nature food

ARTICLES

https://doi.org/10.1038/s43016-022-00493-z



# Profitability of climate-smart soil fertility investment varies widely across sub-Saharan Africa

Ellen B. McCullough[1]✉, Julianne D. Quinn[2]✉ and Andrew M. Simons[3]✉

**Soil fertility investments in sub-Saharan Africa, where budgetary resources are scarce, must be well targeted. Using a causal forest algorithm and an experimental maize trial dataset matched with geocoded rainfall, temperature and soils data, we modelled site-specific, ex ante distributions of yield response and economic returns to fertilizer use. Yield response to fertilizer use was found to vary with growing season temperature and precipitation and soil conditions. Fertilizer use profitability—defined as clearing a 30% internal rate of return in at least 70% of the years—was robust to growing season climate and the fertilizer-to-maize price ratio in several locations but not in roughly a quarter of the analysed area. The resulting profitability-assessment tool can support decision makers when climate conditions at planting are unknown and sheds light on the profitability determinants of different regions, which is key for effective smallholder farm productivity-enhancing strategies.**

In developing countries, the majority of poor people live in rural areas, and most depend on agriculture for their livelihoods[1]. Since 1960, growth in use of inorganic fertilizer along with use of modern seeds that are responsive to fertilizer has contributed to growth in staple crop yields and improved economic indicators such as growth in gross domestic product, growth in labour productivity within and outside of agriculture and reductions in headcount poverty worldwide[2–4]. Increased input intensification, of which fertilizer is a key component, accounted for approximately 60% of gross agricultural output growth in developing countries during this period[5]. In the regions of the world that experienced a prominent Green Revolution (South Asia, East Asia and Latin America), fertilizer use intensity grew by 56×, 17× and 10×, respectively, between 1960–2010[6].

In sub-Saharan Africa, where poverty remains the most concentrated, agriculture is characterized by low yields and low use of productivity-enhancing inputs such as fertilizers and improved seeds[7]. Fertilizer use intensity in sub-Saharan Africa grew by only 5× between 1960 and 2017 from 2.9 kg ha$^{-1}$ to 17.9 kg ha$^{-1}$ (ref. [6]). A marked increase in use of improved seeds and fertilizer is seen as central to the effort to transition African smallholders from their current average staple yields (1.6 t ha$^{-1}$) to the yields that Latin American smallholders have achieved (4.9 t ha$^{-1}$) (ref. [8,9]). In recent decades, African governments have invested heavily in input subsidy programs with total annual expenditures in the region's ten largest countries ranging from US$600 million to US$1.2 billion and accounting for 14–26% of combined annual public expenditures on agriculture[10].

Despite high levels of effort to increase fertilizer use, it is not clear that these investments justify the large costs[10]. Fertilizer profitability is determined by a number of site-specific factors that vary greatly from location to location such as soil type, plot slope and aspect and prices for inputs and outputs[11,12]. Spatial variability in yield response is underestimated when the variation in agronomic conditions and price incentives are ignored. Furthermore, sub-Saharan Africa is characterized by predominantly rainfed cropping systems, which have uncertain yields due to temporal variability in precipitation and temperature. In rainfed systems, farmers must purchase inputs without knowing the weather conditions they will face during the coming growing season[13,14] and, consequently, the benefits they will receive from that investment. Predicting crop response amid this uncertainty is confounded by unobserved factors that can simultaneously influence both a farmer's decision to apply fertilizer and productivity, making it difficult for researchers to credibly estimate the crop response to fertilizer.

In this paper, we integrate extensive agronomic trial data, following ref. [15], with spatially resolved weather data and newly generated soil data to estimate the yield response to inorganic fertilizer use in maize conditional on soil characteristics and growing season climate. Our maize trial dataset spans a wide range of climate and soil conditions, allowing us to explore the heterogeneity of fertilizer response under different growing conditions. Because fertilizer treatment is experimentally assigned in these agronomic trials, we estimate a yield response that is not confounded by farmers' decisions to apply fertilizer. We use a causal forest estimation procedure to predict the yield response to fertilizer given the temperature, precipitation and soil characteristics. We then use this understanding of fertilizer response to develop a forward-looking tool to support investment planning and targeting in the face of climatic risk. We simulate fertilizer response using a 2019 synthetic weather dataset created using each location's historic climate record and temperature trend. Our profitability tool then generates site-specific probabilistic distributions of the returns to fertilizer use. This forward-looking tool allows users to visualize the probability of achieving a user-defined profitability objective, given site-specific growing conditions and stochastic realizations of climate conditions.

Approaching agriculture through a climate-smart lens entails gathering and integrating data at appropriate time frames and geographic scales to enable the creation of analytical tools that support flexible, context-specific decision making[16,17]. Decision tools such

[1]Department of Agricultural and Applied Economics, University of Georgia, Athens, GA, USA. [2]Department of Engineering Systems and Environment, University of Virginia, Charlottesville, VA, USA. [3]Department of Economics, Fordham University, Bronx, NY, USA. ✉e-mail: emccullough@uga.edu; jdq6nn@virginia.edu; asimons5@fordham.edu

as the one we propose here can support policy planners' efforts to develop interventions that are robust in the face of climate uncertainty and that account for fertilizer response heterogeneity across variable growing conditions. With better decision tools, planners can make climate-smart agricultural productivity investments that more efficiently use limited resources.

This study builds on previous attempts to understand the agronomic and/or economic effectiveness of fertilizer use in staple crop production in developing countries, which typically fall into one of four categories—researcher-managed trials, farmer-managed trials, crop modelling and observational studies, as further reviewed in Supplementary Discussion 1. Except for ref. [14], which uses fully mechanistic crop models rather than experimental trial data to model farmers' willingness to pay for fertilizer in sub-Saharan Africa, and ref. [18], which estimates fertilizer response using a farmer-managed trial in Malawi covering 1,205 sites observed over two growing seasons, previous fertilizer-targeting efforts do not condition fertilizer response on both soil and climate conditions, nor do they account for the uncertainty in weather realizations that characterizes rainfed cropping systems.

## Results

**Fertilizer yield response model.** To model fertilizer response that is heterogeneous over site-specific characteristics, management practices and climate realizations, we use a fertilizer response model that is sufficiently flexible in form to capture this heterogeneity. We estimate the yield response function, $\Delta y(\bar{\theta}, X_i, \omega_{it})$, where the yield response to fertilizer ($\Delta y$) depends on technologies and management practices employed ($\bar{\theta}$), site characteristics ($X_i$) and weather realizations ($\omega_{it}$). The full list of predictors is presented in Extended Data Table 1. By avoiding use of a restrictive functional form, which is the standard approach used to estimate production functions, we do not restrict the yield response to be constant over the domain of each predictor—variety planted, site characteristics and growing season weather. Thus, rather than assuming away the possibility of interactions that are well substantiated in the agronomic literature (for example, ref. [19]), we allow the data to uncover the relationships between the predictors and the yield response.

Because our interest is in predicting fertilizer response out of sample, in locations and years where trials have not been conducted, we pool observations across trial sites and estimate a fertilizer response model that includes between-site variation in $X_i$ and site-year variation in $\omega_{it}$. We estimate the model using a causal forest algorithm, an ensemble learning method for estimating heterogeneous treatment effects that is flexible, tends to perform well in generating predictions outside of the estimation sample and accommodates the clustering of trial data into specific locations and years[20,21].

Fertilizer is a binary treatment in our estimation because the vast majority of the trial dataset observations tested only two fertilizer treatments (plots fertilized with 125 kg N ha$^{-1}$, which we refer to as optimal fertilizer, and plots without added nitrogen, which we refer to as no-fertilizer), as described further in Methods. The conditional average treatment effect (yield response to fertilizer use) is 1.82 t ha$^{-1}$ with a standard error of 0.29. This response corresponds with an agronomic use efficiency of 14.5 kg of grain per kg added nitrogen. The average trial site-level predicted fertilizer yield response is 1.49 t ha$^{-1}$, though the site-level average response ranges from 0.44 t ha$^{-1}$ in the lowest response site to 3.19 t ha$^{-1}$ in the highest.

To better understand how the yield difference between fertilized and non-fertilized plots (hereafter, the fertilizer yield response) varies under different growing conditions, we predict the fertilizer response for each trial location over 1,000 growing season temperature and precipitation simulations drawn from each trial site's historical climate record. The mean predicted yield response across sites, weighted by each site's probability of being selected for fertilizer elimination, is 1.50 t ha$^{-1}$. The predicted response at a given site in a given year ranges between 0.31 t ha$^{-1}$ and 3.11 t ha$^{-1}$. The average site-level standard deviation of fertilizer response is 0.64 t ha$^{-1}$, ranging from 0.32 t ha$^{-1}$ in the trial site with lowest variability to 1.15 t ha$^{-1}$ in the trial site with highest variability.

Fertilizer response does indeed vary with growing conditions, a result that bears important implications for the value proposition that farmers face in using fertilizer. We reject the null hypothesis that there is no heterogeneity in the fertilizer response across covariates, with a $P$ value of 0.001. Extended Data Fig. 1 shows the importance of each predictor in explaining the fertilizer treatment effect. The strongest predictor of fertilizer response is precipitation in the first two months of the growing season, followed by soil clay content, soil pH, soil bulk density and soil cation exchange capacity.

Extended Data Fig. 2 shows the predicted fertilizer response plotted over climate and soil characteristics. Expected fertilizer response is increasing over total growing season precipitation and soil clay content. It is decreasing in soil bulk density and soil cation exchange capacity. Fertilizer response follows an inverted U shape in growing season temperature in that it rises over increasing low temperatures and falls as growing season temperature increases beyond 23 °C. It also follows an inverted U shape over elevation and soil silt content. Fertilizer response is largely flat across total soil nitrogen similar to ref. [22]. Fertilizer response is slightly downward sloping as soil pH increases. This varies from the expected response from the literature[23,24], which suggests the highest fertilizer response in a pH range of 6.0–6.8 and then a lower response at lower and higher pH levels. Our data have a low density of observations at the extremes, perhaps contributing to why our response function differs from the literature. Fertilizer response does not follow a clear pattern over the majority of the data distribution for soil exchangeable acidity or soil organic matter.

**Forward-looking measures of fertilizer profitability.** Holding all else equal, a farmer is not expected to adopt a technology if the value of the increased output generated is less than the added cost of the technology[25]. The internal rate of return (IRR) is a single measure that incorporates the discounted private benefits and costs that a farmer faces when considering use of an input at the beginning of the growing season:

$$\mathrm{IRR}(f_q) = \frac{\Delta y \times p^{\mathrm{y}} - \Delta f \times p^{\mathrm{f}} \times (1 + r)}{\Delta f \times p^{\mathrm{f}} \times (1 + r)} \quad (1)$$

where the change in output for crop $y$ is depicted by $\Delta y$, the change in input use is depicted by $\Delta f$ and represents the difference between inorganic fertilizer applied in quantity $q$ ($f_q$) and no fertilizer use, $p^{\mathrm{y}}$ is the output price, $p^{\mathrm{f}}$ is the fertilizer price and $r$ refers to the cost of financing the input purchase. We use monthly maize prices from 99 markets in 20 countries and monthly urea prices from 102 markets in 17 countries (Methods, Extended Data Fig. 3 and Extended Data Table 2 provide more details about the construction of these variables and a map showing these market locations across Africa). We focus on urea prices because urea is generally the least expensive form in which farmers can acquire nitrogen.

The limitations of IRR as a metric to predict input use, profitability and input demand by farmers are well known to researchers, who often adjust the target profitability threshold to account for heterogeneity and uncertainty. In the model described by equation (1), uncertainty arises both from the crop response, depending on stochastic climate realization, and from market fluctuations of output prices because farmers generally do not know the market price they will receive for their outputs at the time inputs are purchased. Though farm profitability is affected by many more forms of uncertainty, including that arising from biotic factors such as pests and disease, from input and output market failures or from human

