Field Crops Research 301 (2023) 109014

Contents lists available at ScienceDirect

Field Crops Research

journal homepage: www.elsevier.com/locate/fcr

How much control do smallholder maize farmers have over yield?

Michael Cecil [a,*], Allan Chilenga [b], Charles Chisanga [c], Nicolas Gatti [d,e], Natasha Krell [f], Noemi Vergopolan [g,h], Kathy Baylis [f], Kelly Caylor [f], Tom Evans [i], Megan Konar [j], Justin Sheffield [k], Lyndon Estes [a]

[a] Graduate School of Geography, Clark University, Worcester, MA 01610, USA
[b] Zambia Agriculture Research Institute (ZARI), Ministry of Agriculture, Mount Makulu Central Research Station, Chilanga, Zambia
[c] School of Natural Resources, Department of Plant & Environmental Sciences, Copperbelt University, Kitwe, Zambia
[d] Instituto Nacional de Tecnología Agropecuaria (INTA), Buenos Aires 1425, Argentina
[e] Universidad del CEMA, Buenos Aires 1054, Argentina
[f] University of California Santa Barbara, Department of Geography, UC Santa Barbara, CA 93106-4060, USA
[g] Atmospheric and Ocean Sciences Program, Princeton University, Princeton, NJ 08540, USA
[h] NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ 08540-6649, USA
[i] School of Geography, Development & Environment, University of Arizona, Tucson, AZ 85721, USA
[j] Department of Civil and Environmental Engineering, Newmark Civil Engineering Laboratory, University of Illinois Urbana-Champaign, Urbana, IL 61801-2352, USA
[k] Department of Geography and Environmental Science, University of Southampton, Southampton SO17 1BJ, United Kingdom

ARTICLE INFO

Keywords:
Yield variance
Crop model
Smallholders
Crop management
Zambia

ABSTRACT

Smallholder agriculture is critical for current and future food security, yet quantifying the sources of smallholder yield variance remains a major challenge. Attributing yield variance to farmer management, as opposed to soil and weather constraints, is an important step to understanding the impact of farmer decision-making, in a context where smallholder farmers use a wide range of management practices and may have limited access to fertilizer. This study used a process-based crop model to simulate smallholder maize (*Zea mays*) yield at the district-level in Zambia and quantify the percent of yield variance (effect size) attributed to soil, weather, and three management inputs (cultivar, fertilizer, planting date). Effect sizes were calculated via an ANOVA variance decomposition. Further, to better understand the treatment effects of management practices, effect sizes were calculated both for all years combined and for individual years. We found that farmer management decisions explained 27–82 % of total yield variance for different agro-ecological regions in Zambia, primarily due to fertilizer impact. Fertilizer explained 45 % of yield variance for the average district, although its effect was much larger in northern districts of Zambia that typically have higher precipitation, where it explained 72 % of yield variance on average. When fixing a specific fertilizer amount, the "low-cost" management options of varying planting dates and cultivars explained 20–28 % of yield variance, with some regional variation. To better understand why management practices impact yield more in particular years, we performed a correlation analysis comparing yearly management effect sizes with four meteorologically based variables: total growing season precipitation, rainy season onset, extreme heat degree days, and longest dry spell. Results showed that fertilizer's impact generally increased under favorable weather conditions, and planting date's impact increased under adverse weather conditions. This study demonstrates how a national yield variance decomposition can be used to understand where specific management interventions would have a greater impact and can provide policymakers with quantification of soil, weather, and management effects. In addition, the variance composition can easily be adapted to a different range of management inputs, such as other cultivars or fertilizer quantities, and can also be used to assess the effect size of management adaptations under climate change.

* Corresponding author.
E-mail addresses: mcecil@clarku.edu (M. Cecil), allanchilenga@gmail.com (A. Chilenga), cbchisanga@gmail.com (C. Chisanga), ngatti11@ucema.edu.ar (N. Gatti), nkrell@ucsb.edu (N. Krell), noemi@princeton.edu (N. Vergopolan), baylis@ucsb.edu (K. Baylis), caylor@ucsb.edu (K. Caylor), tomevans@arizona.edu (T. Evans), mkonar@illinois.edu (M. Konar), justin.sheffield@soton.ac.uk (J. Sheffield), LEstes@clarku.edu (L. Estes).

https://doi.org/10.1016/j.fcr.2023.109014
Received 12 January 2023; Received in revised form 9 June 2023; Accepted 16 June 2023
Available online 29 June 2023
0378-4290/

1. Introduction

Smallholder agriculture is a key component of food security in the developing world. Recent studies estimate that about 30 % of global food production is provided by smallholder farms (defined as farms < 2 ha) (Ricciardi et al., 2018), but their relative share in food production varies substantially between regions. This is particularly true in sub-Saharan Africa (SSA), where smallholder farmers represent about 80 % of all farmers and about 40 % of farmland (Lowder et al., 2016). As smallholder farms in SSA are primarily rainfed, they are particularly vulnerable to climate shocks (Müller et al., 2011). In addition, soil constraints can significantly impact potential yields. Acidic soils may limit crop response to fertilizer (Burke et al., 2017) and shallow soil rooting depth may also limit crop yields (Guilpart et al., 2017). Overall, realized maize (*Zea mays*) yields in SSA are less than 20 % of their potential, among the lowest values globally (Lobell et al., 2009). Increasing the resilience of Africa's smallholder farmers while increasing productivity is essential for this region to become food self-sufficient (Ittersum et al., 2016) and to meet its rapidly increasing demand (Searchinger et al., 2015).

An important first step to achieve these objectives is to understand the degree to which the key factors that govern crop productivity (e.g. soil nutrients, rain, and primary management decisions, such as cultivar selection and fertilization rates) vary over time and space, and how these factors affect crop yield (Ittersum et al., 2013). However, the effect of smallholder management on yield has been difficult to quantify for several reasons. First, smallholder farmers use a wide variety of information and heuristics to decide, for example, planting date and cultivar choice (Waldman et al., 2019, 2017), but the effectiveness of farmers' choices for these inputs is unclear. Smallholder farmers may not seek to optimize yields when making management decisions, instead prioritizing the need to harvest crops early for household food security (Thierfelder et al., 2016), and adapting risk-averse strategies for coping with rainfall variability (Cooper et al., 2008). Farmers also may have unequal access to farm subsidies (Mdee et al., 2020), imperfect information on management inputs such as hybrid seed (Waldman et al., 2017), and may experience delays in access to subsidized fertilizer (Mubanga and Ferguson, 2017). More expensive and time-consuming management methods used by commercial farmers, like soil rehabilitation, irrigation, and intensive fertilizer use, also may be inaccessible or mis-adopted by the poorest of smallholders (Burney and Naylor, 2012; Harris and Orr, 2014).

Second, there is a general scarcity of field-level data for smallholder agriculture. Field-level self-reported yields are often inaccurate (Paliwal and Jain, 2020; Carletto et al., 2016; Gourlay et al., 2019; Paliwal and Jain, 2020) and agricultural censuses may collect data at scales (e.g. district) that are too coarse to capture the impact of field-level management decisions.

Third, variability in soil properties and weather conditions can influence management's effect on yield (Lobell et al., 2002). For example, acidic soils may limit the effectiveness of fertilizer application in SSA (Burke et al., 2017). Farmers may adopt different strategies to cope with these soil and weather conditions, such as using differential management intensities between fields of varying fertility (Tittonell et al., 2007; Vanlauwe et al., 2015), and sowing crops earlier to reduce the impact of heat stress (Jain et al., 2017).

Thus, smallholder farmer management decisions are made in the context of multiple farmer priorities, varying degrees of knowledge and access to inputs, critical soil and weather constraints, and data scarcity. Quantifying the impact of smallholder management on yield is challenging in this context, as different methods and data sets each have their own limitations. Remote sensing (RS) based models (Lobell et al., 2015; Burke and Lobell, 2017; Jain et al., 2017; Jin et al., 2019) estimate the effect of soil, weather, and management on yield over large areas, but these studies have several sources of uncertainty due to cloud cover, small field size relative to spatial resolution, and the need for ground calibration data. Field trials provide a high level of control and precision for inputs and yield measurements (Chisanga et al., 2015; Garcia et al., 2009), but trials can only compare a small number of management options over a few sites. Farmer surveys are invaluable sources that define realistic ranges of management parameters, particularly in smallholder environments (e.g. Waldman et al., 2017 and Giroux et al., 2019). However, such surveys are often limited in spatial scope due to cost, and farmer self-reported yields may be inaccurate (Paliwal and Jain, 2020). Surveys thus represent an essential but incomplete answer to data scarcity for smallholder agriculture. As mentioned above, national and sub-national agricultural statistics (e.g. FAOSTAT) can be used to understand impacts of soil, weather, and management at broad scales (Lobell and Asner, 2003; Lobell and Field, 2007; Ben-Ari and Makowski, 2014; Chabala et al., 2015; Iizumi and Ramankutty, 2016; Zhao et al., 2018; Vergopolan et al., 2021), but lack the field-level detail needed to link specific management practices to yield outcomes.

Physical-based, or mechanistic, crop models (hereafter simply "crop models") are widely used to study the effects of soil, weather, and management factors on crop yield (Lobell et al., 2013; Guan et al., 2017; Frieler et al., 2017). Like other methods, crop models have a unique set of advantages and limitations. Crop models are limited in that their outputs represent potential yield in the absence of pests and disease (Estes et al., 2013), and model results must be interpreted in this context. Moreover, models must also be well-calibrated and use inputs suitable for smallholder agriculture (Grassini et al., 2015). Several recent studies have provided guidelines for using crop models in data-scarce environments (Grassini et al., 2015; Kersebaum et al., 2015). Despite these potential limitations, crop models have the key advantage of being able to simulate many different soil-weather-management combinations, at a scale not possible by other methods (Shelia et al., 2019), with precise control of inputs and measurement of outputs. This precision allows yield variance to be directly attributed to changes in inputs, and not to measurement error or uncaptured variables.

In this study, we leveraged this advantage of crop models in an analysis designed to answer the question of how much control smallholder farmers typically have over their yields. Specifically, we used the Decision Support System for Agrotechnology Transfer (DSSAT) (Jones et al., 2003; Hoogenboom et al., 2019a, 2019b) to simulate maize yields in response to three varying farmer management inputs (cultivar, fertilizer, planting date), while varying soil and weather conditions at the district scale in Zambia. The management practices are representative of smallholder maize farmers across Zambia's three agro-ecological regions, including maize cultivars previously calibrated in Zambia for use in DSSAT (Chisanga et al., 2020).

We used Analysis of Variance (ANOVA) to quantify the effects of each of the five inputs (cultivar, fertilizer, planting date, soil, weather) on yield, which allowed us to quantify the effect of management inputs as compared to soil and weather conditions. We also used a correlation analysis to examine how weather conditions relate to the importance of soil and management inputs in different years.

We focused our investigation on maize farming in Zambia for several reasons. First, smallholder maize agriculture in Zambia is critical to food security domestically and in the region, with Zambia having one of the highest self-sufficiency ratios for cereal production in SSA (Ittersum et al., 2016). Second, smallholder agriculture in Zambia, with predominantly rainfed agriculture, a well-defined rainy season, high yearly climate variability, and high variance in management practices (Waldman et al., 2017) is representative of smallholder agricultural conditions in other parts of SSA (Sheahan and Barrett, 2017). This comprehensive approach and the resulting insights may thus be applicable to other smallholder agriculture systems in SSA and can provide policymakers with a high-level understanding of regional drivers of yield variance. Our findings may also help identify both short-term management interventions and longer-term adaptations to improve resilience to soil and weather constraints.

2. Data and methods

2.1. Study area

Due to its relatively low population density and abundant arable land, Zambia is a potential site of agricultural intensification in order for Africa to become food self-sufficient (Ittersum et al., 2016). Maize is the dominant crop in Zambia, and maize production is dominated by small-scale farmers (Chisanga et al., 2015). While Zambia has among the highest levels of available potential cropland, the economic returns on new cropland may not be as high as initially thought (Chamberlin et al., 2014), and thus understanding the degree to which farmer management practices affect yield is critical. Zambia maize yields also fall far below those of other maize exporting countries (Burke et al., 2012), and the gap between realized and potential maize yield remains high in both Zambia and SSA generally (Lobell et al., 2009; Deininger and Byerlee, 2011). Understanding yield variance in Zambia would thus allow policymakers to determine the impact of specific management interventions (e.g., fertilizer subsidies, extension agency recommendations) and how this impact varies spatially.

Climatically, Zambia has a predominantly humid subtropical climate type, with a well-defined summer rainy season (November – April). Zambia is divided into 3 agro-ecological regions (Fig. 1a), with increasing precipitation and longer growing seasons further north.

2.2. Crop simulations

Crop simulations have been used to better understand smallholder agriculture (Jin et al., 2019; Tovihoudji et al., 2019) and can effectively increase the number of pseudo-observations in data-scarce regions. We used a combination of local datasets and expertise to inform input selection. We obtained information on common management practices from information provided by the Zambia Agricultural Research Institute (ZARI), from previously published studies in Zambia (Burke et al., 2017; Chisanga et al., 2020), and from prior surveys of farming households (e.g. Waldman et al., 2017). In particular, we used observations from the Household Income Consumption and Production Survey (HICPS) (Hadunka and Bayliss, 2022) to inform the range of management practices typical of smallholder maize farming in Zambia. The HICPS data include data on cultivar choice, planting date, and fertilizer use of about 1200 smallholder maize farmers across all agro-ecological regions of Zambia.

