secur* 5:354–360. <https://doi.org/10.1007/s12571-013-0256-x>
- Chamberlin J, Jayne TS, Snapp S (2021) The role of active soil carbon in influencing the profitability of fertilizer use: empirical evidence from smallholder maize plots in Tanzania. *Land Degrad Dev* 32:2681–2694. <https://doi.org/10.1002/ldr.3940>
- CIMMYT (2021) Guiding acid soil management investments in Africa. Bill & Melinda Gates Foundation Investment Grant INV-002829. Tech. rep.
- Coelli TJ, Henningsen A (2013) Frontier: stochastic frontier analysis. R package version 1.1-0
- Descheemaeker K, Oosting SJ, Tui SHK, Masikati P, Falconnier GN, Giller KE (2016) Climate change adaptation and mitigation in smallholder crop-livestock systems in sub-Saharan Africa: a call for integrated impact assessments. *Reg Environ Change* 16:2331–2343. <https://doi.org/10.1007/s10113-016-0957-8>
- Doré T, Sebillotte M, Meynard JM (1997) A diagnostic method for assessing regional variations in crop yield. *Agric Syst* 54:169–188. [https://doi.org/10.1016/S0308-521X\(96\)00084-4](https://doi.org/10.1016/S0308-521X(96)00084-4)
- Dray S, Dufour AB (2007) The ade4 package: implementing the duality diagram for ecologists. *J Stat Softw* 22
- Funk C, Peterson P, Landsfeld M, Pedreros D, Verdin J, Shukla S, Husak G, Rowland J, Harrison L, Hoell A, Michaelsen J (2015) The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Sci Data* 2:150066. <https://doi.org/10.1038/sdata.2015.66>
- Giller KE (2020) The food security conundrum of sub-Saharan Africa. *Glob Food Sec* 26:100431. <https://doi.org/10.1016/j.gfs.2020.100431>
- Giller KE, Delaune T, Silva JV, Descheemaeker K, van de Ven G, Schut AGT, van Wijk M, Hammond J, Hochman Z, Taulya G, Chikowo R, Narayanan S, Kishore A, Bresciani F, Teixeira HM, Andersson JA, van Ittersum MK (2021a) The future of farming: who will produce our food? *Food Secur* 13:1073–1099. <https://doi.org/10.1007/s12571-021-01184-6>
- Giller KE, Delaune T, Silva JV, van Wijk M, Hammond J, Descheemaeker K, van de Ven G, Schut AGT, Taulya G, Chikowo R, Andersson JA (2021b) Small farms and development in sub-Saharan Africa: farming for food, for income or for lack of better options? *Food Secur* 13:1431–1454. <https://doi.org/10.1007/s12571-021-01209-0>
- Hachigonta S, Reason CJC, Tadross M (2008) An analysis of onset date and rainy season duration over Zambia. *Theor Appl Climatol* 91:229–243. <https://doi.org/10.1016/j.agee.2013.08.017>
- Harris D, Orr A (2014) Is rainfed agriculture really a pathway from poverty? *Agric Syst* 123:84–96. <https://doi.org/10.1016/j.agry.2013.09.005>
- Hengl T, Heuvelink GBM, Kempen B, Leenaars JGB, Walsh MG, Shepherd KD, Sila A, MacMillan RA, Mendes de Jesus J, Tamene L, Tondoh JE (2015) Mapping soil properties of Africa at 250 m resolution: random forests significantly improve current predictions. *PLoS One* 10:1–26. <https://doi.org/10.1371/journal.pone.0125814>
- Hengl T, Mendes de Jesus J, Heuvelink GBM, Ruiperez Gonzalez M, Kilibarda M, Blagotic A, Shangguan W, Wright MN, Geng X, Bauer-Marschallinger B, Guevara MA, Vargas R, MacMillan RA, Batjes NH, Leenaars JGB, Ribeiro E, Wheeler I, Mantel S, Kempen B (2017) Soilgrids250m: global gridded soil information based on machine learning. *PLoS One* 12:1–40. <https://doi.org/10.1371/journal.pone.0169748>
- Herrmann SM, Mohr KI (2011) A continental-scale classification of rainfall seasonality regimes in Africa based on gridded precipitation and land surface temperature products. *J Appl Meteorol Climatol* 50:2504–2513. <https://doi.org/10.1175/JAMC-D-11-024.1>