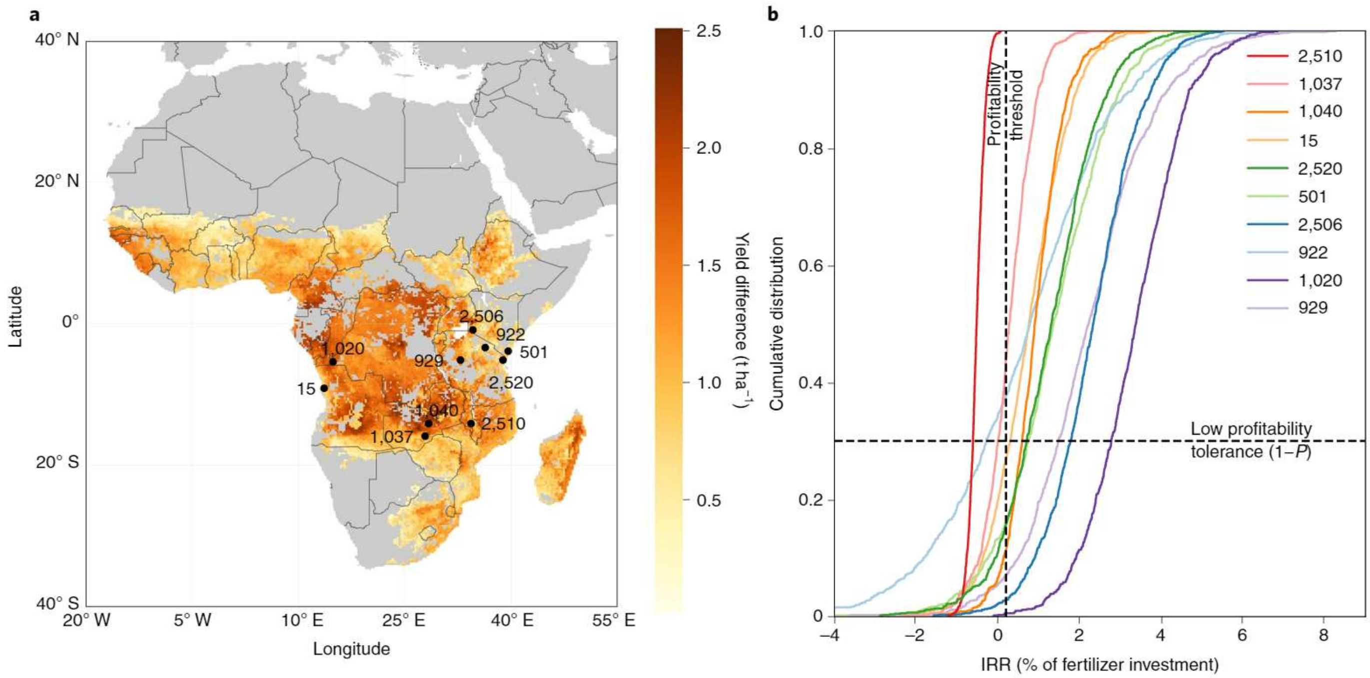



**Fig. 1 | Predicted fertilizer yield response at all African sites modelled and predicted profitability of fertilizer use at randomly selected trial sites.** **a**,**b**, Predicted fertilizer yield response expected to be exceeded at least $P=70\%$ of the time in all African sites modelled (**a**) and the full probabilistic distribution of profitability (IRR) in ten randomly selected trial sites (**b**). The colour shading in **a** indicates the simulated yield response to fertilizer (t ha$^{-1}$) at the probability threshold. The black dots indicate the locations of ten randomly selected trial sites whose IRR distributions are depicted in **b** over simulated temperature and precipitation conditions and simulated output prices. The vertical dashed line represents the profitability threshold ($T$, set to an IRR of 30% of the fertilizer investment) while the horizontal dashed line depicts the risk tolerance for a profitability outcome lower than the threshold. $P$ represents the share of years in which the farmer seeks to achieve the profitability target (70% of the years in this example), and $1-P$ represents the share of years in which a model farmer would tolerate profitability falling below the target. In this case, the model farmer seeks to clear the profitability threshold of >30% IRR at least 70% of the time.

conflict, our current tool focuses on climate and prices. Even though an input's IRR need only exceed 0 for its use to be profitable, a researcher may define a much higher target to be confident that profitability is likely. Agronomists have often used the value–cost ratio as a profitability measure, which is very similar to the IRR, reflecting the ratio between benefits and costs rather than the discounted net benefits as a share of added costs[26,27]. Rather than seeking a larger mean IRR to ensure the investment is robust, we examine the probability over a range of outcomes that the profitability exceeds a desired threshold. For example, given the same expected profitability, one would expect farmers to perceive a technology to be more robust if profitability falls below the threshold once every ten years than if it falls below the threshold three times per decade.

Figure 1b depicts the cumulative distribution of IRR at ten of the trial sites, selected at random, with locations plotted in Fig. 1a. Of these ten sites, seven qualify as robustly profitable in that the profitability forecast exceeds the profitability threshold $T$ (depicted by the dashed vertical line) in a share of the distribution that exceeds $P$ (depicted by the dashed horizontal line). In one site (2510), IRR falls below $T$ across all simulations. In two additional sites (922 and 1037), IRR exceeds $T$ for part of the distribution but less than the target of $P=70\%$. In these sites, the profitability target is achieved over only 63% and 55% of the distributions, respectively. In all seven of the remaining sites, profitability exceeds $T$ for 70% or more of the site distributions.

Generally, fertilizer is believed to be a risk-enhancing input because it can exacerbate the negative impacts of increased temperatures[28]. Our results suggest that the returns to fertilizer use exceed the chosen profitability threshold across much of the distribution of outcomes in most sites but certainly not the full distribution of outcomes in all sites. At the first percentile of outcomes (that is, the worst 1% of simulated outcomes for each trial site), fertilizer use is profitable in only 1 of the 104 trial sites for which we are able to model maize and urea prices. At the fifth percentile, fertilizer use is profitable in only 12% of the sites, and at the 25th percentile, it is profitable in 63% of the sites. At the median simulated outcome, fertilizer use is profitable in 75% of the trial sites. By the 90th percentile, it is profitable in 87% of sites, and by the 99th percentile, it is profitable in 93% of the sites.

**Fertilizer response across sub-Saharan Africa.** To examine the implications of our yield model for fertilizer profitability at additional locations, we predict fertilizer responses conditional on soil characteristics and climate conditions across sub-Saharan Africa. We mask out locations where maize is not grown and locations where climate and soil conditions fall outside the common support of the data used to estimate the yield function. Figure 1a shows the simulated yield response to fertilizer at the probability threshold, meaning the fertilizer response that is expected to be exceeded in at least $P=70\%$ of the simulated years. This 30th percentile predicted fertilizer response ranges from 0.054 t ha$^{-1}$ to 2.50 t ha$^{-1}$ across cells, with a mean of 1.15 t ha$^{-1}$. The within-location standard deviation of fertilizer response over the 1,000 simulated growing seasons is, on average, 0.63 t ha$^{-1}$ (ranging between 0.27 t ha$^{-1}$ and 1.34 t ha$^{-1}$).

Given the challenge of modelling site-specific prices for both maize and fertilizer, we next consider what ratio between fertilizer and maize prices would be required in order for fertilizer use to be profitable according to the robust criteria. We solve for the price ratio that ensures $\text{IRR}\geq T$ in at least $P$ share of years. This price

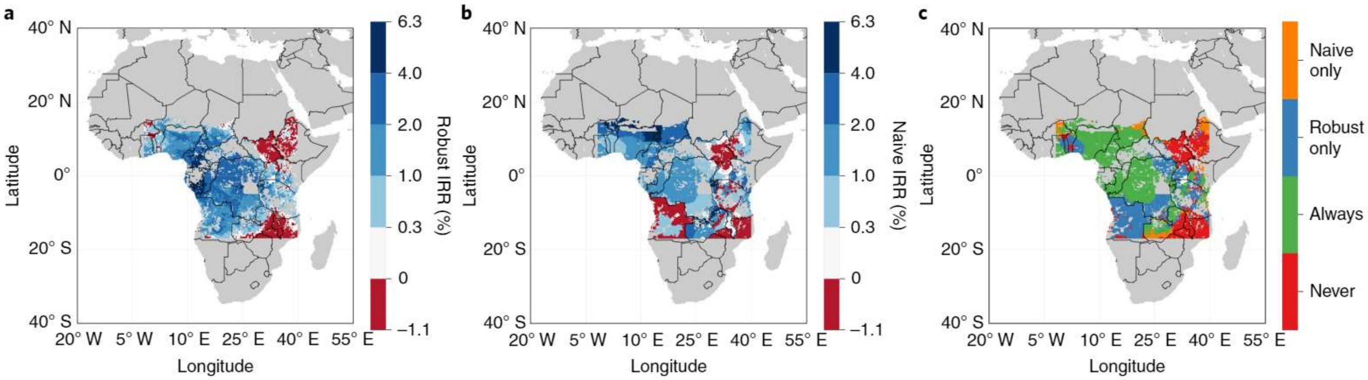



**Fig. 2 | Site-level IRR predictions and a confusion matrix classifying each site by altering profitability criteria. a,b,c**, 'Robust' site-level IRR predictions at the $1-P=0.3$ probability threshold (**a**), site-level IRR forecasts using the 'naive' criteria (**b**) and a confusion matrix classifying each site by its profitability according to both 'robust' and 'naive' criteria (**c**). **a**, Site-level IRR predictions expected to be exceeded at least 70% of the time based on the distribution of modelled yield response to fertilizer use over the synthetic climate dataset, with maize and fertilizer prices generated from local market data using spatial interpolation. **b**, Site-level IRR predictions that are not based on modelling fertilizer response but are derived using the average country-level side-by-side yield comparisons between fertilized and non-fertilized trial plots. Where country-level side-by-side yield comparisons are not available, we use the average side-by-side yield response for trial sites in the same AEZ. Interpolated maize and fertilizer prices are the same as those used in **a**. **c**, Sites classified by whether they are deemed profitable according to both the 'robust' criteria in **a** (IRR ≥30% in at least 70% of years) and the 'naive' criteria in **b** (naive IRR ≥100%).

ratio is a linear transformation of the fertilizer response depicted in Fig. 1a. Extended Data Fig. 4b shows where fertilizer use would be profitable depending on this price ratio. In portions of Angola and Zambia, fertilizer profitability is robust even if the fertilizer is quite expensive, exceeding nine times the price of maize. The fertilizer response is sufficiently high to support profitability, even if the fertilizer is relatively expensive (7–9 times more than the price of maize) in large parts of Angola, Cameroon, Democratic Republic of the Congo, Gabon, Guinea-Bissau, Malawi, Republic of the Congo, Sierra Leone and Zambia. Fertilizer response is the lowest and thus requires a very favourable price ratio of less than three times the price of maize to support profitability in parts of Botswana, Mauritania and Somalia. Extended Data Table 3 describes at the country level the average predicted yield response to fertilizer and the average price ratio that would be required for profitability at the probability threshold.

When we predict the profitability of fertilizer use across sub-Saharan Africa by modelling site-specific IRR distributions, we further restrict our analysis to locations where maize and fertilizer price data are available in sufficient geographic density to support price interpolation using Kriging (Methods). Figure 2a depicts each location's IRR measure at the point in the location's cumulative distribution corresponding to the probability threshold, $1-P=0.3$. Modelled IRR exceeds 200%, on average, in Cameroon, Central African Republic, Democratic Republic of the Congo, Gabon and Republic of the Congo. It exceeds 100%, on average, in Angola, Burundi, Chad, Niger, Nigeria and Rwanda. It is negative, on average, at the profitability threshold and thus mostly not profitable in Eritrea, Ethiopia, Malawi, Mozambique, Sudan and Zimbabwe. Many parts of western Africa, southern Africa and the Horn of Africa are excluded from this analysis due to insufficient coverage of either fertilizer or maize market prices.

**Decision making and fertilizer profitability.** We compare profitability when assessed via the 'robust' ex ante profitability measure developed in this manuscript with a counterfactual 'naive' definition. The 'naive' profitability measure calculates IRR using the average yield difference between fertilized and non-fertilized trial sites in a country, rather than modelling yield differences. If trial data are not available in a given country, we use the average yield difference for trial sites located in the same agro-ecological zone (AEZ), restricting yield measures to side-by-side comparisons between fertilized and non-fertilized plots. This measure is analogous to an ex post measure of profitability as commonly applied in the literature, which also uses a higher target threshold of profitability to compensate for the fact that ex post measures assume away risk. Figure 2b maps IRR according to this 'naive' profitability criteria (IRR ≥1) at the $1-P=0.3$ threshold. While predicted yield response is quite heterogeneous spatially, as depicted in Fig. 1a, the predicted IRR in Fig. 2a appears more spatially homogeneous. When moving from predicted agronomic fertilizer response to predicted profitability, we require input and output prices, which are strongly influenced by country-level trade policies and regulations.

Figure 2c compares the profitability assessment between the two sets of criteria. In the cells for which we are able to predict profitability using both the 'naive' and 'robust' criteria, the assessments are in agreement for 74.9% of the sites, with 60.9% classified as profitable under both sets of criteria and 14.0% classified as not profitable under both sets of criteria. Fertilizer response could be systematically lower in the screened-out locations where we withhold profitability predictions because maize is not grown, because the soil or climate conditions are different than those in our estimation sample or for lack of price data.

Those taking a 'naive' approach to assessing fertilizer profitability would consider fertilizer use to be profitable in 12.5% of sites, where our 'robust' criteria consider it not profitable. In these sites, those following the 'naive' criteria risk promoting fertilizer adoption where it might not be profitable given price incentives and the range of agronomic responses to fertilizer. These sites are concentrated in Burkina Faso, Eritrea, Ethiopia and Zambia. By contrast, under the 'naive' approach, the remaining 12.6% of sites are classified as not profitable, where our 'robust' criteria do consider fertilizer to be profitable. In these sites, those following the 'naive' criteria might not recommend fertilizer where our analysis suggests its use would be profitable. These sites are concentrated in Angola, Benin, Ghana, Democratic Republic of the Congo, South Sudan, Tanzania, Togo and Uganda. We also calculate a 'naive' IRR measure based on whether average fertilizer returns exceed the same profitability criteria used for the 'robust' definition (IRR ≥0.3). These characterizations of profitability, shown in Extended Data Fig. 5, are very similar

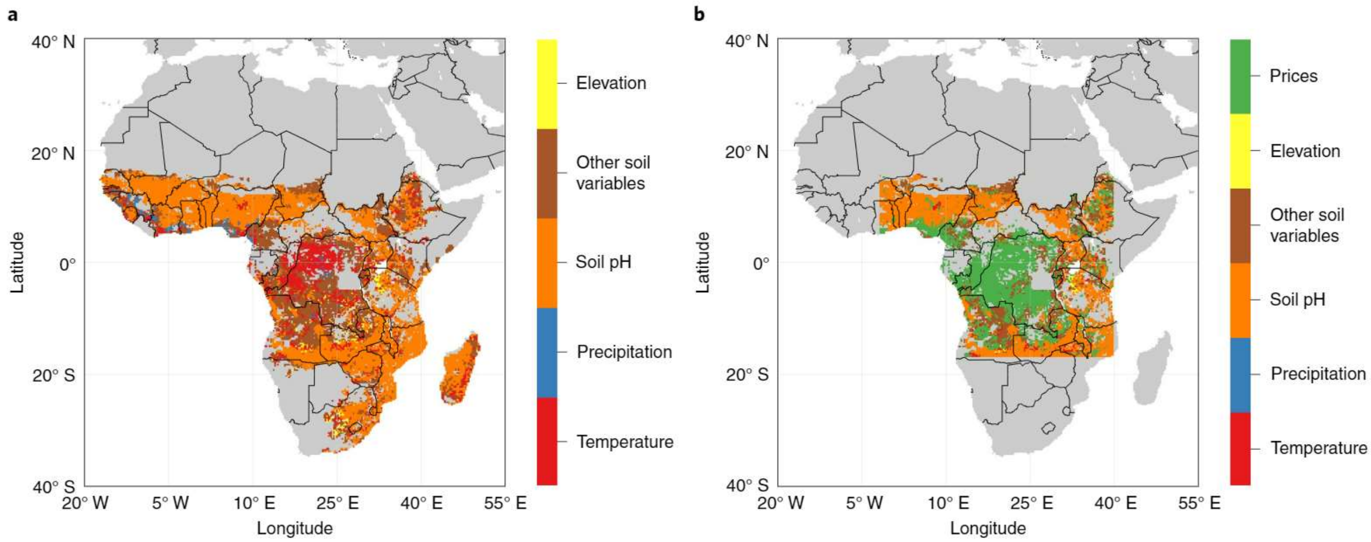



**Fig. 3 | Sensitivity analysis of predicted yield response to fertilizer use and predicted IRR. a,b**, The panels highlight which variable to which predicted yield response is most sensitive (**a**) and to which predicted IRR is most sensitive (**b**). Sensitivity to each input variable is defined using a local, one-at-a-time sensitivity analysis.

to those in Fig. 2c, though they disagree at least 25% of the time in Angola, Benin, Democratic Republic of the Congo, Ethiopia, Ghana, Malawi, Tanzania, Togo and Uganda.