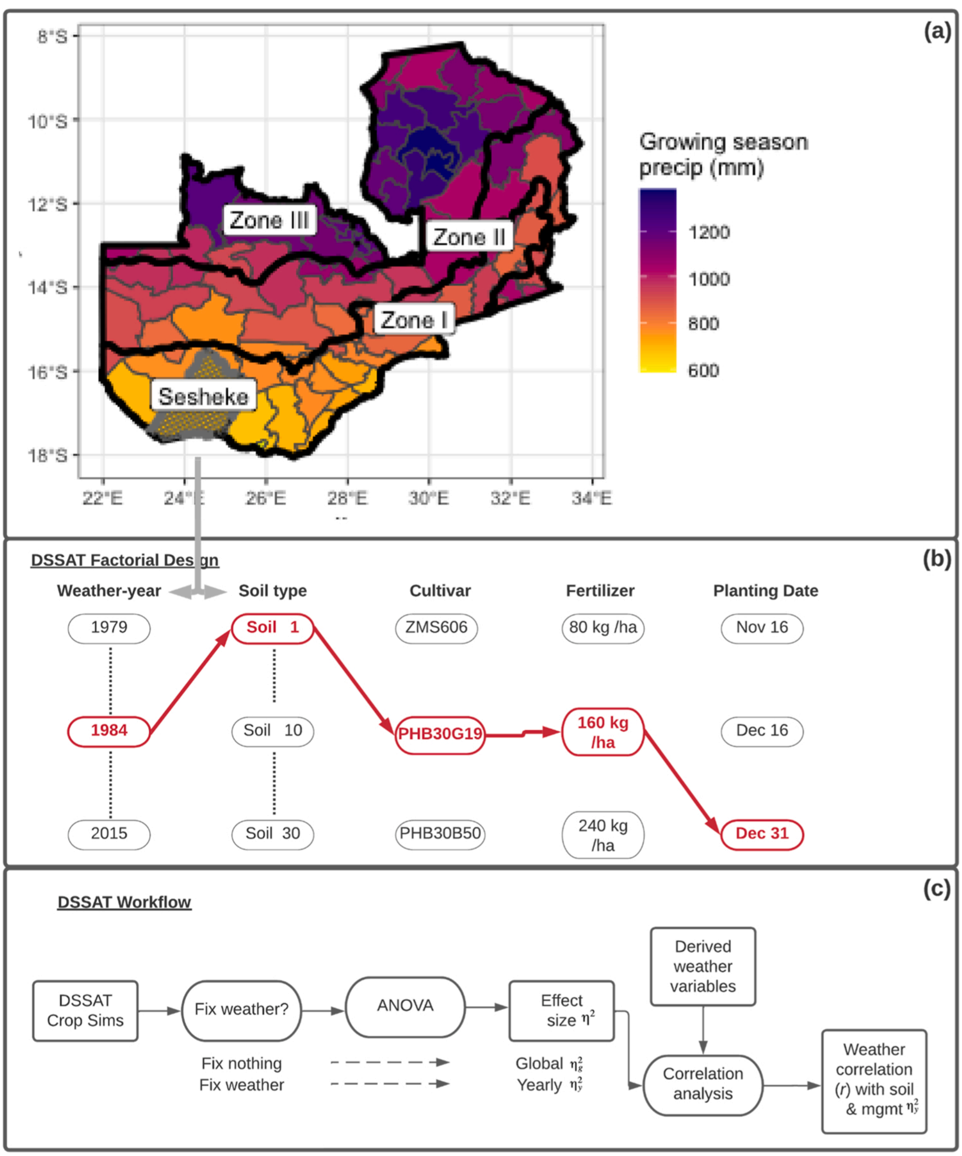

Fig. 1. (a) Zambia map showing average precipitation in the growing season (Nov.- May) (Source: MSWEP). Zambia's three agro-ecological regions are shown with thick borders. (b) Factorial design of crop simulations. Red arrows show one possible permutation of soil, weather, and management inputs. Soil and weather inputs are district-specific, while management inputs are standard across districts. Planting dates vary by agro-ecological region. (c) Crop simulation (DSSAT) variance decomposition workflow. Fixing weather input results in yearly effect sizes, which are then used in a correlation analysis to understand the effect of derived weather variables (e.g. growing season precipitation).

2.2.1. Experiment design

We used the Decision Support System for Agrotechnology Transfer (DSSAT) crop model to simulate crop growth and yield. DSSAT is a process-based model which takes weather, soil, and management inputs to simulate the biophysical processes that drive crop growth. DSSAT simulations can provide both daily and seasonal estimates of crop growth, such as biomass, leaf area index (LAI), and grain yield. Our analysis looked at the effect of changing soil, weather, management practices, and input parameters on end-of-season yield estimates.

The crop simulations used a factorial design (Fig. 1b) that iterated through every permutation of weather-year, soil type, and management (cultivar, fertilizer, planting date) inputs for each district (Table 1). The factorial design allowed for the calculation of yearly effect sizes by fixing weather and allowing other inputs to vary, and the examination of effects when fertilizer level was fixed. The spatial scale of the crop model simulations was the 72 pre-2013 district boundaries of Zambia. Using the pre-2013 boundaries also allowed for backward compatibility with government agricultural surveys, like the Crop Forecast Survey (CFS) and Post-Harvest Survey (PHS), and ready comparison with previous agricultural studies in Zambia (Zhao et al., 2018); (Vergopolan et al., 2021).

2.2.2. Input data for crop simulations

Both the soil and weather inputs for crop modeling leveraged new fine-scaled gridded data sets that allow for spatial modeling in Sub Saharan Africa (SSA). For weather data, we used Multi-Source Weighted-Ensemble Precipitation (MSWEP) data (Beck et al., 2017a) for precipitation inputs, and Princeton Global Forcings (PGF) data (Sheffield et al., 2006) for other weather inputs (temperature, solar radiation, wind, pressure, and specific humidity). MSWEP combines high quality precipitation from seven datasets derived from gauge observations, satellite remote sensing, and atmospheric model reanalysis. MSWEP and PGF were jointly downscaled to a 3-hr 5-km resolution in Zambia for consistency between water and energy fluxes and then converted to daily values for use in the DSSAT crop model. These data sets have been applied (separately) previously in DSSAT simulations (Elliott et al., 2014; Glotter and Elliott, 2016; Yang et al., 2020). To represent natural weather variability and to ensure comparability between simulations, weather data were extracted at district centroids within district boundaries for 37 growing seasons (1979–80 through 2015–16).

Soil inputs were taken from the Global High-Resolution Soil Profile Database for Crop Modeling Applications data set developed specifically for use in the DSSAT crop model (Han et al., 2019). This data set integrates gridded data from the SoilGrids 1 km dataset (Hengl et al., 2014) and HarvestChoice HC27 soil profiles. The SoilGrids profiles are the most extensive gridded soil product available for Africa, based on over 100,000 soil sample locations and using a regression-kriging method with elevation, land cover, and satellite covariates for model fitting. This soil dataset uses a 5 arc-minute resolution (~10 km). Voronoi polygons were used to subdivide each district's area and all grid points whose polygon contained at least 5 % cropland were retained. After this filtering, candidate soil locations were ordered by the proportion of cropland area within their respective Voronoi polygon. The 30 candidate soil locations with the higher proportion of cropland area in each district were retained. For districts with fewer than 30 candidate soil locations, all soil locations were retained. One district, Luangwa had no such candidate points, but this district is primarily contained in the Lower Zambezi National Park. Cropland masking was based on the Global Food Security Support Analysis Data (GFSAD) 30-meter cropland mask (Xiong et al., 2017). For mapping purposes, the simulations of the neighboring district (Chongwe) were duplicated and used for Luangwa. All other districts had at least 4 candidate soil points. For reference, Fig. S1 shows a sample soil profile from Choma district in southern Zambia.

Table 1
Levels of input variables used for DSSAT crop simulations. One crop simulation was run for each unique combination of input variables. In total, there are about 4–30k simulations per district (all permutations of soil, weather, and management inputs).

Input variable	Levels
Weather	37 weather-years extracted at district centroid for seasons starting in 1979–2015.
Soil profiles	4–30 soil profiles per district. Profiles are selected based on highest percent cropland near grid point.
Cultivar	3 medium varieties suitable for all AER (120–130 days to maturity)
Fertilizer	3 fertilizer levels (80, 160, 240 kg/ha each of urea and Compound D)
Planting Date	3 planting dates determined by AER. (Nov.16, Dec. 16, Dec.31 for AER 1 & 2; Dec.1, Dec. 31 for AER III)
Total simulations per district: ~ 4–30k	

We used three medium-maturity cultivars (ZMS606, PHB30G19, PHB30B50) that have been previously calibrated for Zambia. These cultivars were selected due to the extensive field trials and extensive data collection used in their calibration, including weather, soil, management, and crop yield/biomass data (Chisanga et al., 2020). Although the use of early-maturing cultivars is increasing, medium-maturing cultivars were the most commonly used in Zambia over the past decade (Blekking et al., 2021). The ZMS cultivar is produced by ZamSeed and the PHB cultivars are produced by Pioneer, with all cultivars having a typical maturity between 120 and 130 days (Chisanga et al., 2020). These cultivars are also used by smallholder farmers across all three agro-ecological regions (Chisanga et al., 2021).

Cultivar selection was made after extensive review of published studies on cultivars (Chisanga et al., 2021). We used three conditions for selection of cultivars: (i) calibration within Zambia or neighboring country; (ii) all cultivars should be calibrated within a single set of field trials to assure consistency in measurement and crop management; (iii) calibration should include detailed collection of meteorological, soil, management, crop phenology, biomass, and yield data. Based on these constraints, we chose the three medium maturing cultivars mentioned above. Cultivars from other studies were not used because the studies only included two cultivars (Corbeels et al., 2016), were not based in southern African (e.g. Adnan et al., 2020) or did not include cultivar coefficients (Tesfaye et al., 2016). We also considered using data from a limited set of field trials in three Zambia districts (Choma, Kafue, Kabwe). These trials used early, medium, and late maturing cultivars, however the field data collected were not sufficiently detailed for a robust calibration of cultivar parameters in DSSAT. We considered the large difference in yield potential for these cultivars to be an unrealistic of smallholder cultivar choice. We have added a brief discussion of the variance decomposition for these alternative cultivars in Section 4.3.

Using three medium maize cultivars of similar maturity also reflects the difficulty of cultivar choice for smallholder farmers in Zambia in a context of incomplete information (Blekking et al., 2021). Two of the cultivars used (ZMS606 and PHB30G19) were among the most common cultivars used in a previous survey in the Choma district (Waldman et al., 2017) and were also well represented among primary plantings in the HICPS 2016–17 data set (~ 3 % each). The PHB30B50 cultivar is designed for commercial operations growing under both rainfed and irrigated conditions.

We note that the cultivar effect sizes calculated in the variance decomposition are dependent on the specific cultivars used (similar to how other effect sizes are dependent on the specific input levels used). We discuss these caveats and how to interpret effect sizes in greater detail in the Discussion section. The DSSAT phenology and growth coefficients for the cultivars used are listed in Table 2.

Three levels of fertilizer applications were used in all districts, representing 80, 160, and 240 kg/ha of top dressing as Urea (46 % N) at a depth of 2 cm, with applications at 24 and 37 days after planting (DAP). The same quantities were applied as basal fertilizer (compound D [N-P-

Table 2
DSSAT cultivar coefficients for the three cultivars used in simulations.

DSSAT Coefficients	Cultivars		
	ZMS606	PHB30G19	PHB30B50
P1	159	209.9	155.1
P2	1.85	0.441	1.763
P5	810.2	815.9	800.4
G1	945	840.8	795.6
G2	8.559	8.84	15.34
PHINT	59.7	56	59.73

*P1: Degree days (base 8 °C) from emergence to end of juvenile phase; P2: Photoperiod sensitivity coefficient (0/1.0); P5: Degree days (base 8 °C) from silking to physiological maturity; G2: Potential kernel number; G5: Potential kernel growth rate mg/(kernel d); PHINT: Degree days required for a leaf tip to emerge (phyllochron interval) (°C d)

K: 10–20–10]) at a depth of 5 cm at planting. These quantities were selected after consultation with Zambian agronomical experts, and consideration of the following observations. First, these two fertilizer types are by far the most commonly applied in Zambia (Burke et al., 2017) and are supported by the Fertilizer Input Support Program (FISP), which provides 3 bags each (150 kg) of Urea and Compound D to farmers. Given the government support for fertilizer subsidies, we excluded especially low fertilizer quantities. Second, similar ranges have been used in previous studies in Zambia (e.g., 120–240–360 kg/ha, Chisanga et al., 2020). The Crop Forecast Survey from 2006 to 2011 also found that fertilizer users applied on average 150 kg total fertilizer per ha (Burke et al., 2017), which is at the lower end of our scale. The HICPS 2016–17 survey found the median and 95th percentile of fertilizer application to be 100 and 250 kg/ha for top dressing application, with the same values for basal fertilizer. Given these ranges, we decided upon the quantities of 80, 160, and 240 kg/ha for each of urea (N-P-K: 46–0–0) and compound D (N-P-K: 10–20–10). These values are referred to as low, medium, and high fertilizer rates.

Planting dates were determined by agro-ecological regions (AERs) based on ZARI knowledge of management practices. For AER III (northern Zambia), dates of Nov. 1, Dec. 1, and Dec. 31 were used (days 305, 335, 365). For AERs I and II (central and southern Zambia), planting dates Nov. 16, Dec. 16, and Dec. 31 were used (day 320, 350, 365). Planting dates were not adjusted for leap years. These dates correspond with most maize planting dates as determined by household surveys (e.g. (Waldman et al., 2017) for Choma district; HICPS 2016–17 for all regions).

The factorial design included 37 years of weather data (for growing seasons starting in 1979–2015), 4–30 soil profiles per district, three cultivars, three levels of fertilizer input, and three planting dates determined by AERs. Running all permutations of input variables resulted in approximately 4–30 thousand simulations per district (Table 1).

2.3. *Assessing primary sources of yield variance*

We decomposed the variance in DSSAT simulated yields using an Analysis of Variance (ANOVA) sum of squares method, which provides the ability to: (i) consider the effects of multiple inputs simultaneously, and (ii) calculate effect sizes for categorical input variables. With the ANOVA variance decomposition, effect sizes for different inputs (representing the percent of yield variance explained) can be added together, and thus the total effect of all management inputs can be calculated as the sum of individual management effect sizes.