- IAPRI (2012) Rural Agricultural Livelihoods Survey—2012 survey report. Tech. rep. Indaba Agriculture Policy Research Institute, Lusaka, Zambia
- IAPRI (2015) Rural agricultural livelihoods survey—2015 survey report. Tech. rep. Indaba Agriculture Policy Research Institute, Lusaka, Zambia
- IAPRI (2019) Rural Agricultural Livelihoods Survey—2019 survey report. Tech. rep. Indaba Agriculture Policy Research Institute, Lusaka, Zambia
- IAPRI (2020) Zambia: agriculture status report. Tech. rep. Indaba Agriculture Policy Research Institute, Lusaka, Zambia
- Jahnke HE (1982) Livestock production systems and livestock development in tropical Africa. Kieler Wissenschaftsverlag Vauk, Kiel
- Jayne TS, Sanchez PA (2021) Agricultural productivity must improve in sub-Saharan Africa. *Science* 372:1045–1047. <https://doi.org/10.1126/science.abf5413>
- Jayne TS, Sitko NJ, Mason NM, Skole D (2018) Input subsidy programs and climate smart agriculture: current realities and future potential. Springer International Publishing, pp 251–273
- Jayne TS, Snapp S, Place F, Sitko N (2019) Sustainable agricultural intensification in an era of rural transformation in Africa. *Glob Food Sec* 20:105–113. <https://doi.org/10.1016/j.gfs.2019.01.008>

- Kamanga BCG, Waddington SR, Whitbread AM, Almekinders CJM, Giller KE (2014) Improving the efficiency of use of small amounts of nitrogen and phosphorus fertiliser on smallholder maize in central Malawi. *Exp Agric* 50:229–249. <https://doi.org/10.1017/S0014479713000513>
- Keating BA, Herrero M, Carberry PS, Gardner J, Cole MB (2014) Food wedges: framing the global food demand and supply challenge towards 2050. *Glob Food Sec* 3:125–132. <https://doi.org/10.1016/j.gfs.2014.08.004>
- Komarek AM, Kwon H, Haile B, Thierfelder C, Mutenje MJ, Azzarri C (2019) From plot to scale: ex-ante assessment of conservation agriculture in Zambia. *Agric Sys* 173:504–518. <https://doi.org/10.1016/j.agsy.2019.04.001>
- Kumbhakar SC, Lovell CAK (2000) Stochastic frontier analysis. Cambridge University Press, Cambridge
- Leenaars JGB, Hengl T, Gonzalez MR, de Jesus JM, Heuvelink GBM, Wolf J, van Bussel L, Claessens L, Yang H, Cassman KG (2015) Root zone plant-available water holding capacity of the sub-Saharan Africa soil version 1.0. Tech. rep. ISRIC
- Masuka B, Atlin GN, Olsen M, Magorokosho C, Labuschagne M, Crossa J, Bänziger M, Pixley KV, Vivek BS, von Biljon A, Macrobert J, Alvarado G, Prasanna B, Makumbi D, Tarekegne A, Das B, Zaman-Allah M, Cairns JE (2017) Gains in maize genetic improvement in Eastern and Southern Africa: I. CIMMYT hybrid breeding pipeline. *Crop Sci* 57:168–179. <https://doi.org/10.2135/cropsci2016.05.0343>
- MoA/CSO (2019) The 2018/2019 crop forecast survey results presentation. Tech. rep. Ministry of Agriculture and Central Statistical Office. Zambia, Lusaka
- Morgan SN, Mason NM, Levine NK, Zulu-Mbata O (2019) Disincentivizing sustainable intensification? The case of Zambia's maize-fertilizer subsidy program. *World Dev* 122:54–69. <https://doi.org/10.1016/j.worlddev.2019.05.003>
- Mupangwa W, Mutenje M, Thierfelder C, Mwila M, Malumo H, Mujeyi A, Setimela P (2017) Productivity and profitability of manual and mechanized conservation agriculture (CA) systems in Eastern Zambia. *Renew Agr Food Syst* 34:380–394. <https://doi.org/10.1017/S1742170517000606>