**Determinants of investment profitability.** Figure 2 is useful for targeting fertilizer recommendations in regions that are robustly profitable but does not provide insights into how to improve profitability in regions where it is not. To understand what input variables enable some regions to be robustly profitable or prevent other regions from yielding profitable returns to fertilizer adoption, we estimate the sensitivity of both predicted fertilizer yield response and fertilizer profitability (IRR) to the predictors and assumptions that influence them. Figure 3 shows the variable within each pixel to which (1) the yield difference from fertilizer adoption and (2) IRR are most sensitive. Here sensitivity to each input variable is defined as the average absolute change in fertilizer response or IRR across 1,000 simulated years in response to a change in the input variable of +5% and −5%. In Extended Data Fig. 6, we also show the most sensitive variable within particular simulated years corresponding to the fifth percentile, median and 95th percentile yield difference and IRR.

Figure 3a shows that soil pH is the most common variable to which the fertilizer response is most sensitive throughout sub-Saharan Africa (51% of pixels). This is followed next by other soil characteristics (31%), mean temperature (14%), precipitation (2%) and elevation (2%). It is surprising that fertilizer response is most sensitive to precipitation in so few locations given precipitation in the first period of the growing season is the input variable with the greatest importance score in the causal forest model (Extended Data Fig. 1). This could be because spatial variability in precipitation explains spatial variability in fertilizer response but that a small change in precipitation at a particular site explains little variability in fertilizer response locally. Precipitation may also interact substantially with other variables, yielding a high importance score in the causal forest model but a low first-order influence in a one-at-a-time sensitivity analysis.

While the greater local importance of soil characteristics in determining fertilizer response (compared with precipitation) is surprising, it is promising from a management perspective as many soil characteristics can be altered through management (unlike precipitation). Soil management interventions might be most useful in Ethiopia, South Sudan, Malawi and Mozambique, where fertilizer adoption is currently never profitable (Fig. 2) and both fertilizer response and IRR are most sensitive to soil pH. In locations where fertilizer is predicted to be profitable according our 'robust' definition (Democratic Republic of the Congo, Central African Republic, Côte d'Ivoire and Republic of the Congo), the most important predictors explaining fertilizer response are primarily temperature and other soil characteristics. Locations where the fertilizer response is most sensitive to temperature should be monitored as the climate continues to warm. In these locations, under the current climate regime, the impact of temperature on IRR is dominated by prices, specifically urea prices. Because fertilizer response in temperature-sensitive locations is consistently high, profitability responds less to changes in fertilizer response than to the price ratio between maize and urea. This transfer of sensitivity highlights the importance of considering both biophysical and economic influences on profitability. While soil pH is still the variable to which profitability is most commonly most sensitive (40% of pixels), urea prices are a close second (38%), while other soil variables (19%), temperature (2%), precipitation (0%) and elevation (1%) are less influential. Some of the pixels where urea prices are the most dominant factor extend into the never-profitable regions, suggesting interventions to improve the affordability of fertilizer could be promising, perhaps in conjunction with soil amendments.

## Discussion

We propose a new approach to assessing the returns to soil health investments in the face of climate uncertainty and spatial heterogeneity. We find that the response to fertilizer use indeed differs under different soil characteristics, growing season climate realizations and combinations thereof. These results highlight the importance of considering variability in response over time and space when targeting fertilizer recommendations rather than prescribing blanket recommendations based on the mean fertilizer yield response. While the predicted fertilizer response in an agronomic trial setting may not perfectly correspond to the fertilizer response that a farmer might observe on the fields, it is nevertheless informative to explicitly examine fertilizer response interactions with climate realizations and site characteristics. Knowing the range of possible outcomes in a site at the start of a growing season can be helpful, even in the absence of seasonal climate forecasts and field-specific soil information, as it can help decision makers target fertilizer-promotion efforts geographically to locations where the fertilizer response is likely to be large. By failing to incorporate systematic sources of heterogeneity and uncertainty into fertilizer profitability assessments, one risks two types of error—promoting fertilizer in places where its use is not actually profitable and failing to promote it in places where its use could be profitable. Our analysis suggests that decision makers would mischaracterize the returns to fertilizer use in about 25% of the locations for which we can compare assessment approaches if they account for risk by raising the profitability target as opposed to explicitly incorporating heterogeneity and uncertainty.

One key limitation of our approach arises from the imperfect measurement of soil characteristics. We use static soil maps generated at a 250-metre resolution, although soil characteristics are known to vary at much finer scales according to physical, biological and management history[29]. Measurement error of soil characteristics affects the trial dataset used to estimate the fertilizer response and the sub-Saharan Africa-wide dataset that we use to predict fertilizer response. If soils on the plots where the agronomic trials are conducted are systematically better than the map-based predictions for those sites and if soils on smallholders' fields are systematically worse than their map-based predictions, then our estimation of yield response and our out-of-sample predictions could be biased[30]. Further exploration of soil measurement error and further validation of fertilizer yield response predictions on plots with local soil data would help to shed light on possible biases arising from soil data mismeasurement.

A second limitation to our approach arises if there are discrepancies between how crops are managed in agronomic station trials and on smallholders' fields, which could limit the validity of our fertilizer response predictions for smallholder-managed fields. Due to differences in knowledge, price incentives or constraints faced, trial managers may achieve a larger fertilizer response than farmers due to better timing of planting and other activities, weed management, soil water management and other aspects of crop production[31,32]. Alternatively if soils on farmers' plots are more depleted than those on experiment stations, fertilizer response on experiment stations could be smaller than those on farmer-managed plots[33,34]. Our yield predictions are based on agronomic station trials and should not be taken as predictions of smallholders' yields but rather how yields on similarly managed plots would differ with changes in growing season temperature, precipitation and soil characteristics. Finally, we emphasize the shortage of low-nitrogen trial data, which limits researchers' ability to understand heterogeneity in fertilizer response and to evaluate the profitability of fertilizer use.

This new fertilizer profitability-assessment tool can be strengthened as more data become available. Additional fertilizer response trials in cool and dry regions would be especially useful to extend the range of conditions under which fertilizer profitability can be predicted. It would also be interesting to trace out a continuous fertilizer response, as the fertilizer response per kilogram applied is likely to be higher when fertilizer is applied at quantities lower than the full recommended rate. Understanding the continuous fertilizer response will require a major effort to collect fertilizer dosing trial data over multiple sites and years. Medium-range climate forecasts such as the El Niño-Southern Oscillation signal, which is available at the time of planting, could also be used to refine fertilizer profitability recommendations in El Niño and La Niña years when climatic patterns tend to differ[35]. Our tool draws its synthetic climate dataset from the historic climate distributions with extrapolated mean temperature trends to 2019, but we could use projected future distributions that incorporate climate change trends across the whole distribution of both variables to predict fertilizer profitability under future climate scenarios, which will vary considerably in many African countries[28,36].

This new tool has several important applications for research and policy. First, it can be used to help understand technology adoption behaviour by smallholder farmers. Low adoption of fertilizer and other productivity-enhancing inputs by smallholders has motivated a large body of research[37] with explanations that include financial constraints[38], quality uncertainty[39], limited information about the technology[40] and heterogeneous returns[41,42]. Economic theory suggests that farmers will optimally use less fertilizer when facing an output distribution that is more variable[43], though risk and profitability explanations for low adoption are limited beyond a few studies (for example, refs. [44,45]). With better predictions of fertilizer profitability conditioned on soil characteristics and growing season conditions, it is possible to assess whether these fundamental determinants of profitability could explain lack of fertilizer adoption where other explanations have fallen short.

Second, by conditioning profitability estimates on site-specific determinants while also considering the distribution of the response at each site, we develop a fertilizer assessment framework that can be used to decompose the relative contributions of site characteristics, climate risk, input–output price ratios and output price risk in determining fertilizer adoption by smallholder farmers in rainfed settings. Understanding the contributions of different constraints is important for identifying the interventions that are likely to be most successful in increasing fertilizer profitability and, consequently, adoption. For example, if fertilizer profitability is determined by climatic risk, then offering farmers weather-indexed insurance products that cover input purchases could help to increase adoption levels. If instead fixed soil characteristics constrain profitability, then technologies that are adapted to different soil settings are likely to be effective for increasing adoption. Infrastructure investments or fertilizer transport subsidies could promote adoption of fertilizer in places where the input–output price ratio constrains profitability, while contract arrangements that provide output price certainty at the time of planting could promote adoption where output price risk is a constraint.

Third, decision makers can better target their fertilizer-promotion efforts by better understanding fertilizer profitability and its determinants. Because the determinants of input use, profitability, that is, agroclimatic conditions, and market access are often correlated with socio-economic variables such as poverty (for example, refs. [18,45]), improved intervention targeting has both efficiency and distributional implications. While site-specific fertilizer use recommendations alone are unlikely to drive fertilizer adoption by smallholders[46], they do have a role in helping extension agents and non-governmental organizations learn where to most productively target fertilizer-promotion efforts and identify location-specific determinants of fertilizer returns and profitability. Targeting investments in agricultural production (for example, credit to purchase fertilizer) should not be done in complete isolation from understanding output risk and the household's ability to smooth consumption in bad years[45]. If we assume that smallholders in developing countries are limited in their ability to smooth consumption in the face of risk due to financial market underdevelopment, it is important that planners who promote fertilizer use consider the full distribution of possible outcomes that farmers face at the time of planting. Finally, appropriately targeting inorganic fertilizer in high-potential areas (versus low-potential areas) can reduce environmental impacts, as high-potential fields will retain more nitrogen and contribute less to pollution of downstream water bodies[47].

## Methods

We combine data from a variety of sources to create a dataset to execute our analysis. Extended Data Table 2 outlines the datasets used, their provenance, the variables used from each dataset and for which years. Each of these datasets is described in more detail in the following text.

**Agronomic trial data.** We estimate the crop response to fertilizer by combining several datasets of multiple maize trials across eastern and southern sub-Saharan

Africa. We combine the data compiled by ref. [15], the sites used by refs. [48,49] and the sub-Saharan African sites included in the International Maize and Wheat Research Institute's (CIMMYT) publicly accessible International Maize Trial Network data to generate our dataset[50]. The dataset includes trials managed by CIMMYT, national agricultural research institutes, the trials conducted under the Optimization of Fertilizer Recommendations in Africa (OFRA) project[49] and private seed companies. The trials span 13 years (1999–2007 and 2013–2016), 17 countries and eight different AEZs. We focus on maize, which is the most widely grown crop in sub-Saharan Africa, accounting for 27% of all cereal area, 34% of all cereal production and 31% of all calories from cereals in the region[51]. The locations of the maize trials are plotted in Extended Data Fig. 3a.

The different crop trials were conducted for varying purposes. The CIMMYT trials were initiated to test the performance of new maize varieties in different conditions[52], the OFRA project aimed to trace out fertilizer yield response curves based on various fertilizer dosing regimes[48] and the dataset amalgamating different agronomic trials was compiled to study maize response to water and temperature stress[15]. While estimating a fertilizer yield response was not a primary focus of two out of three of these data sources, many of the varieties were tested under a low-N management regime in which crops were planted on fields that were depleted of nitrogen due to continuous cropping of maize over previous seasons, removing all stover after previous harvests and withholding all application of both inorganic and organic fertilizers. Apart from N application, all other crop-management practices were held constant between low-N and optimal N trials. In the estimation dataset, trial observations that fell under different management regimes than low-N and optimal N (for example, the low pH, drought management or streak virus management regimes) were not included in our analysis. The low-N treatment almost never interacted with these other management regimes so including them would not help us to understand the heterogeneity of fertilizer response. After dropping 5,629 observations that are under low pH, drought or maize streak management, our remaining estimation sample contains 21,418 observations.

The application rate of nitrogen in the optimal N management regime was 125 kg N ha$^{-1}$ for the CIMMYT maize trials[52] and 120 kg N ha$^{-1}$ for the OFRA project[48]. In all cases, the low-N regime was 0 kg N ha$^{-1}$. While phosphorus is a very important yield determinant for African maize[53], phosphorus application rates were held constant across low-N and optimal N treatments in the data we use. The two data sources based on CIMMYT maize trials[15,52] used 18 kg P ha$^{-1}$ for both optimal and low-N treatments (refs. [54,55], for example), while the OFRA project used 15 kg P ha$^{-1}$ in both its optimal and low-N treatments[48]). None of the trial treatments included micronutrient additives (for example, Zn, B, Fe, Mo, Cu). While we acknowledge the importance of micronutrients in conditioning crop response to fertilizer[56], we cannot measure or control for their effects using our trial dataset.