In the ANOVA variance decomposition, the variance explained by an input variable is based on how overall variance in a response variable (e.g. crop yield) can be explained by groupings of an input variable. This effect size is represented by eta-squared (η^2), the ratio of between-group sum of squares to the total sum of squares, which provides an estimate of the percent of variance explained by grouping of an input variable (Cohen, 1973). Fig. 2 illustrates an example of η^2 calculation for nine observations and a single grouping variable (soil type). For the DSSAT crop simulations, we simultaneously considered five potential grouping variables (cultivar, fertilizer, planting date, soil type, and weather-year).

2.3.1. *Global effect sizes*

“Global” effect sizes (η_g^2) are calculated from all crop simulations in a district and represent the proportion of yield variance explained when weather, soil, and management inputs vary among all possible permutations. We used Jenks breaks (Jenks, 1977) to divide weather years and soil profiles into three groupings for each district, based on the mean yield level for each year and profile. Jenks breaks were calculated using the R package “BAMMtools”, using the function “*getJenksBreaks()*”. These groupings allow for a more intuitive grouping of weather and soil constraints. We used a second-degree ANOVA to include interaction effects. Each district has one η_g^2 value for each input, representing the percent of variance explained by that input.

2.3.2. *Total management effect*

We also calculated the **total management effect** ($\eta_{g,TM}^2$), which represents the combined effect of management variables (cultivar, fertilizer, planting date) for the global tier. The total management effect is the sum of the effect sizes of the three management variables, including first and second-order contributions of management inputs.

2.3.3. *Fixed fertilizer cases*

For global effects, we also examined scenarios for fixed low, medium, and high fertilizer rates. As changing fertilizer rates is a more cost-intensive management strategy, this process allows us to examine the effect of lower cost management adaptations (cultivar choice and planting date) on yield variance. We use the same notation for global effects (η_g^2) in the fixed fertilizer cases, using context to specify the fertilizer level. We also calculated the total management effect for fixed fertilizer cases, which represents the combined effect of cultivar, planting date, and their interaction as compared to soil, weather, and other interaction effects.

2.3.4. *Yearly effect sizes*

We also calculated **yearly effect sizes** (η_y^2), which fix a given weather-year in a district and calculate effect sizes for all simulations in a given growing season. Yearly effect sizes represent the proportion of yield variance explained by soil or management inputs when weather is held constant. We used the yearly effect sizes to answer two questions about yield variation: (i) how do the impacts of soil and management on yield vary between years? and (ii) which weather variables explain this variation?

As each district has 37 distinct weather-years, each district has 37 η_y^2 for each input factor (e.g. soil and management). As in the global case, we use Jenks’ breaks to separate soil profiles into 3 distinct groups based on mean yield for each profile. As weather is fixed for each year, we do not use Jenks’ breaks for separating weather-years.

To summarize the 37 η_y^2 values, we calculated the median (50th percentile) and standard deviation of the 37 η_y^2 for each input factor. A higher standard deviation of η_y^2 for an input implies that the importance of this input in explaining yield variance varied to a greater degree due to yearly weather changes.

2.4. *Weather correlation analysis*

In addition to assessing how the impacts of soil and management on yield varied by year, we also used the yearly effects to identify which weather variables are most associated with changes in these variables’ importance. Specifically, we computed the Pearson correlation (r),

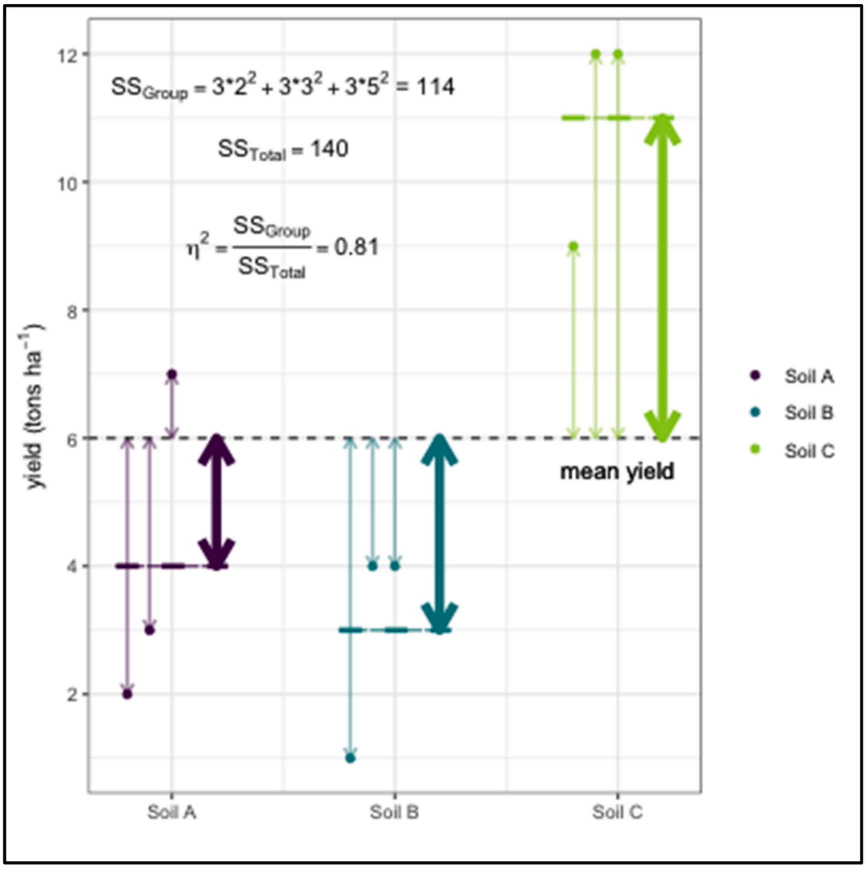

Fig. 2. Schematic of ANOVA eta-squared statistic. The diagram shows 9 individual observations (small dots) grouped by three soil types (colors). Eta-squared (η^2) represents the ratio of group sum of squares (represented by squared distance of thick arrows between group means and overall mean) to the total sum of squares (represented by squared distance of thin arrows between individual observations and overall mean). In the group sum of squares calculation, the squared distance from the group mean to the overall mean is weighted by the size of the group (in this case, each group has 3 observations). In this example, the soil variable explains 81 % of the total yield variance.

between the soil and management yearly effect sizes (η_y^2) and four meteorologically based variables: total growing season precipitation, rainy season onset, extreme heat degree days, and longest dry spell.

These meteorological variables have each been shown to impact maize yield and farmer management decisions. Water is an important limiting factor for rainfed maize (Edreira et al., 2018) and total precipitation has been shown to be more important than rainfall intensity or timing under climate change (Guan et al., 2015). The total growing season precipitation was summed from late October (DOY 298) to the end of March. We selected these dates based on the range of planting dates used (Nov 1. - Dec. 31) and the mean and median harvest dates (DOY 85 and 88, in late March). The start date considered is seven days before the first possible planting date (DOY 305), given the seven-day initialization period for each crop simulation prior to planting.

The definition of rainy season onset is taken from (Tadross et al., 2009) and was defined as the first day when at least 45 mm precipitation accumulated within four days. Similar to the total precipitation variable, we calculated the rainy season onset starting in late October (DOY 298). Among all district-years, there was no rainy season onset in 0.3 % of cases due to low and sporadic rainfall. In these cases, we assigned the rainy season onset as DOY 105 (March 16), which is the latest day of rain onset seen in any district. Smallholder farmer perception of changes in rainy season onset has been an important factor in farmers' adjusting planting date and cultivar selection (Waldman et al., 2017) even though farmers' perceptions of later rain onset is not seen in meteorological data (Waldman et al., 2019, 2017).

Extreme heat degree days were calculated as the number of degree days exceeding 30 ℃, based on daily maximum temperature, for the months of January and February. Extreme heat has been shown to have a greater impact on soil water demand than reduced precipitation (Lobell et al., 2013), and we use this paper's threshold for calculating extreme heat degree days (degree days exceeding 30 ℃). We use daily maximum temperature to calculate degree days exceeding the threshold (as opposed to Lobell's paper which estimates hourly temperatures). These months represent the critical growth stage for maize, with the mean and median anthesis dates for all simulations occurring in early February (DOY 33 and 38). Maize crops are known to be especially vulnerable to drought stress near anthesis (Banziger et al., 2000).

Dry-spells were calculated as consecutive days with 1 mm rain or less, also during January and February. We allowed for 1 mm of rain to account for averaging and interpolation in the gridded precipitation data. Dry-spells of 10 days or longer are common in Zambia (Chabala et al., 2013), which may limit crop yield or result in crop failure (Waldman et al., 2017).

The correlations between yearly effect sizes and meteorological variables can be interpreted as showing whether an increase/decrease in the meteorological variables correlates with an increase or decrease in the percent of variance explained by an agricultural input. For example, if precipitation and planting date are inversely correlated, an increase in precipitation is associated with a decline in the importance of planting date, and vice versa.

3. Results

The district mean yields of DSSAT simulations ranged from 4901 to 7910 kg/ha, with slightly higher yields in northern Zambia (Region III). The coefficient of variation (CV) of yield was on average higher in central and southern Zambia (Regions I & II), where there was a small percent of failed crops (Table 3). Simulated yields matched well with yields from the original crop experiments (6–9 tons per ha; Chisanga et al., 2019). These yields are higher than observed Zambian maize yields, which have averaged about 2600 kg/ha since 2010 (FAO, 2021). This difference between observed and potential yields is common in SSA (Blekking et al., 2021; Ittersum et al., 2016). Thus the effect sizes and

Table 3
Summary statistics for DSSAT crop simulations. Statistics were first calculated within each district, and then the statistics for districts in each agro-ecological region were averaged. Failed crop percent represents the percent of crop simulations with zero yield.

Agro-ecological Region (AER)	Mean yield (kg/ha)	CV yield	Failed crop percent
Region I	6639	0.27	0.03 %
Region II	6873	0.22	0.004 %
Region III	7041	0.19	0 %

variance decomposition should be interpreted in the context of potential (but achievable) grain yield under the specified management conditions and in the absence of weeds, pests and other biotic and abiotic limitations not simulated by DSSAT. Effect sizes are summarized at the agro-ecological region level, as smallholder agricultural recommendations, including crop suitability, planting timing and cultivar choice, are commonly recommended at this scale in Zambia (e.g. publications by the World Bank (CIAT, 2017) and the United Nations Development Programme (UNDP, 2010)).

3.1. Global effect sizes, all fertilizer levels

Fertilizer dominated global effect sizes (η_g^2), explaining on average 45 % of yield variance within a district, but the size of the effect was highly dependent on region and district. By comparison, cultivar, planting date, soil, and weather-year explained 4 %, 3 %, 7 %, and 16 % of yield variance respectively (Table 4). The η_g^2 values for fertilizer were highest in northern Zambia (Fig. 3b), with fertilizer explaining 72 % of yield variance on average in Region III. For comparison, fertilizer explained 17 % and 40 % of yield variance for the mean district in Regions I and II, respectively.

By contrast, weather-year had the highest effect in the drier, hotter conditions of southern Zambia, explaining 30 % of yield-variance in Region I as compared to 16 % and 5 % in Regions II and III, respectively. Other notable findings include the greater importance of cultivar selection in Region III compared to Regions I and II, while planting date was 5–6 times more influential in Regions I and II than in Region III. At a district-level, fertilizer explained the highest percent of yield variance in 54 of 72 districts, weather-year was most important in 17 districts (15 of which are in Region I), and soil in 1 district (also in Region I). The box and whisker plot (Fig. 3f) shows the range of η_g^2 for all districts. We also examined second-order interaction effects which are summarized in the supplementary material (see Table S1). These interaction effects were generally small in magnitude (≤ 1 % for all regions), with the exception of weather-year X planting-date interactions, which explained 3 % of yield variance in all districts, including 4 % in Region II, and 6 % in Region I.

The global effect sizes (η_g^2) represent the percent of variance explained by an input in each district. However, districts may have different levels of yield variability in absolute terms. To examine absolute yield variability, we calculated the coefficient of variation (CV, a normalized measure of variance) of yield and the range of yields in response to each variable. Fig. 4 shows the coefficient of variation (cv) of yield in each district. Clearly yield has higher variability in southern districts, where weather has an increased effect size, although average yield is slightly lower in these districts.

However, as this paper is concerned with understanding yield variance at the district-level, we focus on ANOVA based effect sizes for the rest of this study. The ANOVA effect sizes can be readily compared and allow for calculation of interaction effects. However, understanding the absolute variance of yield magnitude may be helpful in comparing policy interventions across regions.

To quantify the combined effect of all management practices, we calculated the total management effect sizes ($\eta_{g,TM}^2$) for all simulations in each district. This effect size represents the sum of first-order effects for cultivar, planting date, and fertilizer, and the second-order effects involving any two of these factors. These effects can be contrasted with the effects of soil, weather, and other interaction effects. Fig. 5 shows the total management effects for all districts, with a clear trend of higher management effect in northern districts, largely due to fertilizer impact. Region III has a total management effect size of 0.82, meaning that 82 % of yield variance can be explained by management decisions only. By contrast Regions I and II have a mean $\eta_{g,TM}^2$ of 0.27 and 0.50 respectively (Table 4).

3.2. Global effect sizes, fixed fertilizer levels

We also calculated η_g^2 for fixed fertilizer levels (low, medium, and high). In these scenarios, no single input factor dominates the yield variance explanation, but clear spatial patterns exist (Fig. S2). In each of the fertilizer scenarios, the cultivar is more important in northern districts, and weather-year in southern districts. Soil has increased importance in several isolated districts, some in Eastern Zambia. These patterns largely match those seen when combining all fertilizer levels. However, certain effect sizes change in magnitude as fertilizer rate is increased (Fig. 6), with weather becoming more important and soil less important across all regions. Cultivar selection decreases in importance in Regions I and II with increased fertilizer rates, while in Region III, cultivar explains the higher percent of yield variance in the medium fertilizer case. Planting date increases in importance in Regions I & II with increased fertilizer rates, while it decreases in importance in Region III (exact effect sizes are displayed in Table S2).