- Muthoni FK, Odongo VO, Ochieng J, Muglavai EM, Mourice SK, Hoesche-Zeledon I, Mwila M, Bekunda M (2019) Long-term spatial-temporal trends and variability of rainfall over Eastern and Southern Africa. *Theor Appl Climatol* 137:1869–1882. <https://doi.org/10.1007/s00704-018-2712-1>
- Mwila M, Mhlanga B, Thierfelder C (2021) Intensifying cropping systems through doubled-up legumes in Eastern Zambia. *Sci Rep* 11:8101. <https://doi.org/10.1038/s41598-021-87594-0.39>
- Ngoma H, Jayne TS, Chamberlin J (2019) Do changes in farm size distributions promote or retard growth in rural incomes and employment? Evidence from Zambia. In: 6th African Conference of Agricultural Economists. Abuja, Nigeria
- Ngoma H, Lupiya P, Kabisa M, Hartley F (2021) Impacts of climate change on agriculture and household welfare in Zambia: an economy-wide analysis. *Clim Change* 167:55. <https://doi.org/10.1007/s10584-021-03168-775-z>
- Pauw EFD (1994) The management of acid soils in Africa. *Outlook Agric* 23:11–16. <https://doi.org/10.1177/003072709402300104>
- Pelletier J, Ngoma H, Mason NM, Barrett CB (2020) Does smallholder maize intensification reduce deforestation? Evidence from Zambia. *Glob Environ Change* 63:102127. <https://doi.org/10.1016/j.gloenvcha.2020.102127>
- R Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. ISBN 3-900051-07-0
- 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 (2019) Is labour a major determinant of yield gaps in sub-Saharan Africa? A study for cereal-based production systems in Southern Ethiopia. *Agric Sys* 175:39–51. <https://doi.org/10.1016/j.agsy.2019.04.009>
- Silva JV, Reidsma P, Baudron F, Jaleta M, Tesfaye K, van Ittersum MK (2021) Wheat yield gaps across smallholder farming systems in Ethiopia. *Agron Sustain Dev* 41:12. <https://doi.org/10.1007/s13593-020-00654-z>
- Smale M (1995) “Maize is life”: Malawi's delayed Green Revolution. *World Dev* 23:819–831. [https://doi.org/10.1016/0305-750X\(95\)00013-3](https://doi.org/10.1016/0305-750X(95)00013-3)
- 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 MK (2019) Maize crop nutrient input requirements for food security in sub-Saharan Africa. *Glob Food Sec* 23:9–21. <https://doi.org/10.1016/j.gfs.2019.02.001>
- Tittonell P, Giller KE (2013) When yield gaps are poverty traps: the paradigm of ecological intensification in African smallholder agriculture. *Field Crops Res* 143:76–90. <https://doi.org/10.1016/j.fcr.2012.10.007>
- Tittonell P, Muriuki A, Shepherd KD, Mugendi D, Kaizzi KC, Okeyo J, Verchot L, Coe R, Vanlauwe B (2010) The diversity of rural livelihoods and their influence on soil fertility in agricultural systems of East Africa—a typology of smallholder farms. *Agric Sys* 103:83–97. <https://doi.org/10.1016/j.agsy.2009.10.001>
- 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](https://doi.org/10.1016/S0378-4290(97)00037-3)
- van Ittersum MK, Cassman KG, Grassini P, Wolf J, Tittonell P, Hochman Z (2013) 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 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>
- 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 (2013) 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>
- Yang HS, Dobermann A, Lindquist JL, Walters DT, Arkebauer TJ, Cassman KG (2004) Hybrid-maize—a maize simulation model that combines two crop modeling approaches. *Field Crops Res* 87:131–154. <https://doi.org/10.1016/j.fcr.2003.10.003>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.