The first two columns of Extended Data Table 1 show descriptive statistics from the trial estimation dataset separated by whether the observation received the low-N or the optimal N treatment. Yields in the low-N sites are 2.02 t ha$^{-1}$, less than half of yields in optimally managed site that average 4.32 t ha$^{-1}$. The third column shows the normalized difference between fertilizer treatments for each variable, which is the difference in average between the two groups scaled by the square root of the sum of variances across groups. The low-N trial sites differ from the optimal N sites in climate and soil conditions. While the trial dataset covers 142 different trial sites, the low-N treatment takes place in only 51 of those sites. On average, the sites that included low-N observations tended to be slightly warmer and slightly drier than the sites where the low-N treatment was not applied. We address this imbalance by introducing site-level propensity weights into the fertilizer response estimation as discussed in the fertilizer response estimation section of Methods.

**Soils data.** Soil organic matter can influence soil structure, moisture retention and nutrient retention in soil, which is important because applied nitrogen leaches readily through the soil profile, becoming unavailable to crops[57,58]. Yield response to fertilizer has been shown to vary with soil organic matter in Western Kenya[41]. Soil pH also influences nutrient retention and availability to plants, with fertilizer–soil organic matter and fertilizer–mineral interactions typically weakened as soils become more acidic. Soil micronutrients generally become more soluble in acidic soils, which can increase their availability to crops[59]. In researcher-managed fertilizer trials conducted in east Africa, fertilizer recovery was found to be higher in soils with a deep, clayey profile and lower in sandy soils[60].

We match the trial sites with relevant soil characteristics using the Africa Soil Information Service[29]. The 250-metre resolution soil data include estimates of soil characteristics at different soil depths. We use characteristics such as soil cation exchange capacity, soil pH, soil exchangeable acidity, soil organic matter, soil nitrogen content, soil texture (silt and clay content), soil bulk density and soil drainage. Finally, we match the trial data with the 16-zone AEZ classifications developed by the Food and Agriculture Organization (FAO) and the International Institute for Applied Systems Analysis[61].

For our fertilizer profitability simulations throughout sub-Saharan Africa, we upscale the 250-metre soil data to a 0.25-degree grid. We upscale by taking the average for all continuous soil variables and the majority for the categorical soil variable (drainage). The code for processing the soil, climate and price data and for estimating the yield response and simulating fertilizer profitability are all publicly available as described in our data and code availability statements.

**Climate data.** Spatial heterogeneity in precipitation is important to consider when assessing the returns to fertilizer at a given location. The majority of agriculture in sub-Saharan Africa is rainfed, and empirical evidence suggests that rainfall is the common yield-limiting factor among all major cereals[62]. Fertilizer response has been shown to decrease with increasing water stress during the growing season[63]. To the extent that crop responses to soil health interventions are determined by rainfall levels, rainfall conditions in a single year will be a strong determinant of the profitability in that year of soil health interventions such as fertilizer application.

Temperatures also vary in space and are important determinants of crop growth. Negative relationships between temperature (growing degree days) and African maize yields have been well documented[15,64]. Nitrogen application can help mediate the effects of heat stress[15]. In a study using side-by-side comparisons of fertilizer-treated and non-treated on-farm experimental plots across a large sample of Malawi model farms over multiple growing seasons, fertilizer response was shown to vary with temperature and rainfall[18]. When agronomic trials were repeated over multiple years, fertilizer response was shown to vary with both rainfall and temperature[65].

Using the location of each experiment site, we match each crop trial observation with climate data. We use daily mean temperature and total precipitation data computed from hourly observations at 0.5-degree resolution from the fifth generation of the European Centre for Medium-Range Weather Forecasts over the period 1979–2018[66]. In the crop-growing season, which we define as five 30-day periods, or months, after planting[52], we calculate total precipitation and average temperature in each month for each site. The third month generally coincides with flowering and silking, a period that is considered especially sensitive to water and temperature stress. Because our model does not assume that the fertilizer response is linear in parameters across the full parameter space, we are able to capture the negative effects of high temperatures on fertilizer response using mean monthly temperature as a predictor rather than separately modelling growing degree days or extreme heat degree days[67]. The predictive power of an alternative specification (with growing degree days and heat degree days instead of average temperature) is very similar.

For 0.25-degree grid cells in sub-Saharan Africa, we generate 1,000 year-long synthetic climate datasets using each cell's historical climate distribution from the European Centre for Medium-Range Weather Forecasts climate forcing dataset, with the mean temperature for 2019 estimated from a linear trend since 1979. The climate dataset is at 0.5-degree resolution, so clusters of four 0.25-degree cells within each 0.5-degree pixel have the same climate statistics. However, the 0.25-degree resolution allows us to better capture spatial heterogeneity in soil characteristics. Using the planting date for each cell, we compute monthly growing season temperature and precipitation values over five consecutive months after the planting date described above. A linear model was then fit to each month's mean temperature time series. Next, a Gaussian copula was fit to each month's precipitation total and mean temperature anomalies from the linear trend. Precipitation totals were modelled by Gamma marginals, while temperature anomalies were modelled by normal marginals. A first-order vector auto-regression model was fit to the observed standardized anomalies of the Gaussian copula over the five growing season months. One thousand years of synthetic standardized anomalies over the five growing season months were then generated from the first-order vector auto-regression model. These were then transformed back to real space through their marginals, and each month's mean temperature estimate for 2019 from the linear trend was added to the corresponding synthetic temperature anomaly.

**Price data.** We downloaded monthly maize and monthly urea price series for all available African countries. Maize price data were sourced from the FAO Global Information and Early Warning System (GIEWS) dataset[68] and urea retail prices from the africafertilizer.org initiative[69]. Maize data included 99 markets in 20 countries with price data available as far back as 2000 in some markets. Urea data included 102 markets in 17 countries with price data available as far back as 2010 in some markets. The locations of the markets with price data are plotted in Extended Data Fig. 3b. We used monthly world maize prices (the US Gulf price) from the FAO GIEWS dataset and the monthly world urea price (the Black Sea, bulk, spot, free on board price) as reference world market prices[68,70]. The local maize price data were downloaded in US$ kg$^{-1}$ (the conversion from local currency to US$ was done by FAO GIEWS). We converted local retail urea prices to US$ kg$^{-1}$ using monthly local currency exchange rates from OANDA. For maize, we ran a regression of the form:

$$\ln\left(\frac{p_{it}^{\mathrm{m}}}{\mathrm{pw}_t^{\mathrm{m}}}\right) = \alpha_i^{\mathrm{m}} d_i + \beta_i \mathrm{mkt}_i + \chi \mathrm{yr} + \delta \mathrm{yr}^2 + \phi_{ct}(\mathrm{month} * \mathrm{country}),$$

and for urea we ran a regression of the form:

$$\ln\left(\frac{p_{it}^{\mathrm{u}}}{\mathrm{pw}_t^{\mathrm{u}}}\right) = \alpha_i^{\mathrm{u}} d_i + \chi \mathrm{yr} + \delta \mathrm{yr}^2 + \phi_{ct}(\mathrm{month} * \mathrm{country}),$$

where $p_{it}^{m}$ ($p_{it}^{\mathrm{u}}$) is the local market price for maize (urea) in market location $i$ at monthly time $t$, $\mathrm{pw}_t^{\mathrm{m}}$ ($\mathrm{pw}_t^{\mathrm{u}}$) is the world market price for maize (urea) at monthly time $t$, $d_i$ is a categorical variable for market location $i$ and $\mathrm{mkt}_i$ is an indicator variable for local market type (retail or wholesale, only for the maize data). The variable yr is the number of years since the initial year of the dataset, month is the month of the year (1–12) and country is a categorical variable for the country. The interaction between month and country allows us to control for a different price seasonality cycle in each country for each commodity. The parameters $\alpha_i^m$ ($\alpha_i^u$, $\beta$, $\chi$, $\delta$ and $\psi$) represent the marginal effects of maize (urea) market location, maize market type, year, year squared, and month-country interaction on the local maize (and urea) market price.

Using these regressions, we predict the percent increase over the world price for each market location (and commodity). We multiply this predicted price increase by the world price to predict prices in wholesale maize markets across Africa for the approximate harvest month in 2019. To assist in identifying harvest month, we examine the monthly cycle of maize prices in each country to find the month with the lowest relative price (which indicates the post-harvest flux of commodity onto the market). Fertilizer markets also have monthly price cycles reflecting typically higher demand (prices) in the planting season and lower demand (prices) during the harvest season. We examine the monthly cycle of urea prices in each country to find the month with the highest relative price, which is our indication of the planting season peak price. Using this information, we predict each market location's maize and urea price coefficient in 2019.

The maize regression gives us 99 different predicted $\alpha^m$ coefficients, and the urea regression gives us 102 different predicted $\alpha^u$ coefficients, one for each commodity market location. To predict $\alpha^m$ and $\alpha^u$ for all gridded 0.25-degree cells on the map, we interpolate $\alpha^m$ and $\alpha^u$ using ordinary Kriging with a spherical variogram for the maize prices and a Gaussian variogram for the urea prices. We do the same for their standard errors. Ordinary Kriging uses spatial gradients and works for out-of-sample market price predictions if we have a good spatial distribution of observations across the relevant gradients of conditions that affect price in our data. In our sample we have, on average, five (six) market locations per country for maize (urea) markets, so we know that our data is not focused only on one or two major cities in each country. Furthermore, the underlying maize price data were generated as part of a famine early-warning system, and the market locations represent not only large, easy-to-reach markets but also many rural markets that are more difficult to access.

Our market-level predicted maize and fertilizer price can be thought of as local wholesale prices as opposed to a 'farm gate' price. Farmers probably receive prices less favourable than what we predict because they face farm-level (idiosyncratic) transaction costs in procuring inputs and selling outputs which can be quite large[71]. To correct for farm-level transaction costs in acquiring inputs and marketing outputs, we assume that the local wholesale fertilizer price is 75% of the farm gate fertilizer price (following ref. [72]) and that the farm gate maize price is 25% lower than the local wholesale maize price (following ref. [73]). Similar to a recent study on the spatial variation in fertilizer prices across sub-Saharan Africa[12], we did not find fertilizer price data for countries in southern African (for example, South Africa, Zimbabwe) and so our fertilizer price layer lacks coverage in southern Africa, limiting our ability to predict fertilizer prices (and thus simulate IRR) in the southernmost parts of sub-Saharan Africa. Finally, we assume that farmers face an annual interest rate of 31% to finance fertilizer purchases. This is the average annual (real) interest rate for microfinance institutions in African countries over the past decade (2010–2019) according to the Microfinance Information Exchange dataset from the World Bank, which we convert to a 14% seasonal rate assuming a six month seasonal loan period[74].

**Fertilizer response model.** To predict a yield response to fertilizer use that is heterogeneous over observable covariates (that is, soil characteristics and climate conditions), we estimate a causal forest model using the grf (Generalized Random Forest) package[75] in R[76]. We use yield rather than log yield as the dependent variable because the treatment effect we estimate (fertilizer response) involves applying fertilizer in a fixed quantity regardless of baseline yields. A priori, we have no reason to believe that adding 120 kg ha$^{-1}$ of fertilizer would result in the same multiplicative effect on yields in a location with very low baseline yields as it would in a location with very high baseline yields. A level specification seems appropriate because fertilizer response is often discussed in terms of nitrogen use efficiency (or the rate of conversion of applied fertilizer to harvested grain)[77]. When we use the very similar regression forest to predict yields with and without fertilizer, we find that the root mean square error (RMSE) as a share of the outcome variable is comparable and thus the models perform similarly. When the outcome variable is level yields, the RMSE on model predictions is 73.9% of the mean (level) yield. When the outcome variable is log yields, the RMSE on model predictions is 80.4% of the mean (log) yield. Therefore we do not believe that our model performs worse when we use level yields versus log yields.

The causal forest algorithm generates a non-parametric prediction of the local average treatment effect based on a local maximum likelihood estimator trained with a weighted set of nearby observations, with the weights derived by the causal forest algorithm. The predictions are consistent and asymptotically Gaussian[21]. Our estimation approach accommodates the clustered design of the trial datasets in that we generate out-of-bag predictions for each trial observation without using any observations from the same trial site and year to generate those fertilizer response predictions. This helps us to ensure that prediction errors are appropriately assessed given the clustered nature of the trial data.

To account for differences between sites where the fertilizer elimination treatment was implemented and where it was not, we orthoganolize our causal forest using propensity weights that describe the predicted probability that each trial's site included some no-fertilizer observations and on each observation's predicted yield conditional on the covariates. We estimate propensity weights using trial site-level mean temperature, precipitation and soil characteristics as predictors. Summary statistics of the trial data, re-weighted by the site-level predicted probability of conducting no-fertilizer trials, are depicted in columns 4–6 of Extended Data Table 1. When we incorporate these site-level weights into summary statistics, the normalized difference between no-fertilizer and optimal fertilizer sites for all variables is within 25%. When this difference is greater than 25%, treatment effect estimates are sensitive to model specification[78]. Even though the differences remain statistically significant for almost all climate and soil variables, they are not especially large in magnitude.

In our causal forest algorithm, we orthoganolize our data using both these site-level propensity weights and using yield predictions generated from a random forest model that includes the covariates (excluding fertilizer) as arguments. Because nitrogen treatment was experimentally assigned, rather than selected by farmers based on expected response, prior knowledge or a financing constraint, a clean identification of the effect of nitrogen on crop growth is assured.