The total management effect ($\eta_{g,TM}^2$) for the fixed fertilizer cases were calculated slightly differently than when all fertilizer levels are considered (dashed line in Fig. 6). In the fixed fertilizer cases, fertilizer choice was excluded from the yield variance calculations. The total management effect thus represents the effects of cultivar, planting date, and their interaction, as compared to soil, weather, and other interaction effects (e.g. soil x management, weather x management, soil x weather). It is important to note that the total management effect sizes for fixed fertilizer rates should not be compared directly to the total management effect sizes for all fertilizer levels.

3.3. Yearly effect sizes

As in the case for global effect sizes, fertilizer dominated the spatial pattern of median values for yearly effect sizes (η_y^2, Fig. 7, left), particularly for northern districts. Notably, both fertilizer and planting date had high yearly standard deviations (Fig. 7, right) in central and southern districts, indicating that the importance of these factors varied greatly from year to year in these districts.

Table 4
Global effect sizes averaged across all districts and averaged within each agro-ecological region. (Standard deviation within each region shown in parentheses).

Mean district η_g^2 across:	Soil	Weather- year	Cultivar	Fertilizer	Planting Date	Total Management Effect
all districts	0.07 (0.07)	0.16 (0.13)	0.04 (0.03)	0.45 (0.27)	0.03 (0.03)	0.55 (0.27)
Region I	0.09 (0.07)	0.30 (0.11)	0.02 (0.01)	0.17 (0.13)	0.06 (0.02)	0.27 (0.13)
Region II	0.08 (0.08)	0.16 (0.06)	0.04 (0.03)	0.40 (0.14)	0.05 (0.04)	0.50 (0.13)
Region III	0.04 (0.06)	0.05 (0.01)	0.07 (0.02)	0.72 (0.09)	0.01 (0.01)	0.82 (0.10)

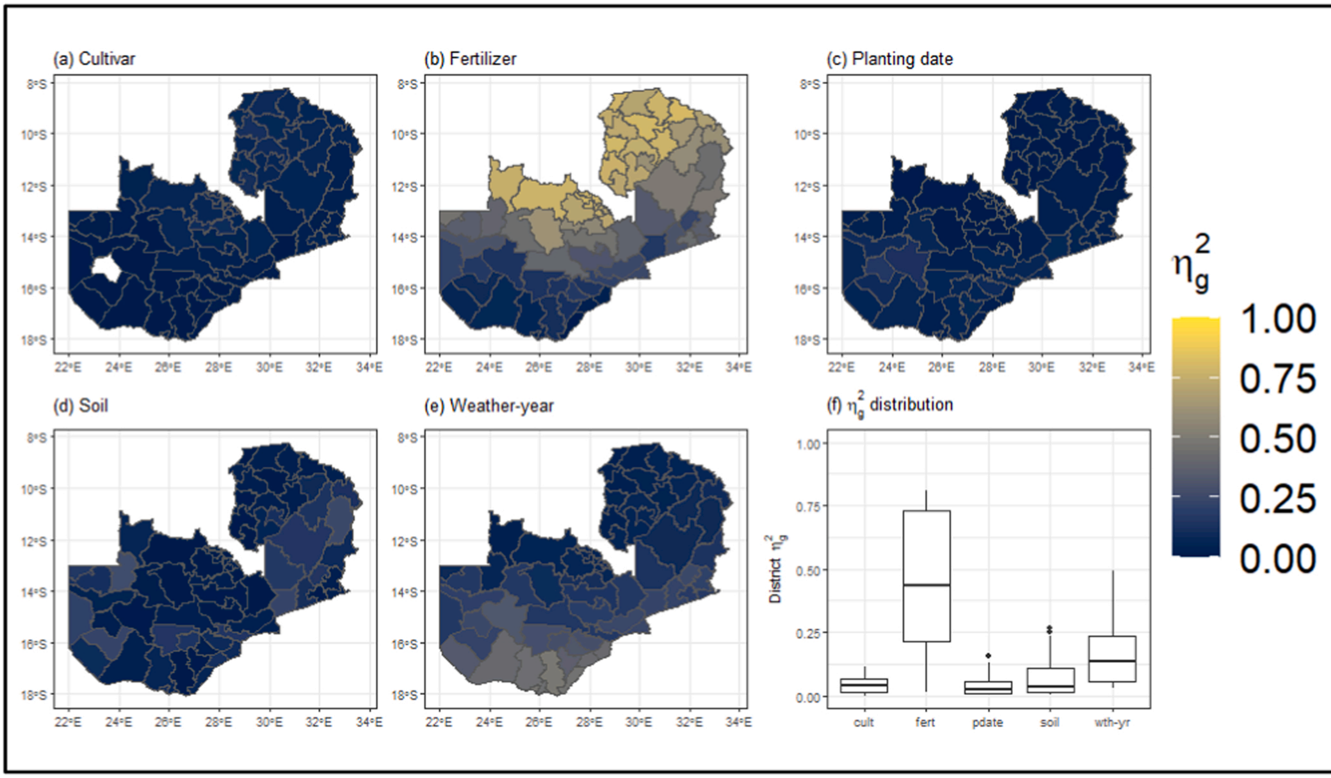

Fig. 3. (a-e) Global effect size (η_g^2) for all DSSAT crop simulations by district. Only districts with significant effect sizes ($p < 0.01$) are shown. (f) Box and whisker plot showing district median, interquartile range, and outliers for each global effect size.

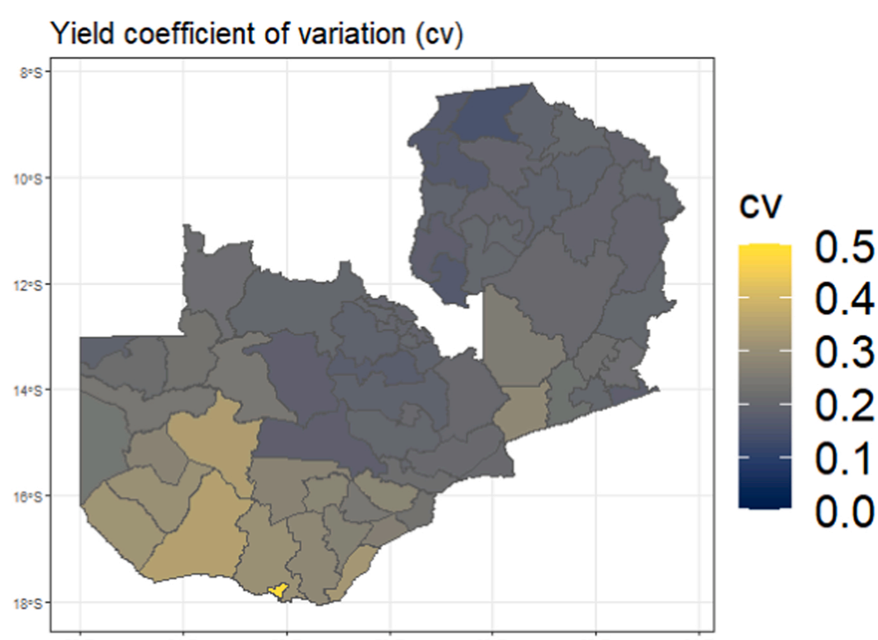

Fig. 4. Coefficient of variation (cv) for yield (kg/ha) among all simulations in each district. Southern districts have increased yield variability compared to northern districts.

As a case study, we examined the yearly variability for one district in each of Zambia's agro-ecological regions: Choma in Region I (southern Zambia), Kapiri Mposhi in Region II (central Zambia) and Mpulungu in Region III (northern Zambia). These three districts are also outlined in Fig. 9 for context.

Choma is known for having shorter growing seasons, with less reliable precipitation and higher maximum temperatures than districts further north. The yearly effect sizes for cultivar, fertilizer, planting date, and soil for Choma are shown in Fig. 8 (bottom), arranged by year. Fertilizer and planting date clearly had the dominant effect size in most years, with some degree of trade-off in the importance of these two variables. Cultivar was also important in a few isolated years, while soil

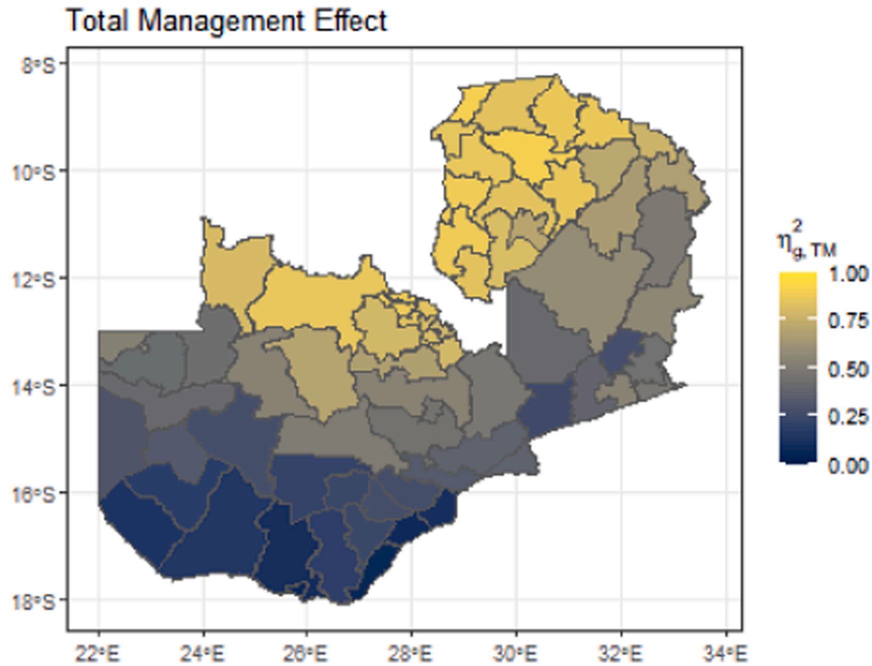

Fig. 5. Total management effect size for all simulations. This effect size represents the combined effect of cultivar, planting date, and fertilizer, and second-order interactions of these inputs.

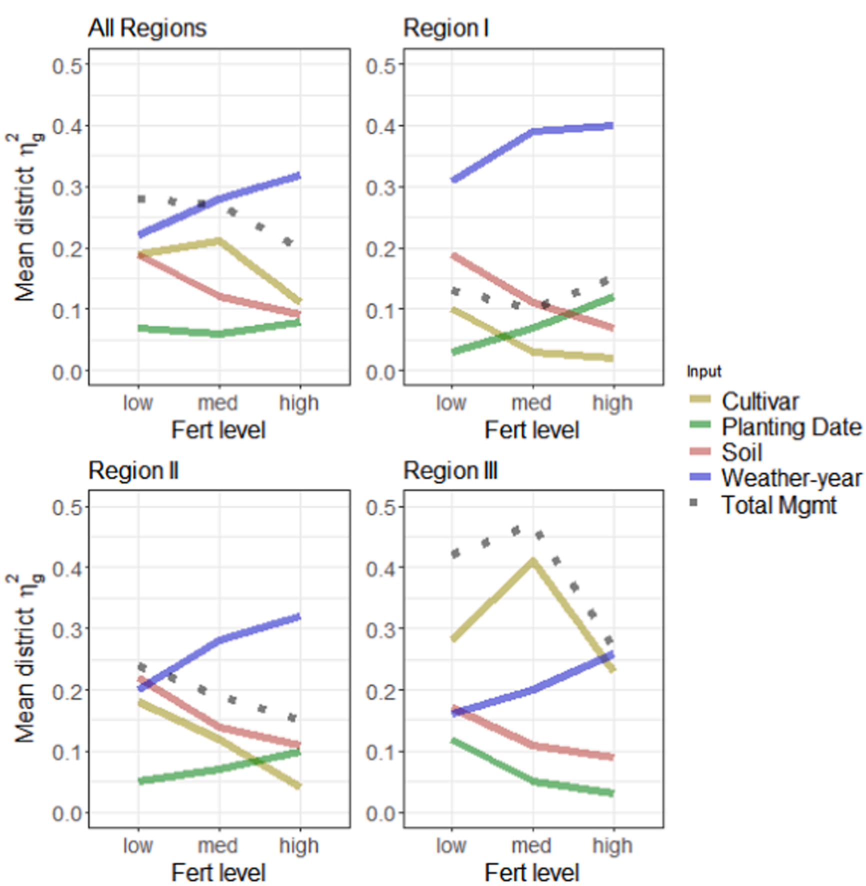

Fig. 6. Effect size change for low, medium, and high fertilizer rates. The four plots show results for all districts (top-left) and for Regions I, II, III (top-right, bottom-left, bottom-right).

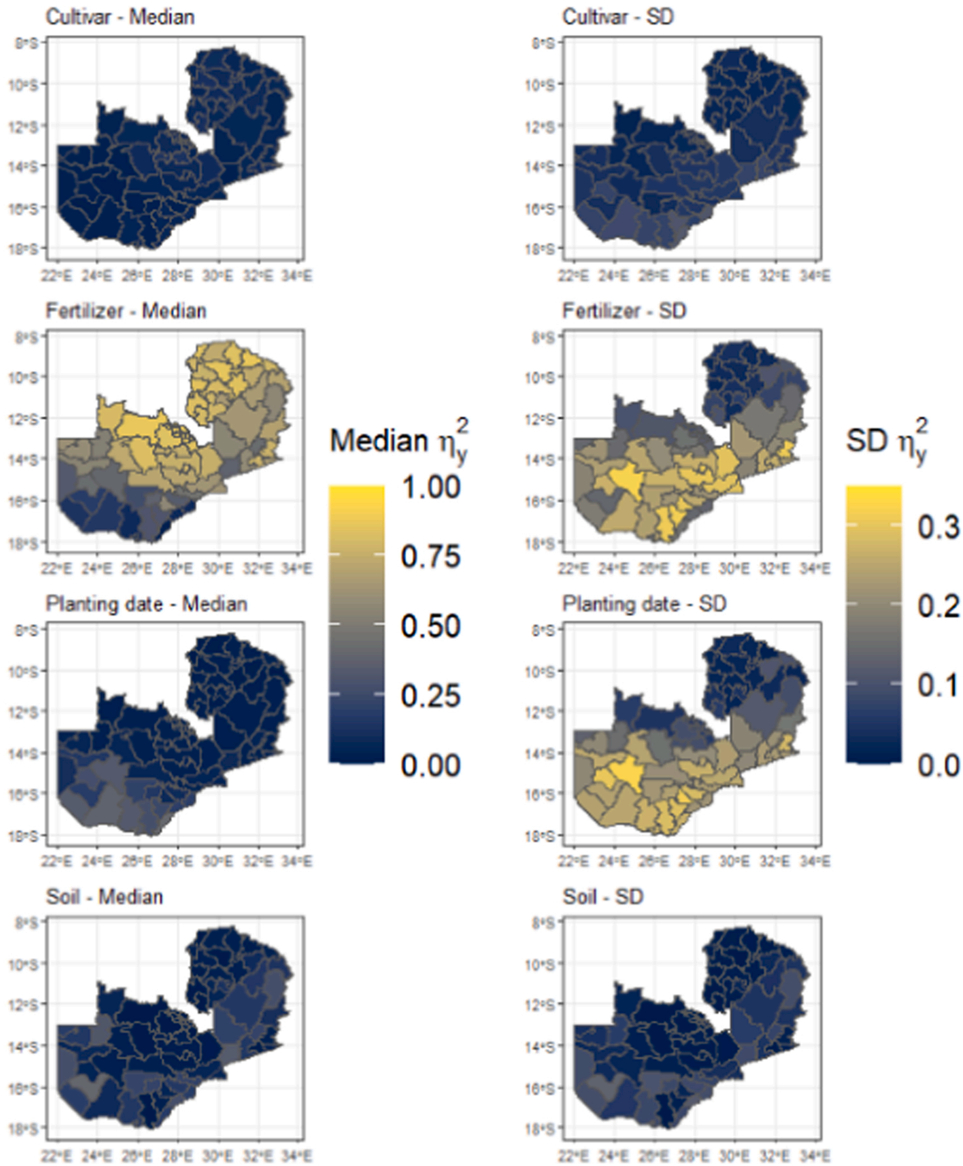

Fig. 7. Median (left) and standard deviation (right) of yearly effect sizes (η_y^2).

had consistently low importance. By contrast, in Kapiri Mposhi (Fig. 8, middle), there was also a degree of trade-off in fertilizer and planting date effect sizes, although there were fewer years where planting date had a larger effect, compared to Choma district. There were also not any years where cultivar had a large (> 0.20) effect. In Mpulungu district (Fig. 8, top), the effect sizes were almost perfectly consistent, with a large effect size for fertilizer in all years, and very small amounts of variation in effect size for all inputs.

3.4. Weather-management correlation analysis

For each district, we determined the strength of the association between the four meteorological variables (total growing season precipitation, rainy season onset, extreme heat days, and longest dry spell) and yearly effect sizes. Fig. 9 shows the district-level correlations between meteorological variables and effects sizes for soil and management. For example, the third row and second column of Fig. 9 show the correlation between total precipitation and planting date effect size, with positive correlations in gold and negative correlations in blue. For much of Regions I and II, an increase in total precipitation is negatively correlated with planting date effect size, thus planting date is more important when conditions are drier. There are a handful of districts in northeastern Zambia with a countervailing effect, where increased precipitation correlated with an increase in planting date importance.

To illustrate these correlations for a single district, Fig. S3 shows the scatterplots comparing yearly effect sizes with the meteorological variables for Choma district. Two such correlations were significant ($p < 0.05$). There was one outlier with a high cultivar effect size, which occurred in a year with low total precipitation and a high number of heat stress days (although correlations between cultivar η_y^2 and these meteorological variables were not significant).

4. Discussion

Our analysis is the first national-scale assessment of factors that drive maize yield variance under typical smallholder management conditions in Zambia, a regionally significant agricultural producer that is a bellwether for agricultural development and a potential breadbasket for Africa (Ittersum et al., 2013).

In particular, this study builds on previous work investigating the factors influencing smallholder yield variance in three notable ways. First, we examined five variables, including three distinct management

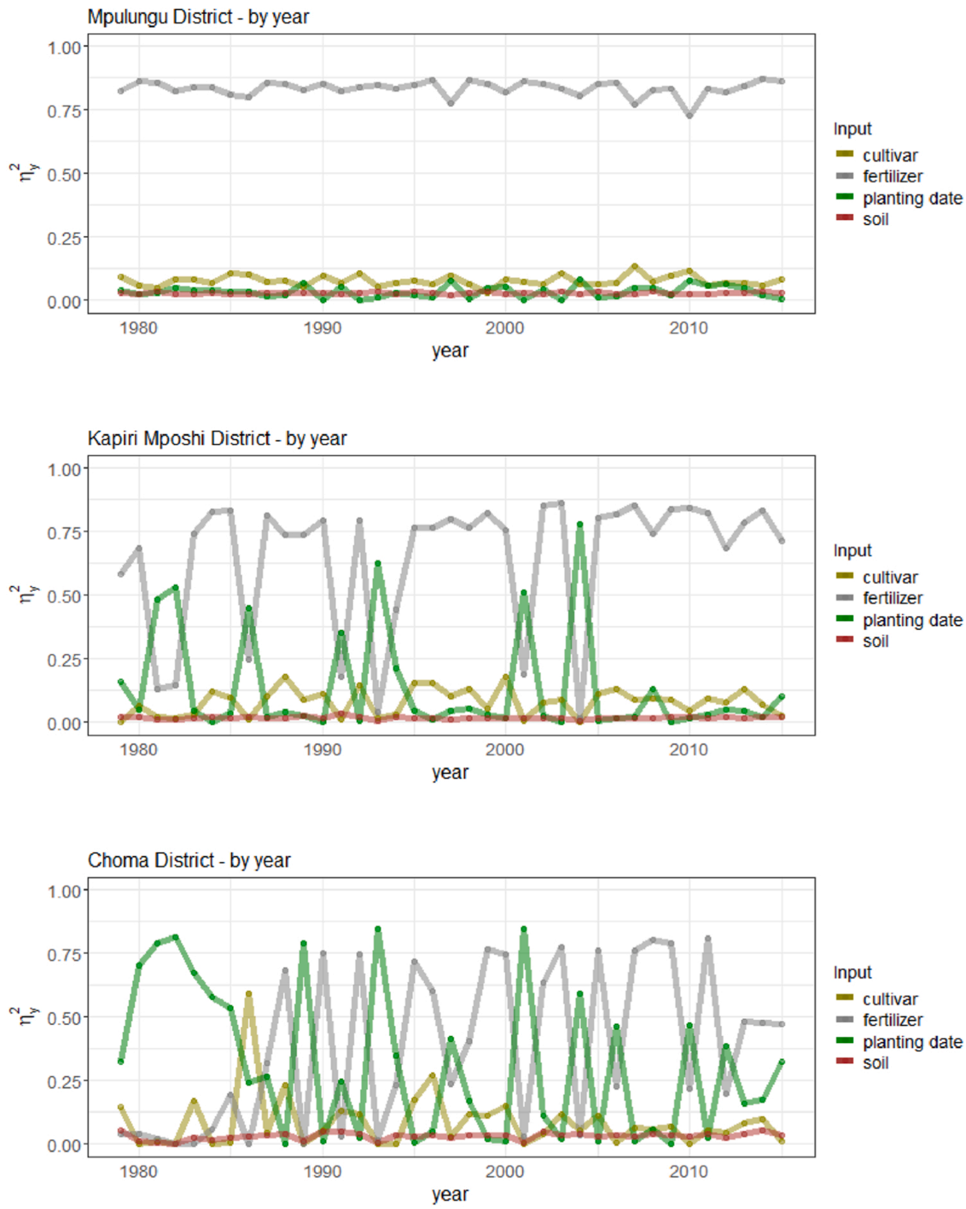

Fig. 8. Yearly effect sizes in Mpulungu (Region III, northern Zambia), Kapiri Mposhi (Region II, central Zambia), and Choma (Region I, southern Zambia) districts, listed by year.

inputs. Prior studies quantifying yield variance often look at a single factor (often climate, e.g. Lobell and Field, 2007; Iizumi and Ramankutty, 2016), while those that considered management may not evaluate specific inputs (e.g. (Lobell and Asner, 2003; Ben-Ari and Makowski, 2014)), or instead assumed that variance unexplained by soil and/or weather was attributable to management (Lobell et al., 2002).

Second, this study quantifies soil-weather-management interaction effects, including correlations between agriculturally targeted meteorological variables and yearly effect sizes. Other studies that considered interaction effects focused on a specific type of interaction, such as soil fertility and fertilizer (Burke et al., 2017), or cultivar and fertilizer (Sileshi et al., 2010), and did not quantify the effects of multiple interactions.

Third, this study quantifies yield variance at a national level, using the district as the unit of analysis. This scale allows for regional trends to be readily apparent. Notably, other studies that evaluated complex soil-weather-management (and socio-economic) interactions are often based on comprehensive analysis of a small number of study sites (Chisanga et al., 2019; Dutta et al., 2020) and may be challenging to scale nationally.

The combination of these three factors allows this study to uniquely provide insights on how smallholder management decisions impact yield variance regionally, and where specific management decisions have a greater impact.

4.1. Spatial gradients of management effects

The relative importance of management variables varied substantially across Zambia, due to the different weather and soils used in each district. Fertilizer's global effect size was largest in northern districts ($\eta_g^2 = 0.72$ in Region III) but decreased in other regions ($\eta_g^2 = 0.17$, 0.40 in Regions I, II). The variable productivity of fertilizer in Zambia has been studied previously in Zambia (Burke et al., 2019, 2017), but those studies focused more on the specific relations between soil properties (e.g. ph, organic carbon) and fertilizer responsiveness, rather than analyzing spatial variation of fertilizer responsiveness. In particular, Burke found that maize yield had higher fertilizer response in less acidic

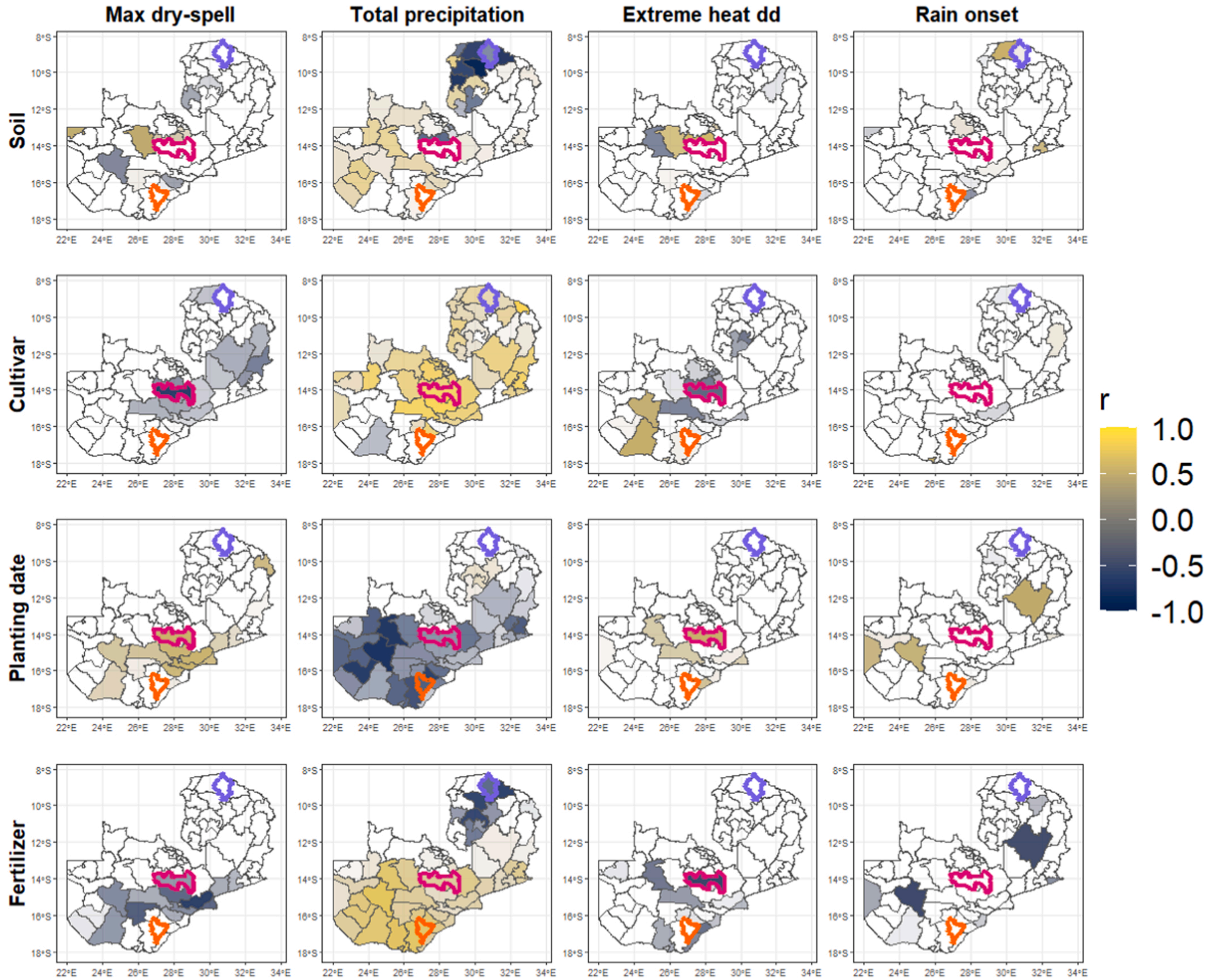

Fig. 9. Maps showing Pearson's correlation coefficients between yearly effect sizes (rows: soil, cultivar, planting date, fertilizer) and meteorological variables (columns: max dry spell length, total precipitation, extreme heat degree days, rainy season onset). Gold colors represent positive correlations and blue colors represent negative correlations. Highlighted districts are Choma (orange), Kapiri Mposhi (pink) and Mpulungu (purple). Only statistically significant ($p < 0.05$) correlations are shown.

soils (higher ph), and in soils with higher organic matter, although the organic matter impact was larger in magnitude (Burke et al., 2019, 2017). Sandy soils, more common in southern regions, also showed a lower responsiveness to fertilizer than acrisols with higher clay content (Burke et al., 2019, 2017). Among the soil properties used in simulations, we see that soils in Region III are both more acidic and have a higher organic carbon percent (Fig. S4). The higher precipitation in Region III also may lead to increased leaching of nutrients, resulting in higher importance for increased fertilizer quantities to replenish soils of nutrients (Burke et al., 2017). The increased effect size of fertilizer in further north may thus be due to a combination of factors including: (i) higher fertilizer responsiveness due to organic carbon % (mitigating effect of higher ph soils), (ii) higher fertilizer responsiveness due to increased precipitation and availability of soil water, (iii) more consistent soil/weather constraints leading to lower yield variability due to soil and weather, (iv) leaching of nutrients increasing fertilizer importance and (v) more favorable soil/weather constraints leading to higher yields and allowing fertilizer to bring yields closer to their potential.