Because the causal forest model predicts fertilizer response at the observation level, where it is not possible to observe the true fertilizer response, we cannot compare the causal forest model predictions to those generated with a different model by comparing RMSE on a held-out testing sample. To facilitate such a comparison, we generate yield predictions using the very similar random forest algorithm, which we estimate using the same covariates and weights and the same software[75]. Our comparison model is a linear regression model that includes a subset of covariates selected from the full set of heterogeneity interactions (interacting fertilizer treatment with covariates) and additional interactions (interacting the covariate vector with itself). From the full covariate set, we select the subset to include based on Akaike information criterion using a variable selection algorithm[79]. We estimate this linear parametric model using fivefold cross-validation, each time leaving out 20% of the observations grouped into site-year clusters. The random forest model performs much better (smaller RMSE) in predicting out-of-cluster yields than the parametric linear regression model. The average RMSE on out-of-sample yield prediction is 5.20 t ha$^{-1}$ with the linear model, compared with 2.26 t ha$^{-1}$ for the random forest model. Extended Data Table 4 offers a side-by-side comparison of yield response predictions using the three estimation approaches described here (the causal forest model, the random forest model and a parametric production function model). The $R^2$ of a simple model regressing actual on predicted yields for out-of-cluster predictions is 0.422 (compared with 0.772 with the random forest model).

**Forecasting profitability.** Point estimates of profitability necessarily imply that crop response and output prices are certain[26,80,81]. However, assuming a certain crop response and known output prices is only appropriate when determining ex post whether using an input was profitable. Outcome uncertainty can explain low adoption rates even when it appears, on average or in a given year, that a technology is profitable. Moving into an ex ante framework, IRR takes the form of a random variable with probability distribution $f$(IRR) and reflects the distributional aspects of returns to fertilizer at a given location. Input use decisions, both regarding fertilizer use ($f_q$) and other inputs ($\bar{\theta}$), are determined exogenously in this framework, meaning that the fertilizer response function applies to a representative farmer who is hypothetically assigned to either use fertilizer or not use it, and thus the fertilizer use decision is not modelled. At each location $i$, the distribution of $\mathrm{IRR}_i$ is a function of output prices ($p_{it}^Y$) and fertilizer prices ($p_{it}^f$) in addition to the predicted fertilizer yield response based on the causal forest model. For the sake of simulating profitability, we assume that output prices are orthogonal to climate variables. In reality, African markets are characterized by very high trade costs, with local production levels affecting local market prices[82,83]. By assuming independence between maize price and climate, we might overestimate the returns to fertilizer use in good years and underestimate the returns to fertilizer use in bad years. Examining the dynamics of local price fluctuations interacted with trade frictions would be a valuable area of future enquiry.

We compute a site- and year-specific internal rate of return for a representative farmer in location $i$ at time $t$ applying the recommended quantity of fertilizer (that is, $q_f = 100\,\mathrm{kg\,ha^{-1}}$) at price $p_{it}^f$ to crop $y$, which incorporates the modelled yield response to fertilizer: $y\left(X_i, \omega_{it}; f_q, \bar{\theta}\right) - y\left(X_i, \omega_{it}; f_0, \bar{\theta}\right)$. For a given target profitability threshold that the farmer might seek ($T$), we evaluate robustness according to the probability that IRR is expected to exceed $T$. For example, if a farmer in location $i$ hopes to achieve a 30% return on fertilizer investment, we compute the probability that $\mathrm{IRR}_i > 0.3$. This can be derived by evaluating the probability (Pr) that $\mathrm{IRR}_i \leq T$ using the simulated cumulative distribution of IRR in location $i$:

$$\Pr\left(\mathrm{IRR}_i\left(X_i,\omega_{it},p_{it}{}^{y};\bar{\theta},f_q,p_i{}^{f}\right)\leq T\right)$$
$$=\iint_{\{\omega_{it},p_{it}{}^{y};\mathrm{IRR}_i\leq T\}}\mathrm{IRR}_i\left(X_i,\omega_{it},p_{it}{}^{y};\bar{\theta},f_q,p_i{}^{f}\right)\delta\omega_{it}\delta p_{it}{}^{y} \quad (2)$$
$$\approx\frac{1}{N}\sum_{t=1}^{N}\mathbb{1}\left[\mathrm{IRR}_i\left(X_i,\omega_{it},p_{it}{}^{y};\bar{\theta},f_q,p_i{}^{f}\right)\leq T\right]$$

where $f(\mathrm{IRR}_i)$ is the probability density function (PDF) of IRRs at location $i$ across years with different weather $\omega_{it}$ and crop prices $p_{it}{}^{y}$, and $N$ represents the number of years (indexed by $t$) simulated per location.

We approximate $f(\mathrm{IRR}_i)$ using a Monte Carlo simulation of $N=1{,}000$ years and count the number of years in which the simulated IRR exceeds the threshold $T$. Using a 0.25-degree gridded map of sub-Saharan Africa, we limit our simulations to locations that meet the following criteria. First, we mask out areas where maize is not grown using Harvest Choice's spatially disaggregated crop statistics data[84]. Second, we mask out areas where the temperature and precipitation distributions fall outside of the range used to estimate the fertilizer response model, excluding all cells where probability of fertilizer elimination is lower than the minimum predicted probability that a fertilizer exclusion trial was conducted in that cell.

For each cell that remains in our simulation sample, we use the causal forest parameters to predict fertilizer response over 1,000 growing season climate iterations sampled with replacement from the synthetic climate dataset described above. We sample a random error for each predicted fertilizer response based on the standard error of the out-of-sample causal forest prediction. To simulate profitability, we also sample a random error for each cell's maize price prediction using the Kriged standard error of the market-level maize price predictions.

Our ex ante profitability analysis focuses on the distribution of the stochastic IRR variable. For the purposes of this analysis, we assess the binary outcome $\mathbb{1}[1-F_{\mathrm{IRR}}(0.3)>0.7]$, that is, whether the farmer achieves at least a 30% return on the fertilizer investment ($T=0.3$) in at least 70% of the seasons ($P=0.7$). Our thresholds are illustrative—any decision support tool user could select different profitability and probability thresholds.

**Sensitivity analysis of yield response and profitability.** To determine which variables most control the yield response to fertilizer and profitability in different locations throughout sub-Saharan Africa, we perform a local, one-at-a-time sensitivity analysis. This allows us to estimate the first-order influence of each variable, but does not consider the influence of variable interactions.

For each synthetically generated season of weather and prices, we perturb each of the continuous variables used to predict yield response and profitability by +5% and −5%, while holding all other variables constant. Following the terminology of ref. [85], we denote the change in yield difference ($\Delta y$) or the IRR between these two perturbations as the elementary effect (EE) of the $j$th input variable $X_j$:

$$\mathrm{EE}_{\Delta y}(X_j)=\Delta y(1.05X_j)-\Delta y(0.95X_j) \quad (3)$$

$$\mathrm{EE}_{\mathrm{IRR}}(X_j)=\mathrm{IRR}(1.05X_j)-\mathrm{IRR}(0.95X_j). \quad (4)$$

To see which variable is most influential across the 1,000 synthetically generated seasons, we compute the average absolute value of the EE for $\Delta_y$ and *IRR* and variable $X_j$ as $\mu^*_{\Delta_y}(X_j)$ and $\mu^*_{IRR}(X_j)$, respectively.

$$\mu_{\Delta y}{}^{*}(X_j)=\frac{1}{1,000}\sum_{i=1}^{1,000}|\mathrm{EE}_{\Delta y}(X_{j,i})| \quad (5)$$

$$\mu_{\mathrm{IRR}}{}^{*}(X_j)=\frac{1}{1,000}\sum_{i=1}^{1,000}|\mathrm{EE}_{\mathrm{IRR}}(X_{j,i})|. \quad (6)$$

Figure 3 shows the variable with the greatest value of $\mu_{\Delta y}{}^{*}$ (Fig. 3a) and $\mu_{\mathrm{IRR}}{}^{*}$ (Fig. 3b) in each pixel. We also find the variable with the greatest |EE| in the synthetic years yielding each pixel's 5th, 50th and 95th percentile yield difference and IRR. This is shown in Extended Data Fig. 6.

## Data availability

Data used in this study are openly available, and we have provided citations that include a either a DOI or URL at the point in the methods section where each dataset is described. Our processed estimation dataset is available on Github (https://github.com/julianneq/Africa_Fertilizer_Profitability).

## Code availability

All of the code that we have written, including code to generate our estimation dataset, estimate the fertilizer response function, simulate fertilizer response in trial sites and across sub-Saharan Africa, analyse sensitivity of profitability to the diiferent predictors and generate the manuscript tables and figures, are available on Github (https://github.com/julianneq/Africa_Fertilizer_Profitability).

Received: 7 April 2021; Accepted: 18 March 2022;
Published online: 25 April 2022

## References


1. *Rural Development Report 2016: Fostering Inclusive Rural Transformation* (IFAD, 2016).
2. Evenson, R. E. & Gollin, D. Assessing the impact of the Green Revolution, 1960 to 2000. *Science* **300**, 758–762 (2003).
3. McArthur, J. W. & McCord, G. C. Fertilizing growth: agricultural inputs and their effects in economic development. *J. Dev. Econ.* **127**, 133–152 (2017).
4. Gollin, D., Hansen, C. W. & Wingender, A. Two blades of grass: The Impact of the green revolution. *J. Polit. Econ.* **129**, 2344–2384 (2021).
5. Fuglie, K. O. Is agricultural productivity slowing? *Glob. Food Sec.* **17**, 73–83 (2018).
6. *World Development Indicators* (World Bank Group, 2020); http://databank.worldbank.org/
7. Sheahan, M. & Barrett, C. B. Ten striking facts about agricultural input use in sub-Saharan Africa. *Food Policy* **67**, 12–25 (2017).
8. Sanchez, P. A. En route to plentiful food production in Africa. *Nat. Plants* **1**, 2014–2015 (2015).
9. *FAOSTAT Statistical Database* (Food and Agriculture Organization of the United Nations, 2020); https://www.fao.org/faostat/en/#home
10. Jayne, T. S., Mason, N. M., Burke, W. J. & Ariga, J. Review: taking stock of Africa's second-generation agricultural input subsidy programs. *Food Policy* **75**, 1–14 (2018).
11. Burke, W. J., Frossard, E., Kabwe, S. & Jayne, T. S. Understanding fertilizer adoption and effectiveness on maize in Zambia. *Food Policy* **86**, 101721 (2019).
12. Cedrez, C. B., Chamberlin, J., Guo, Z. & Hijmans, R. J. Spatial variation in fertilizer prices in sub-Saharan Africa. *PLoS ONE* **15**, 1–20 (2020).
13. Chavas, J.-P. & Holt, M. T. Economic behavior under uncertainty: a joint analysis of risk preferences and technology. *Rev. Econ. Stat.* **78**, 329–335 (1996).
14. Hurley, T., Koo, J. & Tesfaye, K. Weather risk: how does it change the yield benefits of nitrogen fertilizer and improved maize varieties in sub-Saharan Africa? *Agric. Econ.* **49**, 711–723 (2018).
15. Lobell, D. B., Banziger, M., Magorokosho, C. & Vivek, B. Nonlinear heat effects on African maize as evidenced by historical yield trials. *Nat. Clim. Change* **1**, 42–45 (2011).
16. Lipper, L. et al. Climate-smart agriculture for food security. *Nat. Clim. Change* **4**, 1068–1072 (2014).
17. Neufeldt, H. et al. Beyond climate-smart agriculture: toward safe operating spaces for global food systems. *Agric. Food Sec.* **2**, 1–6 (2013).
18. Harou, A. P., Liu, Y., Barrett, C. B. & You, L. Variable returns to fertiliser use and the geography of poverty: experimental and simulation evidence from Malawi. *J. Afr. Econ.* **26**, 342–371 (2017).
19. Vanlauwe, B. & Giller, K. Popular myths around soil fertility management in sub-Saharan Africa. *Agric. Ecosyst. Environ.* **116**, 34–46 (2006).
20. Athey, S. & Wager, S. Estimating treatment effects with causal forests: an application. *Observational Studies* **5**, 37–51 (2019).
21. Athey, S., Tibshirani, J. & Wager, S. Generalized random forests. *Ann. Stat.* **47**, 1148–1178 (2019).
22. Shepherd, M., Ghani, A., Rajendram, G., Carlson, B. & Pirie, M. Soil total nitrogen concentration explains variation in pasture response to spring nitrogen fertiliser across a single farm. *Nutr. Cycling in Agroecosyst.* **101**, 377–390 (2015).
23. Mallarino, A. P., Sawyer, J. E. & Barnhart, S. K. *A General Guide for Crop Nutrient and Limestone Recommendations in Iowa*. Report No. PM 1688 (Iowa State Univ. Extension and Outreach, 2013).
24. Mamo, M., Wortmann, C. S. & Shapiro, C. A. G03-1504 Lime Use for Soil Acidity Management (Historical Materials from Univ. of Nebraska-Lincoln Extension, 2003).
25. Feder, G., Just, R. E. & Zilberman, D. Adoption of agricultural innovations in developing countries: a survey. *Econ. Dev. Cult. Change* **33**, 255–298 (1985).
26. Morris, M., Kelly, V. A., Kopicki, R. J. & Byerlee, D. *Fertilizer Use in African Agriculture: Lessons Learned and Good Practice Guidelines* (World Bank, 2007).
27. *From Agronomic Data to Farmer Recommendations: An Economics Training Manual* 27 (CIMMYT, 1988).
28. Lobell, D. Climate Change and Agricultural Adaptation. in *Frontiers in Food Policy: Perspectives on sub-Saharan Africa* (eds. Falcon, W., Naylor, R.) Ch. 9 (Create Space, 2014).
29. Hengl, T. et al. SoilGrids250m: global gridded soil information based on machine learning. *PLoS ONE* https://doi.org/10.1371/journal.pone.0169748 (2017).
30. Hyslop, D. R. & Imbens, G. W. Bias from classical and other forms of measurement error. *J. Bus. Econ. Stat.* **19**, 475–481 (2001).