The other two management inputs, cultivar and planting date, had far smaller effect sizes, representing 3–4 % of yield variance among all districts (Table 4). The decreased importance of planting date in northern districts ($\eta_g^2 = 0.01$ in Region III as opposed to 0.06, 0.06 in Regions I, II) is likely due to the more consistent rains in Region III. The importance of planting date in Region I matches previous studies on Choma district, where optimal planting dates based on rainy season onset vary from late November to early January (Chabala et al., 2013) and farmers planting across a broader range of dates (Waldman et al., 2017).

Cultivar selection also explained a higher percent of yield variance in Region III. We selected three medium-maturing cultivars due to their previous calibration for DSSAT and the common use of medium-maturing cultivars throughout Zambia (Blekking et al., 2021). However, as Blekking notes, and as evidenced in other studies (Waldman et al., 2017), early-maturing cultivars are increasingly supplanting medium-maturing cultivars even when a region has sufficient climate conditions to support higher-yielding, later maturing cultivars. This shift could be due to economic necessities of farmers to plant and harvest early to provide food during the lean season (Thierfelder et al., 2016). Most of Zambia has adequate rainfall to support growing medium-maturity cultivars (Blekking et al., 2021), and these cultivars are expected to produce higher yields than earlier-maturing cultivars.

We thus attribute the increased importance of cultivar in more northern districts to the favorable conditions in Region III allowing better-performing cultivars to achieve higher yields, creating increased differentiation among cultivar performance. The importance of cultivar selection (and data selection generally) is discussed in Section 4.3. This section includes results from an alternative variance decomposition using a different set of cultivars.

Overall, management explained a far higher percent of yield variance in Region III (82 %), compared to regions I and II (27 % and 50 %), largely due to fertilizer application. These results match the common narrative of soil and weather constraints in Zambia, that districts in Region III have better conditions for longer maturing cultivars with higher yield potential. Soil and weather explain a lesser percentage of yield variance, although their effect sizes both increase further south, with weather explaining 30 % of yield variance in Region I. This larger effect size for weather in Region I is likely due to the lower and less reliable precipitation during the growing season. Soil's effect size is relatively low in most districts, representing less than 10 % of yield variance explained in 52 of 72 districts (72 %). Districts with higher soil effect size (above 10 %) are not geographically clustered, including districts in far western, southern, and eastern Zambia. The greater effect size for soil in these districts is likely due to higher within-district variability of soil properties, increasing yield variance due to soil type. _Importantly, the soil and weather data sets are by necessity derived from gridded data products, and thus may not reflect the local (subpixel) variability in soil fertility and weather that can influence yield.

4.2. Soil-weather-management interactions

4.2.1. Impact of different fertilizer rates

We also examined how effect sizes varied when fertilizer rates were fixed. As an increase in fertilizer rate should generally increase yields, farmers can be expected to increase fertilizer application when possible. However, farmers may not receive timely deliveries of fertilizer through the FISP subsidy program (Mubanga and Ferguson, 2017), or may have other constraints that limit the amount of fertilizer they apply. It is thus reasonable to consider the causes of yield variance when the fertilizer level is fixed at different levels (low, medium, and high levels).

As fertilizer rate increased, the relative effect sizes of other inputs varied (Fig. 6). The importance of weather consistently increased in all three regions, while soil's effect consistently decreased. The decrease in soil importance at high fertilizer levels has been seen previously in a crop model-based study (Folberth et al., 2016). These trends can be explained as increased fertilizer rates allowed yields to overcome soil constraints, while optimal weather allowed fertilizer to achieve higher yields.

Increasing fertilizer rate had a varied impact on other management effects, depending on region. Planting date's effect generally increased with higher fertilizer in Regions I and II while it decreased in Region III. This increased importance of planting date in Region III for low fertilizer may be due to the importance of timing planting correctly, in order to allow small quantities of fertilizer to have a positive effect on grain yield under good soil/weather conditions.

Cultivar's effect size decreased with higher fertilizer rate in Regions I and II but was largest at medium fertilizer levels in Region III. The peak in cultivar effect at medium levels of fertilizer may indicate that, at high levels of fertilizer, weather's impact simply dominates yield variance.

4.2.2. Weather-management correlations

The weather correlation analysis showed several clear patterns in terms of how meteorological variables correlated with yearly effect sizes. First, correlations with total precipitation were the most widespread for all input effect sizes. These results match a previous study on sorghum yield in West Africa, which found that total precipitation impacted yields to a greater degree than the timing of rain onset or the intensity of rains (Guan et al., 2015).

The associated decline in planting date importance and increase in fertilizer importance with increased total precipitation in much of central and southern Zambia (Regions I and II) showed how improved weather conditions (higher rainfall) alter the relative importance of these two key management variables. This pattern is visible more generally, as planting date and fertilizer had opposite sign correlations with meteorological variables (rows 3 and 4 in Fig. 9). Generally, an increase in planting date's importance was correlated with more adverse weather conditions (lower total precipitation, longer maximum dry spell, more heat stress days, later onset of rains), while an increase in fertilizer's importance correlated with more favorable conditions.

Other variables (cultivar, soil) had more localized correlations with meteorological variables. Cultivar showed a decrease in importance when the maximum dry spell length was longer for some districts in central Zambia. Previous studies have highlighted the importance of maize cultivars resistant to heat stress in Africa (Lobell et al., 2011), however there are no widespread correlations (positive or negative) between the number of heat stress days and cultivar's effect size. A few districts in central Zambia have a significant negative correlation between cultivar effect and heat stress days, indicating that increased heat stress correlates with decreased yield variance due to cultivar. The absence of correlations for cultivar may in part be due to the relatively similar potential yield for the three cultivars, and the similar performance of these cultivars under different meteorological conditions. As discussed in the next section, cultivar's importance can increase when cultivars have a larger difference in potential yield.

Understanding soil's impact on yield variance was more challenging due to both soil's lower overall effect on yield, and the differing number of soil profiles used in each district (between 4 and 30 for each district). Soil only had noticeable correlations with total precipitation. Some districts in western Zambia had positive correlations between total precipitation and soil's effect, indicating that in rainier years, soil was likely to have a larger impact in yield. Some of these districts had low clay content (Fig. S4, bottom), so the increased soil importance may represent variability in soil texture in these districts, with sandy soils less able to retain water in higher precipitation seasons. One such district in western Zambia, Kaoma, had the lowest water retention capacity compared to districts in central Zambia (Cornelissen et al., 2013). Some northeastern districts had negative correlations, indicating that increased total precipitation correlated with an increased importance of soil while some western districts had positive correlations, indicating the opposite effect. One possible explanation is that when total precipitation is higher, all soils in these districts retain sufficient water for crops, decreasing the difference in productivity among soil profiles.

4.3. Limitations

Our study had several limitations due to the limitations of crop modeling and selecting representative data sets for management, soil, and weather. First, our analysis did not consider economic costs, such as agricultural inputs (cost of fertilizer, seed, labor, etc.). Each management option in the experiment design is treated as equally likely, although in reality farmer decision-making is complex, and farmers may make decisions not designed for yield-optimization. Research in Zambia has found links between higher socioeconomic well-being and hybrid seed use (Smale and Mason, 2014), and that fertilizer subsidies for the FISP subsidy program may more often be allocated to farmers with larger holdings (Mason and Jayne, 2013). The analysis also did not consider farmer motivation to plant and harvest early, which may be necessary to provide food during the lean season (Thierfelder et al., 2016). Rather, the analysis is based upon final yield at maturity. Effect sizes are also based on potential yield using the prescribed management and may not be identical with those seen for observed yields, where a wide variety of factors may affect productivity (e.g. pests, disease, early harvest). Further, we also did not consider irrigation, an important constraint to agricultural intensification (Jayne et al., 2014) in Africa.

Irrigation will also be increasingly important for sub-Saharan Africa to become food self-sufficient (Ittersum et al., 2016), although as of 2017, only a small percent (~1 %) of Zambia smallholder farmers used irrigation on field crops (Ngoma et al., 2019), and thus focusing on rainfed agriculture is not a major limitation of this study.

Data selection has important ramifications for crop model-based studies. This study used management inputs representative of smallholder farmers, corroborated by local expert knowledge. The crop model simulations are sensitive to selection of inputs, and results for the variance decomposition would change if a different set of inputs for fertilizer, cultivar, or other factors were used. To examine the sensitivity of the variance decomposition to input choices, we used a separate set of trial data from field sites in three districts (Choma, Kafue, Kabwe) to calibrate early, medium, and late-maturing (EML) cultivars. We chose not to use these three cultivars in the main results as the data collected were not sufficiently detailed (e.g. only limited soil properties measured at a single depth). However, the effect sizes for these cultivars illustrate the importance of considering management inputs when interpreting effect sizes, and that using cultivars with a large difference in potential yields will increase the effect size for cultivar.

The variance decomposition for the EML cultivars showed a much larger cultivar effect size (mean cultivar $\eta_g^2 = 0.31$ with the EML cultivars, as compared to 0.04 for the main set of cultivars). Other spatial patterns resembled those in this study (e.g. weather more important in southern districts, fertilizer in northern districts, see Fig. S5). The increased effect size for cultivar is likely due to a wider range in cultivar parameters (akin to those in Table 2) and the higher yield potential for medium and late maturing cultivars in the EML set. We considered using three cultivars with large differences in yield potential an unrealistic choice for smallholder farmers. However, choosing among three cultivars with similar (but not identical) maturity and yield potential represents a realistic selection dilemma for smallholders in a market saturated with different cultivars (Waldman et al., 2017, 2019; Blekking et al., 2021).

Negative effects of excessive soil water on crop yield are also important to consider. Even in areas with low or moderate annual precipitation (< 1000 mm / year) (e.g. cowpea crops in Sudan savanna (Iseki et al., 2021), negative effects of excessive water can still occur which are not always captured by crop models (Li et al., 2019). We performed a correlation analysis comparing mean yearly growing season precipitation for each district with mean yearly crop yields (averaged across soil and management permutations). The results (Fig. S6) show a clear south-north gradient, with significant positive correlations in drier conditions of southern Zambia (Region I), and significant negative correlations in wetter conditions of northern Zambia. It should be noted that DSSAT did capture some negative impacts of excessive precipitation in Li's study (based in the United States), although not to the full extent seen in observed data (Li et al., 2019). Further field trials would be needed to determine if DSSAT fully captures these negative effects to the extent seen in Zambia.

Given the lack of weather stations and extensive soil sampling in SSA (Dobardzic et al., 2019), it was not possible to use ground or station measured soil and weather data. We thus used gridded soil and weather data with relatively fine spatial resolutions (~10 km) that were used in prior crop modeling studies. The soil database (Han et al., 2019) is a state-of-the-art gridded data set based on the SoilGrids 1 km data set (Hengl et al., 2014), which integrates an extensive soil sample databases with satellite-based covariates, and is designed for use in the DSSAT crop model. However, it does not include estimates of phosphorus or potassium, and generally may not capture the fine-scale gradients in fertility between different farmer fields (Tittonell et al., 2008). The soil effects thus should be interpreted as interpreting how soil fertility changes within a district, and not how historical management may have improved or degraded soils locally.

Similarly, the gridded weather data sets were selected due to their relatively fine spatial resolution and their accuracy. MSWEP precipitation data performed best in terms of correlation with gauge precipitation measurements (Beck et al., 2017b). Like soil data, however, the weather data do not fully capture sub-grid variability in weather conditions, which can be substantial in agricultural systems with high weather variability (Baron et al., 2005). This study does use soil and weather grids with approximately a 10 km grid cell, a substantial improvement on previous products (e.g. General Circulation Models with 200 km resolution). As mentioned above, improved soil measurements and weather station observations can be used as inputs provided these data are available. Continued improvements in local data collection are crucial for accurately modeling soil and weather constraints in smallholder agriculture.

5. Conclusions

A few general conclusions emerge from our study. In total, management explained 53 % of yield variance, although this percentage increased greatly in northern districts with higher precipitation and more favorable growing conditions, in large part due to the effect of fertilizer. Management choices explained 27 %, 50 %, and 82 % of yield variance in different agro-ecological regions Soil and weather constraints also affected which management interventions were more important. Generally, fertilizer's effect was largest in more favorable growing conditions, explaining 45 % of the yield variability on average, while planting date's effect was larger in less favorable conditions (lower precipitation, higher heat stress), explaining 5–6 % of yield variance in southern and central agro-ecological regions. The effect of cultivar selection was more muted on average (4 %) due to the relatively similar potential yields of the three cultivars selected. The soil effect was largest in southern and central agro-ecological regions, where soil explained 8–9 % of district-scale yield variance.

Effective policy interventions require an understanding of yield variance at a variety of spatial scales, from field and farm to district and region. Zambia has a substantial farm subsidy program providing fertilizer and hybrid seed to smallholder farmers, often with recommendations for management at the agro-ecological region level. The approach and results of this study can inform where particular management interventions have a greater impact, both at the regional and district level and provide policymakers with actionable insights on yield variance that can be complemented with studies at a local scale.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This work was supported by the National Science Foundation (grant nos. SES-1534544, SES-1360463, DEB-1924309).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fcr.2023.109014.

References

Adnan, A.A., Diels, J., Jibrin, J.M., Kamara, A.Y., Shaibu, A.S., Garba, I.I., Craufurd, P., Maertens, M., 2020. Optimizing sowing density-based management decisions with different nitrogen rates on smallholder maize farms in Northern Nigeria. Exp Agr 56, 866–883. https://doi.org/10.1017/s001447972000037x.

Banziger, M., Edmeades, G.O., Beck, D.L., Bellon, M.R., 2000. Breeding for Drought and Nitrogen Stress Tolerance in Maize: From Theory to Practice. CIMMYT, Mexico, D.F.

Baron, C., Sultan, B., Balme, M., Sarr, B., Traore, S., Lebel, T., Janicot, S., Dingkuhn, M., 2005. From GCM grid cell to agricultural plot: scale issues affecting modelling of climate impact. Philos. Trans. R. Soc. B Biol. Sci. 360, 2095–2108. https://doi.org/10.1098/rstb.2005.1741.

Beck, H.E., Dijk, A.I.J.M., van, Levizzani, V., Schellekens, J., Miralles, D.G., Martens, B., Roo, A. de, 2017a. MSWEP: 3-hourly 0.25° global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. Hydrol. Earth Syst. ScI. 21, 589–615. https://doi.org/10.5194/hess-21-589-2017.

Beck, H.E., Vergopolan, N., Pan, M., Levizzani, V., Dijk, A.I.J.M., van, Weedon, G.P., Brocca, L., Pappenberger, F., Huffman, G.J., Wood, E.F., 2017b. Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. Hydrol. Earth Syst. ScI. 21, 6201–6217. https://doi.org/10.5194/hess-21-6201-2017.

Ben-Ari, T., Makowski, D., 2014. Decomposing global crop yield variability. Environ. Res Lett. 9, 114011 https://doi.org/10.1088/1748-9326/9/11/114011.

Blekking, J., Waldman, K.B., Evans, T., 2021. Hybrid-maize seed certification and smallholder adoption in Zambia. J. Environ. Plann Man 64, 359–377. https://doi.org/10.1080/09640568.2020.1764342.

Burke, M., Lobell, D.B., 2017. Satellite-based assessment of yield variation and its determinants in smallholder African systems. Proc. Natl. Acad. Sci. 114, 2189–2194. https://doi.org/10.1073/pnas.1616919114.

Burke, W.J., Jayne, T.S., Black, R.J., 2012. Getting More "Bang for the Buck": Diversifying Subsidies Beyond Fertilizer and Policy Beyond Subsidies. FSRP Policy Synth. 52.

Burke, W.J., Jayne, Thom, S., Black, J.R., 2017. Factors explaining the low and variable profitability of fertilizer application to maize in Zambia. Agr. Econ. 48, 115–126. https://doi.org/10.1111/agec.12299.

Burke, W.J., Frossard, E., Kabwe, S., Jayne, T.S., 2019. Understanding fertilizer adoption and effectiveness on maize in Zambia. Food Policy 86, 101721. https://doi.org/10.1016/j.foodpol.2019.05.004.

Burney, J.A., Naylor, R.L., 2012. Smallholder irrigation as a poverty alleviation tool in Sub-Saharan Africa. World Dev. 40, 110–123. https://doi.org/10.1016/j.worlddev.2011.05.007.

Carletto, C., Gourlay, S., Murray, S., Zezza, A., 2016. Cheaper, Faster, and More Than Good Enough: Is GPS the New Gold Standard in Land Area Measurement? https://doi.org/10.1596/1813–9450-7759.

Chabala, L.M., Kuntashula, E., Kaluba, P., 2013. Characterization of temporal changes in rainfall, temperature, flooding hazard and dry spells over Zambia. Univers J. Agric. Res 1, 134–144. https://doi.org/10.13189/ujar.2013.010403.

Chabala, L.M., Kuntashula, E., Kaluba, P., Miyanda, M., 2015. Assessment of maize yield variations due to climatic variables of rainfall and temperature. J. Agr. Sci. 7. https://doi.org/10.5539/jas.v7n11p143.

Chamberlin, J., Jayne, T.S., Headey, D., 2014. Scarcity amidst abundance? Reassessing the potential for cropland expansion in Africa. Food Policy 48, 51–65. https://doi.org/10.1016/j.foodpol.2014.05.002.

Chisanga, C.B., Phiri, E., Chinene, V.R.N., 2019. Evaluation of sowing date and fertilization with nitrogen in maize cultivars in rainy conditions in Zambia. Afr. J. Plant Sci. 13, 221–230. https://doi.org/10.5897/ajps2019.1792.

Chisanga, C.B., Phiri, E., Chinene, V.R.N., 2021. Evaluating APSIM-and-DSSAT-CERES-maize models under rainfed conditions using zambian rainfed maize cultivars. Nitrogen 2, 392–414. https://doi.org/10.3390/nitrogen2040027.

Chisanga, C.B., Phiri, E., Shepande, C., Sichingabula, H., 2015. Evaluating CERES-Maize Model Using Planting Dates and Nitrogen Fertilizer in Zambia. J. Agr. Sci. 7. https://doi.org/10.5539/jas.v7n3p79.

Chisanga, C.B., Phiri, E., Chinene, V.R.N., Chabala, L.M., 2020. Projecting maize yield under local-scale climate change scenarios using crop models: sensitivity to sowing dates, cultivar, and nitrogen fertilizer rates. Food Energy Secur 9. https://doi.org/10.1002/fes3.231.

CIAT; World Bank. 2017. Climate-Smart Agriculture in Zambia. CSA Country Profiles for Africa Series. International Center for Tropical Agriculture (CIAT), Washington, D.C. 25 p.

Cohen, J., 1973. Eta-Squared and Partial Eta-Squared in Fixed Factor Anova Designs. Educ. Psychol. Meas. 33, 107–112. https://doi.org/10.1177/001316447303300111.

Cooper, P.J.M., Dimes, J., Rao, K.P.C., Shapiro, B., Shiferaw, B., Twomlow, S., 2008. Coping better with current climatic variability in the rain-fed farming systems of sub-Saharan Africa: An essential first step in adapting to future climate change? Agric. Ecosyst. Environ. 126, 24–35. https://doi.org/10.1016/j.agee.2008.01.007.

Corbeels, M., Chirat, G., Messad, S., Thierfelder, C., 2016. Performance and sensitivity of the DSSAT crop growth model in simulating maize yield under conservation agriculture. Eur. J. Agron. 76, 41–53. https://doi.org/10.1016/j.eja.2016.02.001.

Cornelissen, G., Martinsen, V., Shitumbanuma, V., Alling, V., Breedveld, G.D., Rutherford, D.W., Sparrevik, M., Hale, S.E., Obia, A., Mulder, J., 2013. Biochar effect on maize yield and soil characteristics in five conservation farming sites in Zambia. Agronomy 3, 256–274. https://doi.org/10.3390/agronomy3020256.

Deininger, K., Byerlee, D., 2011. Rising Global Interest in Farmland. https://doi.org/10.1596/978–0-8213–8591-3.

Dobardzic, S., Dengel, C., Gomes, A.M., Hansen, J., Bernardi, M., Fujisawa, M., Heureux, A.M., Kanamaru, H., Neretin, L., Rojas, O., Intsiful, J., Barnwal, A., Iqbal, F., Kull, D., Bogdanova, A.M., Fara, K., Pergolini, G., Aich, B., Alexieva, A., Baddour, O., Delju, A., Coning, E.D., Devillier, R., Eggleston, S., Gallo, I., Harou, A., Hechler, P., Hovsepyan, A., Jepsen, L.A., Okia, W.M., Msemo, N., P, P., Richter, C., Riishojgaard, L.P., Rixen, M., Ruti, P., Santamaria, L., Stefanski, R., Watkins, J., Wright, W., 2019. 2019 State of Climate Services: Agriculture and Food Security. World Meteorological Organization (WMO), Geneva, Switzerland.

Dutta, S., Chakraborty, S., Goswami, R., Banerjee, H., Majumdar, K., Li, B., Jat, M.L., 2020. Maize yield in smallholder agriculture system-an approach integrating socio-economic and crop management factors. Plos One 15, e0229100. https://doi.org/10.1371/journal.pone.0229100.

Edreira, J.I.R., Guilpart, N., Sadras, V., Cassman, K.G., Ittersum, M.K., van, Schils, R.L.M., Grassini, P., 2018. Water productivity of rainfed maize and wheat: a local to global perspective. Agr. For. Meteor. 259, 364–373. https://doi.org/10.1016/j.agrformet.2018.05.019.

Elliott, J., Müller, C., Deryng, D., Chryssanthacopoulos, J., Boote, K.J., Büchner, M., Foster, I., Glotter, M., Heinke, J., Iizumi, T., Izaurralde, R.C., Mueller, N.D., Ray, D. K., Rosenzweig, C., Ruane, A.C., Sheffield, J., 2014. The Global Gridded Crop Model Intercomparison: data and modeling protocols for Phase 1 (v1.0). Geosci. Model Dev. 8, 261–277. https://doi.org/10.5194/gmd-8-261-2015.

Estes, L.D., Beukes, H., Bradley, B.A., Debats, S.R., Oppenheimer, M., Ruane, A.C., Schulze, R., Tadross, M., 2013. Projected climate impacts to South African maize and wheat production in 2055: a comparison of empirical and mechanistic modeling approaches. Glob. Change Biol. 19, 3762–3774. https://doi.org/10.1111/gcb.12325.

Folberth, C., Skalský, R., Moltchanova, E., Balkovič, J., Azevedo, L.B., Obersteiner, M., Velde, M. van der, 2016. Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. Nat. Commun. 7, 11872. https://doi.org/10.1038/ncomms11872.

Food and Agriculture Organization (FAO) of the U.N**., 2021. Crops and Livestock Products [WWW Document] . URL http://www.fao.org/faostat/en/#data/QCL (accessed 8.18.21)**.

Frieler, K., Schauberger, B., Arneth, A., Balkovič, J., Chryssanthacopoulos, J., Deryng, D., Elliott, J., Folberth, C., Khabarov, N., Müller, C., Olin, S., Pugh, T.A.M., Schaphoff, S., Schewe, J., Schmid, E., Warszawski, L., Levermann, A., 2017. Understanding the weather signal in national crop-yield variability: WEATHER-SIGNAL IN CROP-YIELD VARIABILITY. Earth's Futur 5, 605–616. https://doi.org/10.1002/2016ef000525.

Garcia, A.G., Guerra, L.C., Hoogenboom, G., 2009. Impact of planting date and hybrid on early growth of sweet corn. Agron. J. 101, 193–200. https://doi.org/10.2134/agronj2007.0393.

Giroux, S.A., Kouper, I., Estes, L.D., Schumacher, J., Waldman, K., Greenshields, J.T., Dickinson, S.L., Caylor, K.K., Evans, T.P., 2019. A high-frequency mobile phone data collection approach for research in social-environmental systems: applications in climate variability and food security in Sub-Saharan Africa. Environ. Model. Softw. 119, 57–69. https://doi.org/10.1016/j.envsoft.2019.05.011.

Glotter, M., Elliott, J., 2016. Simulating US agriculture in a modern Dust Bowl drought. Nat. Plants 3, 16193. https://doi.org/10.1038/nplants.2016.193.

Gourlay, S., Kilic, T., Lobell, D.B., 2019. A new spin on an old debate: Errors in farmer-reported production and their implications for inverse scale - Productivity relationship in Uganda. J. Dev. Econ. 141, 102376 https://doi.org/10.1016/j.jdeveco.2019.102376.

Grassini, P., Bussel, L.G.J., van, Wart, J.V., Wolf, J., Claessens, L., Yang, H., Boogaard, H., Groot, H., de, Ittersum, M.K., van, Cassman, K.G., 2015. How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. Field Crop Res 177, 49–63. https://doi.org/10.1016/j.fcr.2015.03.004.

Guan, K., Sultan, B., Biasutti, M., Baron, C., Lobell, D.B., 2015. What aspects of future rainfall changes matter for crop yields in West Africa? Geophys Res Lett. 42, 8001–8010. https://doi.org/10.1002/2015gl063877.

Guan, K., Sultan, B., Biasutti, M., Baron, C., Lobell, D.B., 2017. Assessing climate adaptation options and uncertainties for cereal systems in West Africa. Agr. For. Meteor. 232, 291–305. https://doi.org/10.1016/j.agrformet.2016.07.021.

Guilpart, N., Grassini, P., Wart, J. van, Yang, H., Ittersum, M.K., van, Bussel, L.G.J., van, Wolf, J., Claessens, L., Leenaars, J.G.B., Cassman, K.G., 2017. Rooting for food security in Sub-Saharan Africa. Environ. Res Lett. 12, 114036 https://doi.org/10.1088/1748-9326/aa9003.

Hadunka, P., Bayliss, K., 2022. Staple crop pest damage and natural resources exploitation: fall army worm infestation and charcoal production in Zambia, in: Agricultural Applied Economics Association Annual Meeting. https://doi.org/https://hdl.handle.net/2142/106355.