31. Yanggen, D. et al. Incentives for Fertilizer Use in sub-Saharan Africa: A Review of Empirical Evidence on Fertilizer Response and Profitability. International Development Working Paper 70 (Michigan State Univ., 1998).
32. Heisey, P. W. & Mwangi, W. in *Africa's Emerging Maize Revolution* (eds Byerlee, D. & Eicher, C.K.) 193–212 (Lynne Rienner Publishers, 1997).
33. Vanlauwe, B., Tittonell, P. & Mukalama, J. Within-farm soil fertility gradients affect response of maize to fertiliser application in western Kenya. *Nutr. Cycling Agroecosyst.* **76**, 171–182 (2006).
34. Vanlauwe, B. et al. Agronomic use efficiency of N fertilizer in maize-based systems in sub-Saharan Africa within the context of integrated soil fertility management. *Plant Soil* **339**, 35–50 (2011).
35. Korecha, D. & Barnston, A. G. Predictability of June–September rainfall in Ethiopia. *Mon. Weather Rev.* **135**, 628–650 (2007).
36. Burke, M. B., Lobell, D. B. & Guarino, L. Shifts in African crop climates by 2050, and the implications for crop improvement and genetic resources conservation. *Glob. Environ. Change* **19**, 317–325 (2009).
37. Foster, A. D. & Rosenzweig, M. R. Microeconomics of technology adoption. *Ann. Rev. Econ.* **2**, 395–424 (2010).
38. Duflo, E., Kremer, M. & Robinson, J. Nudging farmers to use fertilizer: theory and experimental evidence from Kenya. *Am. Econ. Rev.* **101**, 2350–2390 (2011).
39. Bold, T., Kaizzi, K., Svensson, J. & Yanagizawa-Drott, D. Lemon tecnologies and adoption: measurement, theory, and evidence from agricultural markets in Uganda. *Q. J. Econ.* **132**, 1055–1100 (2017).
40. Conley, T. G. & Udry, C. R. Learning about a new technology: pineapple in Ghana. *Am. Econ. Rev.* **100**, 35–69 (2010).
41. Marenya, P. P. & Barrett, C. B. State-conditional fertilizer yield response on western Kenyan farms. *Am. J. Agric. Econ.* **91**, 991–1006 (2009).
42. Suri, T. Selection and comparative advantage in technology adoption. *Econometrica* **79**, 159–209 (2011).
43. Anderson, J. R. & Hardaker, J. B. Risk aversion in economic decision making: Pragmatic guides for consistent choice by natural resource managers. in *Risk and Uncertainty in Environmental and Natural Resource Economics* (eds Wesseler, J. et al.) 171–188 (Edward Elgar Publishing, 2003).
44. Emerick, K., De Janvry, A., Sadoulet, E. & Dar, M. H. Technological innovations, downside risk, and the modernization of agriculture. *Am. Econ. Rev.* **106**, 1537–1561 (2016).
45. Dercon, S. & Christiaensen, L. Consumption risk, technology adoption and poverty traps: evidence from Ethiopia. *J. Dev. Econ.* **96**, 159–173 (2011).
46. Harou, A. et al. The joint effects of information and financing constraints on technology adoption: Evidence from a field experiment in rural Tanzania. *J. Dev. Econ.* **155**, 102702 (2022).
47. Palm, C., Neill, C., Lefebvre, P. & Tully, K. Targeting sustainable intensification of maize-based agriculture in East Africa. *Trop. Conserv. Sci.* **10**, 1–4 (2017).
48. Wortmann, C. S. et al. Maize–nutrient response functions for Eastern and Southern Africa. Dryad https://doi.org/10.5061/dryad.fg15tg2 (2018).
49. Wortmann, C. S. et al. Maize–nutrient response functions for Eastern and Southern Africa. *Agron. J.* **110**, 2070–2079 (2018).
50. International maize trial network data series. CIMMYT https://data.cimmyt.org/dataverse/root (2020).
51. Smale, M., Byerlee, D. & Jayne, T. in *An African Green Revolution* (eds Otsuka, K. & Larson, D. F.) Ch. 8 (Springer, 2013).
52. Bänziger, M., Setimela, P. S., Hodson, D. & Vivek, B. Breeding for improved abiotic stress tolerance in maize adapted to southern Africa. *Agric. Water Manag.* **80**, 212–224 (2006).
53. Van der Velde, M. et al. African crop yield reductions due to increasingly unbalanced nitrogen and phosphorus consumption. *Glob. Change Biol.* **20**, 1278–1288 (2014).
54. Bänziger, M., Betrán, F. & Lafitte, H. Efficiency of high-nitrogen selection environments for improving maize for low-nitrogen target environments. *Crop Sci.* **37**, 1103–1109 (1997).
55. Bänziger, M., Edmeades, G. & Lafitte, H. Selection for drought tolerance increases maize yields across a range of nitrogen levels. *Crop Sci.* **39**, 1035–1040 (1999).
56. Kihara, J., Bolo, P., Kinyua, M. & Piikki, K. Micronutrient deficiencies in African soils and the human nutritional nexus: opportunities with staple crops. *Environ. Geochem. Health* **42**, 3015–3033 (2020).
57. Magdoff, F. & van Es, H. *Building Soils for Better Crops* 2nd edn (Sustainable Agriculture Network, 2000).
58. Bationo, A. et al. in *Knowing the African Soils to Improve Fertilizer Recommendations* (eds Batiano, A. et al.) 19–42 (Springer, 2012).
59. Sarkar, A. N. & Wynjones, R. G. Effect of rhizosphere pH on the availability and uptake of Fe, Mn and Zn. *Plant Soil* **66**, 361–372 (1982).
60. Tully, K. L., Hickman, J., McKenna, M., Neill, C. & Palm, C. A. Effects of fertilizer on inorganic soil N in East Africa maize systems: vertical distributions and temporal dynamics. *Ecol. Appl.* **26**, 1907–1919 (2016).
61. *CELL5M: A Multidisciplinary Geospatial Database for Africa South of the Sahara* (Harvest Choice, International Food Policy Research Institute & Univ. of Minnesota, 2017).
62. You, L. et al. *What is the Irrigation Potential for Africa?* (International Food Policy Research Institute, 2010).
63. Haefele, S. M. et al. Factors affecting rice yield and fertilizer response in rainfed lowlands of northeast Thailand. *Field Crops Res.* **98**, 39–51 (2006).
64. Lobell, D. B. & Schlenker, W. Robust negative impacts of climate change on African agriculture. *Environ. Res. Lett.* **5**, 14010 (2010).
65. Uyovbisere, E. O. & Lombim, G. Efficient fertilizer use for increased crop production: the sub-humid Nigeria experience. *Fertilizer Res.* **29**, 81–94 (1991).
66. Near surface meteorological variables from 1979 to 2019 derived from bias-corrected reanalysis. European Centre for Medium-Range Weather Forecasts https://cds.climate.copernicus.eu/cdsapp/#!/dataset/derived-near-surface-meteorological-variables?tab=overview (2020).
67. Roberts, M. J., Schlenker, W. & Eyer, J. Agronomic weather measures in econometric models of crop yield with implications for climate change. *Am. J. Agric. Econ.* **95**, 236–243 (2013).
68. Food Price Monitoring and Analysis (FPMA) Tool. Food and Agriculture Organization of the United Nations https://fpma.apps.fao.org/giews/food-prices/tool/public/#/home (2021).
69. Retail fertilizer prices. AfricaFertilizer.org Initiative https://africafertilizer.org/ (2021).
70. Commodity prices. World Bank https://www.worldbank.org/en/research/commodity-markets (2021).
71. Minten, B., Koru, B. & Stifel, D. The last mile(s) in modern input distribution: pricing, profitability, and adoption. *Agric. Econ.* **44**, 629–646 (2013).
72. *Fertilizer Supply and Costs in Africa* (Chemonics and International Fertilizer Development Center, 2007).
73. Yamano, T. & Arai, A. in *The Maize Farm–Market Price Spread in Kenya and Uganda* (eds Yamano, T. et al.) 23–38 (Springer, 2011).
74. MIX Market Data. World Bank https://datacatalog.worldbank.org/search/dataset/0038647 (2021).
75. Tibshirani, J., Athey, S. & Wager, S. grf: Generalized Random Forests. R package version 1.1.0 https://CRAN.R-project.org/package=grf (2020).
76. R Core Team *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2019).
77. Ichami, S. M., Shepherd, K. D., Sila, A. M., Stoorvogel, J. J. & Hoffland, E. Fertilizer response and nitrogen use efficiency in African smallholder maize farms. *Nutr. Cycling Agroecosyst.* **113**, 1–19 (2019).
78. Imbens, G. W. & Rubin, D. B. *Causal Inference in Statistics, Social, and Biomedical Sciences* (Cambridge Univ. Press, 2015).
79. Lindsey, C. & Sheather, S. Variable selection in linear regression. *Stata J.* **13**, 65–76 (2013).
80. Spielman, D. J., Kelemwork, D. & Alemu, D. in *Food and Agriculture in Ethiopia: Progress and Policy Challenges* (eds Dorosh, P. & Rashid, S.) 84–122 (Univ. of Pennsylvania Press, 2012).
81. Nziguheba, G. et al. The African Green Revolution: results from the Millennium Villages Project. *Adv. Agron.* **109**, 75–115 (2010).
82. Porteous, O. High trade costs and their consequences: an estimated dynamic model of African agricultural storage and trade. *Am. Econ. J.: Appl. Econ.* **11**, 327–366 (2019).
83. Porteous, O. Trade and agricultural technology adoption: evidence from Africa. *J. Dev. Econ.* **144**, 102440 (2020).
84. Koo, J. & Pardey, P. G. *HarvestChoice: Supporting Strategic Investment Choices in Agricultural Technology Development and Adoption* (International Food Policy Research Institute, 2020).
85. Campolongo, F., Saltelli, A. & Cariboni, J. From screening to quantitative sensitivity analysis. A unified approach. *Comput. Phys. Commun.* **182**, 978–988 (2011).


## Acknowledgements

We are grateful to the International Food Policy Research Institute for partially funding this research. We thank C. Barrett for his leadership and guidance and J. Woodard and L. Verteramo for collaboration on earlier related research focused on Ethiopia. Initial seed funding was provided through the National Science Foundation Integrative Graduate Education and Research Traineeship (IGERT) program. We also thank S. Gameda, J. Warner, T. Benson, J. Koo, B. Minten, A. S. Taffesse and S. Rashid of the International Food Policy Research Institute for their contributions to this effort. We thank C. Wortmann for generously sharing data. We are also grateful to S. Kidanu, J. Reid and K. Bomba of the Ethiopian Agricultural Transformation Agency, M. Sheahan, D. Lobell, L. Xue, A. Perello, D. Flannelly, Z. Zeng, J. Mullen, D. Levine and E. Tjernstrom. Any errors are our own.

## Author contributions

E.B.M. conceived the initial idea for this manuscript. E.B.M., J.D.Q. and A.M.S. collaborated to create the research design. J.D.Q. prepared the historical and synthetic climate data, and A.M.S. prepared the price data. E.B.M. estimated the fertilizer response model and the profitability assessment. J.D.Q. completed the sensitivity analysis. E.B.M. led the writing of the manuscript. E.B.M., J.D.Q. and A.M.S. contributed to editing the manuscript.

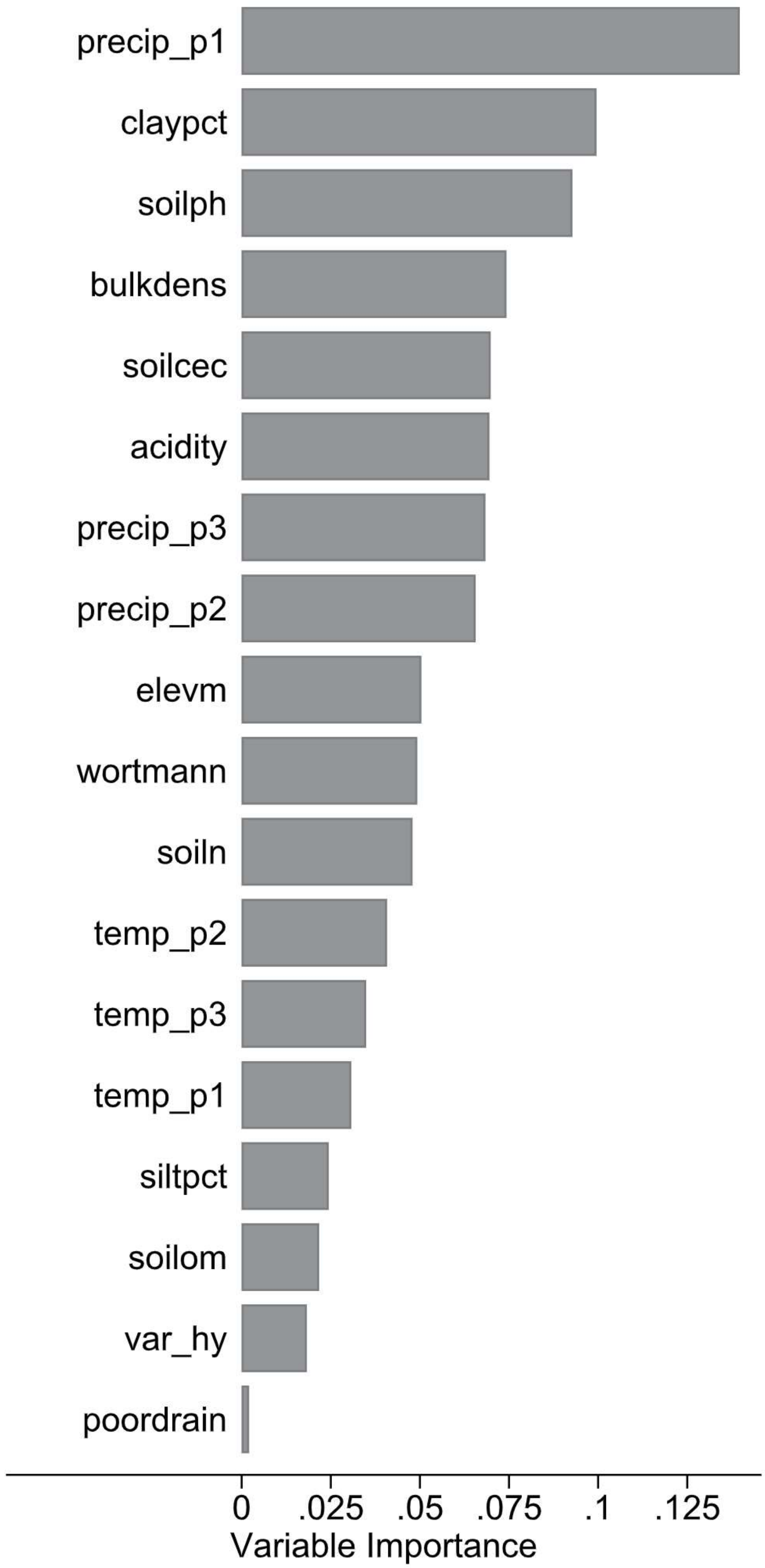
## Competing interests
The authors declare no competing interests.