Han, E., Ines, A.V.M., Koo, J., 2019. Development of a 10-km resolution global soil profile dataset for crop modeling applications. Environ. Model. Softw. 119, 70–83. https://doi.org/10.1016/j.envsoft.2019.05.012.

Harris, D., Orr, A., 2014. Is rainfed agriculture really a pathway from poverty? Agr. Syst. 123, 84–96. https://doi.org/10.1016/j.agsy.2013.09.005.

Hengl, T., Jesus, J.M., de, MacMillan, R.A., Batjes, N.H., Heuvelink, G.B.M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Gonzalez, M.R., 2014. SoilGrids1km — Global Soil Information Based on Automated Mapping. Plos One 9, e105992. https://doi.org/10.1371/journal.pone.0105992.

Hoogenboom, G., Porter, C.H., Boote, K.J., Shelia, V., Wilkens, P.W., Singh, U., White, J. W., Asseng, S., Lizaso, J.I., Moreno, L.P., Pavan, W., Ogoshi, R., Hunt, L.A., Tsuji, G. Y., Jones, J.W., 2019b. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.7.5. DSSAT Foundation, Gainesville, Florida, USA.

Hoogenboom, Gerrit, Boote, K., Porter, C.H., Boote, K.J., Shelia, V., Wilkens, P.W., USA, I.F.D.C., Singh, U., White, J.W., USA, U.-A., Asseng, S., Lizaso, J.I., Spain, U.P. de M., Moreno, L.P., Pavan, W., Brazil, U. de P.F., Ogoshi, R., USA, U. of H., Hunt, L. A., Canada, U. of G., Tsuji, G.Y., Jones, J.W., 2019a. The DSSAT crop modeling ecosystem. In: Boote, K. (Ed.), Advances in Crop Modeling for a Sustainable Agriculture. Burleigh Dodds Series in Agricultural Science, pp. 173–216. https://doi.org/10.19103/as.2019.0061.10.

Iizumi, T., Ramankutty, N., 2016. Changes in yield variability of major crops for 1981–2010 explained by climate change. Environ. Res Lett. 11, 034003 https://doi.org/10.1088/1748-9326/11/3/034003.

Iseki, K., Ikazaki, K., Batieno, J.B., 2021. Cowpea yield variation in three dominant soil types in the Sudan Savanna of West Africa. Field Crop Res. 261, 108012 https://doi.org/10.1016/j.fcr.2020.108012.

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

Ittersum, M.K., van, Bussel, L.G.J., van, Wolf, J., Grassini, P., Wart, J., van, Guilpart, N., Claessens, L., Groot, H., de, Wiebe, K., Mason-D'Croz, D., Yang, H., Boogaard, H., Oort, P.A.J., van, Loon, M.P., van, Saito, K., Adimo, O., Adjei-Nsiah, S., Agali, A., Bala, A., Chikowo, R., Kaizzi, K., Kouressy, M., Makoi, J.H.J.R., Ouattara, K., Tesfaye, K., Cassman, K.G., 2016. Can sub-Saharan Africa feed itself? Proc. Natl. Acad. Sci. 113, 14964–14969. https://doi.org/10.1073/pnas.1610359113.

Jain, M., Singh, B., Srivastava, A.A.K., Malik, R.K., McDonald, A.J., Lobell, D.B., 2017. Using satellite data to identify the causes of and potential solutions for yield gaps in India's Wheat Belt. Environ. Res Lett. 12, 094011 https://doi.org/10.1088/1748-9326/aa8228.

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

Jenks, G.F., 1977. Optimal data classification for choropleth maps. University of Kansas, Lawrence, KS.

Jin, Z., Azzari, G., You, C., Tommaso, S.D., Aston, S., Burke, M., Lobell, D.B., 2019. Smallholder maize area and yield mapping at national scales with Google Earth Engine. Remote Sens Environ. 228, 115–128. https://doi.org/10.1016/j.rse.2019.04.016.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron. 18, 235–265. https://doi.org/10.1016/s1161-0301(02)00107-7.

Kersebaum, K.C., Boote, K.J., Jorgenson, J.S., Nendel, C., Bindi, M., Frühauf, C., Gaiser, T., Hoogenboom, G., Kollas, C., Olesen, J.E., Rötter, R.P., Ruget, F., Thorburn, P.J., Trnka, M., Wegehenkel, M., 2015. Analysis and classification of data sets for calibration and validation of agro-ecosystem models. Environ. Model. Softw. 72, 402–417. https://doi.org/10.1016/j.envsoft.2015.05.009.

Li, Y., Guan, K., Schnitkey, G.D., DeLucia, E., Peng, B., 2019. Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. Glob. Change Biol. 25, 2325–2337. https://doi.org/10.1111/gcb.14628.

Lobell, D.B., Asner, G.P., 2003. Climate and management contributions to recent trends in U.S. agricultural yields, 1032–1032 Science 299. https://doi.org/10.1126/science.1077838.

Lobell, D.B., Field, C.B., 2007. Global scale climate–crop yield relationships and the impacts of recent warming. Environ. Res Lett. 2, 014002 https://doi.org/10.1088/1748-9326/2/1/014002.

Lobell, D.B., Cassman, K.G., Field, C.B., 2009. Crop yield gaps: their importance, magnitudes, and causes. Annu. Rev. Env Resour. 34, 179–204. https://doi.org/10.1146/annurev.environ.041008.093740.

Lobell, D.B., Ortiz-Monasterio, J.I., Addams, C.L., Asner, G.P., 2002. Soil, climate, and management impacts on regional wheat productivity in Mexico from remote sensing. Agr. For. Meteor. 114, 31–43. https://doi.org/10.1016/s0168-1923(02)00138-7.

Lobell, D.B., Bänziger, M., Magorokosho, C., Vivek, B., 2011. Nonlinear heat effects on African maize as evidenced by historical yield trials. Nat. Clim. Change 1, 42–45. https://doi.org/10.1038/nclimate1043.

Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop ield mapper. Remote Sens Environ. 164, 324–333. https://doi.org/10.1016/j.rse.2015.04.021.

Lobell, D.B., Hammer, G.L., McLean, G., Messina, C., Roberts, M.J., Schlenker, W., 2013. The critical role of extreme heat for maize production in the United States. Nat. Clim. Change 3, 497–501. https://doi.org/10.1038/nclimate1832.

Lowder, S.K., Skoet, J., Raney, T., 2016. The number, size, and distribution of farms, smallholder farms, and family farms worldwide. World Dev. 87, 16–29. https://doi.org/10.1016/j.worlddev.2015.10.041.

Mason, N.M., Jayne, T.S., 2013. Fertiliser subsidies and smallholder commercial fertiliser purchases: crowding out, leakage and policy implications for Zambia. J. Agr. Econ. 64, 558–582. https://doi.org/10.1111/1477-9552.12025.

Mdee, A., Ofori, A., Chasukwa, M., Manda, S., 2020. Neither sustainable nor inclusive: a political economy of agricultural policy and livelihoods in Malawi, Tanzania and Zambia. J. Peasant Stud. 1–24. https://doi.org/10.1080/03066150.2019.1708724.

Mubanga, K.H., Ferguson, W., 2017. Threats to food sufficiency among smallholder farmers in Choma, Zambia. Food Secur 9, 745–758. https://doi.org/10.1007/s12571-017-0700-4.

Müller, C., Cramer, W., Hare, W.L., Lotze-Campen, H., 2011. Climate change risks for African agriculture. Proc. Natl. Acad. Sci. 108, 4313–4315. https://doi.org/10.1073/pnas.1015078108.

Ngoma, H., Hamududu, B., Hangoma, P., Samboko, P., Hichaambwa, M., Kabaghe, C., 2019. Irrigation development for climate resilience in Zambia: The known knowns and known unknowns. Feed the Future Innovation Lab for Food Security Policy. Research Paper 144.

Paliwal, A., Jain, M., 2020. The accuracy of self-reported crop yield estimates and their ability to train remote sensing algorithms. Front. Sustain Food Syst. 4, 25. https://doi.org/10.3389/fsufs.2020.00025.

Ricciardi, V., Ramankutty, N., Mehrabi, Z., Jarvis, L., Chookolingo, B., 2018. How much of the world's food do smallholders produce? Glob. Food Secur. 17, 64–72. https://doi.org/10.1016/j.gfs.2018.05.002.

Searchinger, T.D., Estes, L., Thornton, P.K., Beringer, T., Notenbaert, A., Rubenstein, D., Heimlich, R., Licker, R., Herrero, M., 2015. High carbon and biodiversity costs from converting Africa's wet savannahs to cropland. Nat. Clim. Chang. 5, 481–486. https://doi.org/10.1038/nclimate2584.

Sheahan, M., Barrett, C.B., 2017. Ten striking facts about agricultural input use in Sub-Saharan Africa. Food Policy 67, 12–25. https://doi.org/10.1016/j.foodpol.2016.09.010.

Sheffield, J., Goteti, G., Wood, E.F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. J. Clim. 19, 3088–3111. https://doi.org/10.1175/jcli3790.1.

Shelia, V., Hansen, J., Sharda, V., Porter, C., Aggarwal, P., Wilkerson, C.J., Hoogenboom, G., 2019. A multi-scale and multi-model gridded framework for forecasting crop production, risk analysis, and climate change impact studies. Environ. Model. Softw. 115, 144–154. https://doi.org/10.1016/j.envsoft.2019.02.006.

Sileshi, G., Akinnifesi, F.K., Debusho, L.K., Beedy, T., Ajayi, O.C., Mong'omba, S., 2010. Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-Saharan Africa. Field Crop Res 116, 1–13. https://doi.org/10.1016/j.fcr.2009.11.014.

Smale, M., Mason, N., 2014. Hybrid seed and the economic well-being of smallholder maize farmers in Zambia. J. Dev. Stud. 50, 680–695. https://doi.org/10.1080/00220388.2014.887690.

Tadross, M., Suarez, P., Lotsch, A., Hachigonta, S., Mdoka, M., Unganai, L., Lucio, F., Kamdonyo, D., Muchinda, M., 2009. Growing-season rainfall and scenarios of future change in southeast Africa: implications for cultivating maize. Clim. Res 40, 147–161. https://doi.org/10.3354/cr00821.

Tesfaye, K., Sonder, K., Cairns, J., Magorokosho, C., Erenstein, O., 2016. Targeting drought-tolerant maize varieties in southern Africa: a geospatial crop modeling approach using big data. Climatic Change 1–18.

Thierfelder, C., Rusinamhodzi, L., Setimela, P., Walker, F., Eash, N.S., 2016. Conservation agriculture and drought-tolerant germplasm: reaping the benefits of climate-smart agriculture technologies in central Mozambique. Renew. Agr. Food Syst. 31, 414–428. https://doi.org/10.1017/s1742170515000332.

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

Tittonell, P., Vanlauwe, B., Corbeels, M., Giller, K.E., 2008. Yield gaps, nutrient use efficiencies and response to fertilisers by maize across heterogeneous smallholder farms of western Kenya. Plant Soil 313, 19–37. https://doi.org/10.1007/s11104-008-9676-3.

Tovihoudji, P.G., Akponikpè, P.B.I., Agbossou, E.K., Bielders, C.L., 2019. Using the DSSAT model to support decision making regarding fertilizer microdosing for maize production in the sub-humid region of Benin. Front. Environ. Sci. 7, 13. https://doi.org/10.3389/fenvs.2019.00013.

UNDP, 2010. Adaptation to the effects of drought and climate change in Agro-ecological Regions I and II in Zambia. UNDP,, Lusaka, Zambia. ⟨https://www.adaptation-undp.org/projects/ldcf-drought-zambia⟩.

Vanlauwe, B., Descheemaeker, K., Giller, K.E., Huising, J., Merckx, R., Nziguheba, G., Wendt, J., Zingore, S., 2015. Integrated soil fertility management in sub-Saharan Africa: unravelling local adaptation. Soil 1, 491–508. https://doi.org/10.5194/soil-1-491-2015.

Vergopolan, N., Xiong, S., Estes, L., Wanders, N., Chaney, N.W., Wood, E.F., Konar, M., Caylor, K., Beck, H.E., Gatti, N., Evans, T., Sheffield, J., 2021. Field-scale soil moisture bridges the spatial-scale gap between drought monitoring and agricultural yields. Hydrol. Earth Syst. Sci. 25, 1827–1847. https://doi.org/10.5194/hess-25-1827-2021.

Waldman, K.B., Blekking, J.P., Attari, S.Z., Evans, T.P., 2017. Maize seed choice and perceptions of climate variability among smallholder farmers. Glob. Environ. Change 47, 51–63. https://doi.org/10.1016/j.gloenvcha.2017.09.007.

Waldman, K.B., Vergopolan, N., Attari, S.Z., Sheffield, J., Estes, L.D., Caylor, K.K., Evans, T.P., 2019. Cognitive biases about climate variability in smallholder farming systems in Zambia Cognitive biases about climate variability in smallholder farming systems in Zambia. Weather Clim. Soc. 11, 369–383. https://doi.org/10.1175/wcas-d-18-0050.1.

Xiong, J., Thenkabail, P.S., Gumma, M.K., Teluguntla, P., Poehnelt, J., Congalton, R.G., Yadav, K., Thau, D., 2017. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. Isprs J. Photo 126, 225–244. https://doi.org/10.1016/j.isprsjprs.2017.01.019.

Yang, M., Wang, G., Ahmed, K.F., Adugna, B., Eggen, M., Atsbeha, E., You, L., Koo, J., Anagnostou, E., 2020. The role of climate in the trend and variability of Ethiopia's cereal crop yields. Sci. Total Environ. 723, 137893 https://doi.org/10.1016/j.scitotenv.2020.137893.

Zhao, Y., Vergopolan, N., Baylis, K., Blekking, J., Caylor, K., Evans, T., Giroux, S., Sheffield, J., Estes, L., 2018. Comparing empirical and survey-based yield forecasts in a dryland agro-ecosystem. Agr. For. Meteorol. 262, 147–156. https://doi.org/10.1016/j.agrformet.2018.06.024.