## Additional information
**Extended data** is available for this paper at https://doi.org/10.1038/s43016-022-00493-z.

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s43016-022-00493-z.

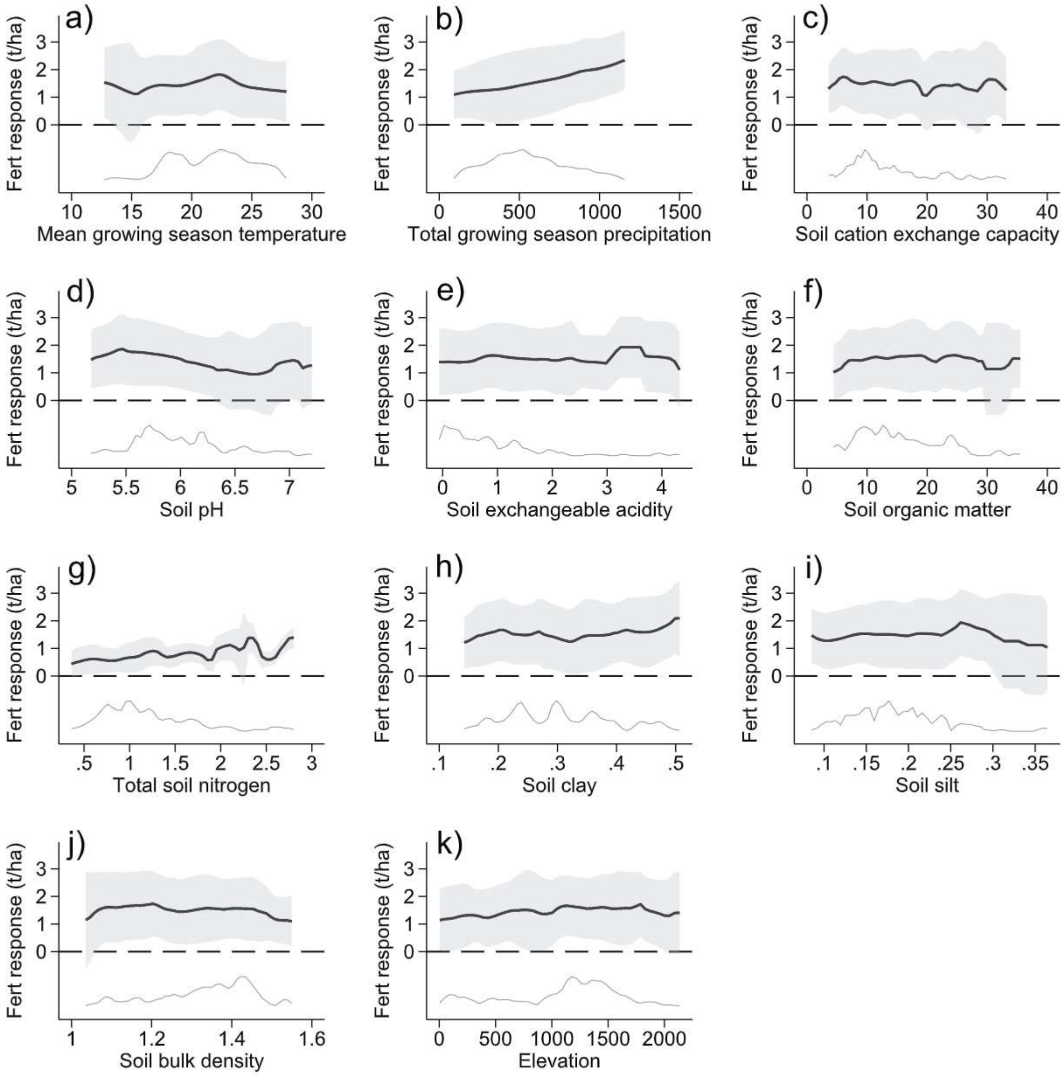
**Correspondence and requests for materials** should be addressed to Ellen B. McCullough, Julianne D. Quinn or Andrew M. Simons.

**Peer review information** *Nature Food* thanks David Lobell, Aurelie Harou and Jordan Chamberlin for their contribution to the peer review of this work.

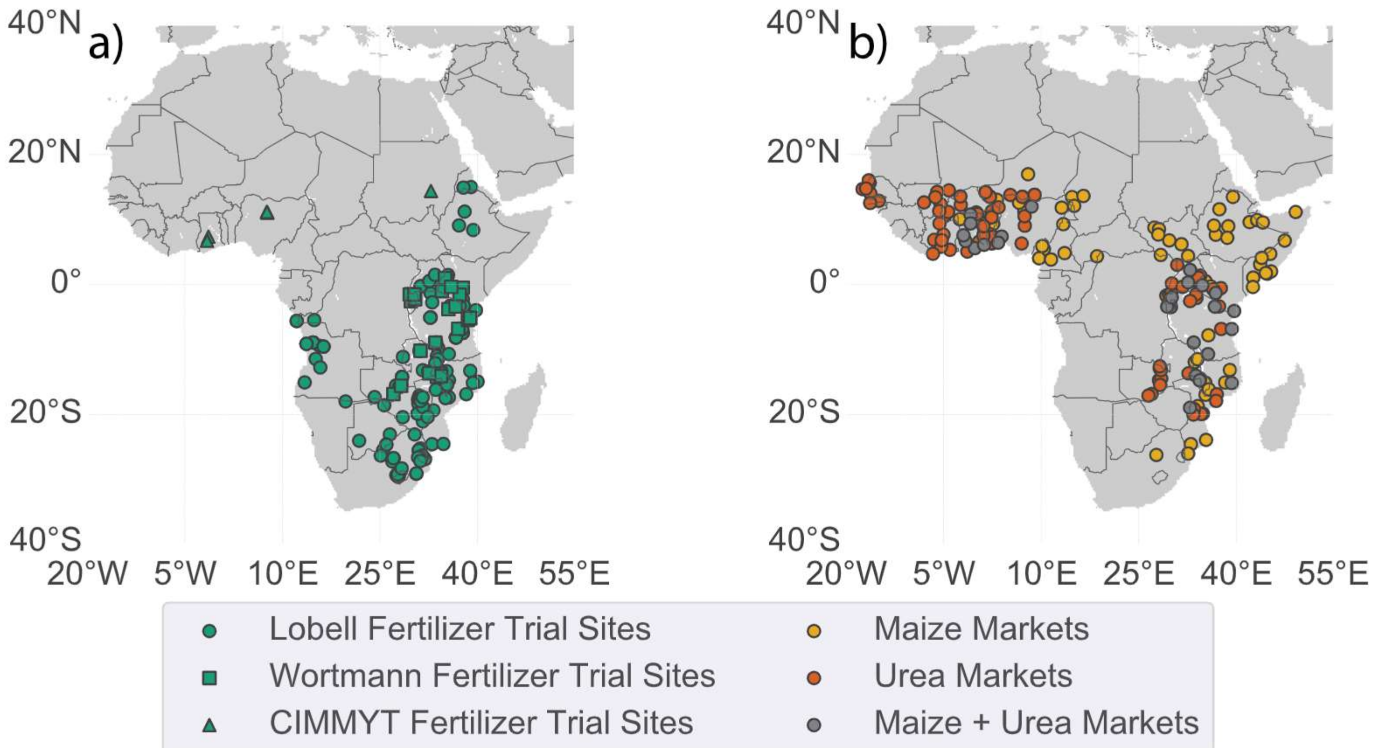



**Extended Data Fig. 1 | See next page for caption.**

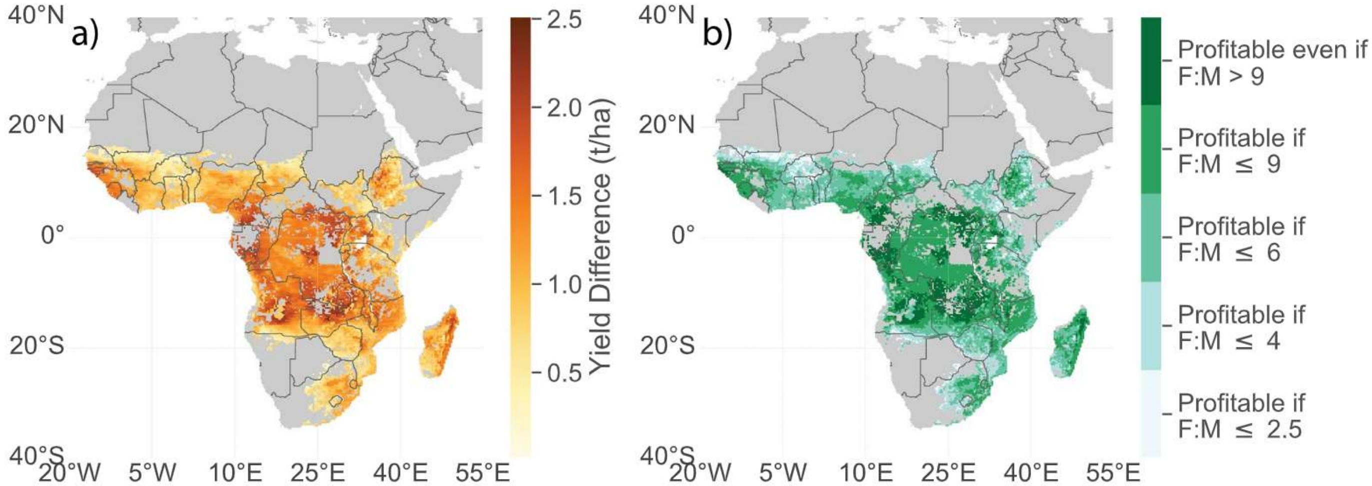
**Extended Data Fig. 1 | Ranking, by importance, of predictors included in the causal forest model used to predict the maize yield response to fertilizer, estimated using the grf package in R [75].** Importance is quantified as a weighted sum of the number of times that variable was used to split the data at each depth in the forest. The climate variables include: total precipitation in the 1st and 2nd months of the growing season (precip p1), during the 3rd month (precip p2), and during the 4th and 5th months (precip p3), average daily temperature during the 1st and 2nd months of the growing season (temp p1), during the 3rd month (temp p2), and during the 4th and 5th months (temp p3). The soil variables include: soil cation exchange capacity in centimol charge per kg soil (soilcec), soil pH as determined in a soil/water mixture (soilph), soil clay content share by volume (claypct), soil silt content share by volume (siltpct), soil bulk density in kg per cubic decimeter (bulkdens) soil exchangeable acidity in centimols charge per kg soil (acidity), soil organic matter in g per kg soil (soilom), soil nitrogen content in g per kg soil (soiln), the site's elevation (elevm), and a binary variable indicating whether the soil is characterized as having poor drainage (poordrain). The management variables include an indicator for whether a hybrid variety was used (compared with an open populated variety) and an indicator for whether the data come from the OFRA trial (wortmann, [15]) rather than the CIMMYT-supervised trials [15,52]. All continuous variables are standardized prior to estimation to a mean of zero and standard deviation of 1.



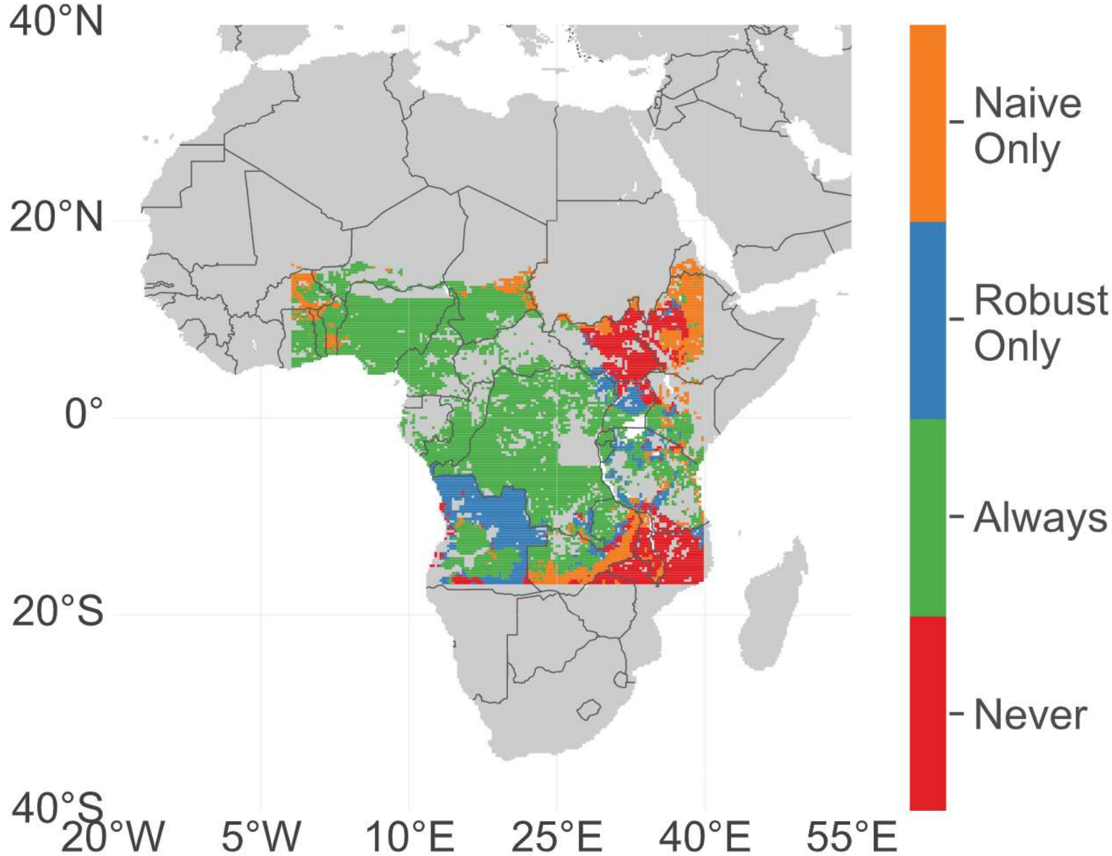

**Extended Data Fig. 2 | Mean predicted fertilizer yield response conditional on each predictor (solid line), bounded by 95% confidence intervals (shaded area).** The gray line in the lower section of each graph shows that explanatory variable's density over its range. These responses are plotted over average growing season temperature (a), total growing season precipitation (b), soil cation exchange capacity (c), soil pH (d), soil exchangeable acidity (e), soil organic matter (f), total soil nitrogen (g), soil clay content (h), soil silt content (i), soil bulk density (j), and elevation (k). The fertilizer response is predicted using a simulation dataset matching each trial site with 1000 years of climate data sampled from historical (1979-2018) climate distributions for that site, centered on predicted mean 2019 temperatures estimated from a linear trend between 1979-2018. An error is sampled for each prediction based on the standard error of the out-of-sample yield prediction for each observation. The treatment effects are plotted with a local linear regression over the range of each explanatory variable.



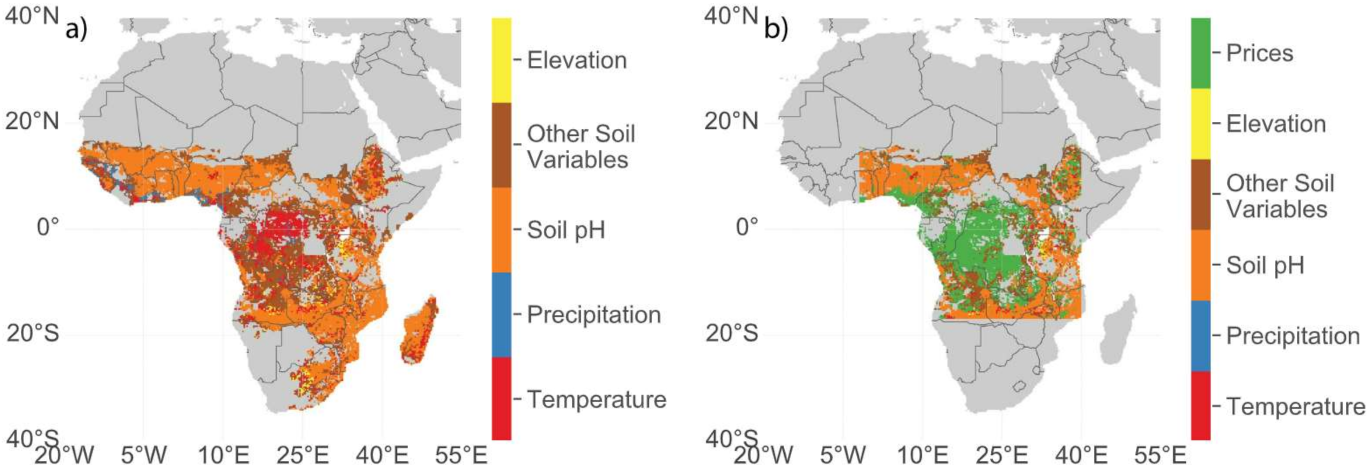

**Extended Data Fig. 3 | Map depicting locations of fertilizer trial data used to estimate the fertilizer response model and maize and urea price data used to predict fertilizer and maize prices.** Panel (a) maps the fertilizer trial data from the three different sources that we have compiled – CIMMYT-supervised trials that were included in the Lobell et al. study [15] (circles), additional CIMMYT-supervised trials accessed through the International Maize Trial Network [52] (triangles), and the OFRA trial data by Wortmann et al. [49] (squares). Panel (b) depicts the locations from which we observe monthly maize prices (in yellow), monthly urea prices (in orange) or both maize and urea prices (in grey). Maize prices were downloaded from the FAO Global Information and Early Warning System (GIEWS) dataset [68]. Urea prices were downloaded from AfricaFertilizer.org [69].




**Extended Data Fig. 4 | Predicted yield response to fertilizer that is expected at least 70% of the time (a), and the price ratio between fertilizer and maize that would be required to satisfy the "robust" profitability criteria.** Panel (a) yield responses are simulated using the random forest yield response model and the synthetic climate dataset. The color shading indicates the simulated yield response to fertilizer (t/ha) at the P=0.7 probability threshold. Given the modeled yield response in panel (a), panel (b) depicts the price ratio between fertilizer and maize that would be required in order for the IRR to exceed 30% at the P=0.7 probability threshold.




**Extended Data Fig. 5 | Alternate version of Fig. 2c, where "naive" profitability is based on IRR≥30% instead of IRR≥100%.** This figure offers an alternate version of Fig. 2c that applies the same profitability criteria as the "robust" profitability definition defined in the manuscript.




**Extended Data Fig. 6 | Sensitivity analysis at different points in the yield distribution (a-c) and IRR distribution (d-f).** Panels (a-c) show the predictor to which predicted yield response is most sensitive at the 5th percentile, 50th percentile, and 95th percentile of the yield distribution, respectively. Panels (d-f) show the input variable to which IRR is most sensitive at the 5th percentile, 50th percentile, and 95th percentile of the IRR distribution, respectively. Sensitivity to each input variable is defined using a local, one-at-a-time sensitivity analysis. The variable to which yield difference is most sensitive does not change significantly across its distribution, while the sensitivity of IRR to prices (specifically urea prices) increases as IRR increases. This, and the similarity in sensitivity of yield difference and IRR at low percentiles, suggest that yields are more limiting when profitability is low.

**Extended Data Table 1 | Summary statistics of estimation dataset by fertilizer management strategy.** The left three columns show raw means (and standard deviations) of the explanatory variables by fertilizer use, along with the normalized difference between fertilized and non fertilized observations. The right three columns show mean (and standard deviation) of the explanatory variables weighted (wgt) by the site-level probability of including some no-fertilizer observations. Technologies and management practices predictors $\overline{\theta}$ include use the Fertilized plot dummy, the OFRA trial indicator, and the hybrid variety indicator. Site characteristics predictors ($X_i$) include cation exchange capacity, soil pH, soil exchangeable acidity, soil organic matter, total soil nitrogen, soil clay share, soil silt share, soil bulk density, elevation, and a poor drainage indicator. The weather realization predictors ($\omega_i t$) include average temperature over three growing season segments and total precipitation over three growing season segments.

| | No Fert. | Opt. Fert. | Norm. Diff. | No Fert. (wgt) | Opt. Fert. (wgt) | Norm. Diff. (wgt) |
|---|---|---|---|---|---|---|
| Yield (t/ha) | 2.02<br>(1.33) | 4.32<br>(2.64) | 0.78 | 2.02<br>(1.34) | 4.38<br>(2.61) | 0.80 |
| Fertilized plot (indicator) | 0.00<br>(0.00) | 1.00<br>(0.00) | | 0.00<br>(0.00) | 1.00<br>(0.00) | |
| Temp, months 1-2 (mean, degrees C) | 23.57<br>(1.92) | 23.11<br>(2.30) | -0.16 | 23.32<br>(1.81) | 23.08<br>(2.26) | -0.08 |
| Temp, month 3 (mean, degrees C) | 22.97<br>(2.07) | 22.63<br>(2.42) | -0.11 | 22.63<br>(2.01) | 22.56<br>(2.37) | -0.03 |
| Temp, months 4-5 (mean, degrees C) | 21.54<br>(1.92) | 21.35<br>(2.20) | -0.07 | 21.29<br>(1.95) | 21.28<br>(2.19) | -0.00 |
| Precip, months 1-2 (tot, mm) | 329.13<br>(125.63) | 330.29<br>(163.01) | 0.01 | 317.74<br>(119.04) | 317.10<br>(155.05) | -0.00 |
| Precip, month 3 (tot, mm) | 177.91<br>(117.83) | 149.03<br>(99.76) | -0.19 | 164.30<br>(117.00) | 145.71<br>(97.78) | -0.12 |
| Precip, months 4-5 (tot, mm) | 148.01<br>(121.83) | 142.79<br>(134.99) | -0.03 | 145.39<br>(121.62) | 146.15<br>(133.93) | 0.00 |
| Cation exch. capacity (cmol charge/kg soil) | 12.93<br>(6.23) | 13.39<br>(7.74) | 0.05 | 13.42<br>(6.49) | 13.42<br>(7.39) | -0.00 |
| Soil pH (pH from soil/water mixture) | 5.94<br>(0.43) | 5.99<br>(0.41) | 0.07 | 5.93<br>(0.44) | 5.97<br>(0.41) | 0.07 |
| Exchangeable acidity (cmol charge/kg soil) | 0.91<br>(0.85) | 0.89<br>(0.96) | -0.02 | 0.95<br>(0.95) | 0.91<br>(1.02) | -0.03 |
| Soil organic matter (g/kg soil) | 16.87<br>(7.60) | 15.80<br>(8.21) | -0.10 | 18.26<br>(8.40) | 16.99<br>(9.01) | -0.10 |
| Total soil nitrogen (g/kg soil) | 1.27<br>(0.67) | 1.27<br>(0.65) | -0.01 | 1.41<br>(0.79) | 1.35<br>(0.72) | -0.05 |
| Soil clay (share by volume) | 0.34<br>(0.12) | 0.30<br>(0.11) | -0.24 | 0.35<br>(0.12) | 0.31<br>(0.11) | -0.24 |
| Soil silt (share by volume) | 0.18<br>(0.05) | 0.18<br>(0.06) | 0.03 | 0.18<br>(0.05) | 0.18<br>(0.06) | -0.03 |
| Soil bulk density (kg per cubic decimeter) | 1.37<br>(0.13) | 1.35<br>(0.13) | -0.11 | 1.36<br>(0.14) | 1.35<br>(0.14) | -0.05 |
| Elevation (M) | 1,144.86<br>(528.43) | 1,117.38<br>(480.61) | -0.04 | 1,188.31<br>(499.41) | 1,136.13<br>(470.16) | -0.08 |
| Poor drainage (indicator) | 0.12<br>(0.32) | 0.17<br>(0.37) | 0.10 | 0.08<br>(0.26) | 0.12<br>(0.32) | 0.11 |
| OFRA trial (Wortmann) | 0.08<br>(0.27) | 0.01<br>(0.12) | -0.22 | 0.11<br>(0.31) | 0.02<br>(0.15) | -0.25 |
| Hybrid variety | 0.62<br>(0.49) | 0.63<br>(0.48) | 0.01 | 0.60<br>(0.49) | 0.65<br>(0.48) | 0.07 |
| $N$ | 3,069 | 18,349 | | 3,069 | 18,349 | |

**Extended Data Table 2 | Datasets used to model and simulate fertilizer yield response and profitability.** This table identifies and describes the datasets used to model and simulate the yield response and profitability of fertilizer.

| Dataset Name | Dataset Original Purpose | Notes | Years Used | Source |
|---|---|---|---|---|
| Lobell Trial Data | Combines different agronomic trials to study maize response to water and temperature stress. | Management regime labeled as Drought, Low N, Low pH, Maize Streak Virus, and Optimal. We only use the Low N and Optimal trials. We only use 125kg N, 18kg P, 0kg K (which we label Optimal N) and 0kg N, 18kg P, 0kg K (which we label Low N). | 1999-2007 | [15] |
| Optimized Fertilizer Recommendations for Africa (OFRA) | Trace out fertilizer response curves for N, P, and K. | OFRA executed maize trials while varying doses of N, P, and K. We only use the subset of these trials with 120kg N, 15kg P, 0kg K (which we label Optimal N) and 0kg N, 15kg P, 0kg K (which we label Low N). | 2013-2016 | [49] |
| International Maize Trial Network | Test performance of new maize varieties. | CIMMYT sent maize seeds to collaborators in various countries, those collaborators plant the seeds in an optimal N setting and a low N setting, while holding all other factors fixed between the two trials. | 2005-2007 | [51] |
| Africa Soil Information Service (AfSIS) | Mapping effort to map soils across Africa, resolution 250 m. | We used soil cation exchange capacity, soil pH, soil exchangeable acidity, soil organic matter, soil nitrogen content, soil texture (silt and clay content), soil bulk density, and soil drainage | | [29] |
| HarvestChoice | IFPRI's HarvestChoice generates knowledge products to help guide strategic decisions to improve the well-being of the poor in sub-Saharan Africa. | We used the 16-zone agro-ecological zone (AEZ) classifications developed by the FAO and the International Institute for Applied Systems Analysis (IIASA), which are included in the HarvestChoice data product. | | [62] |
| European Centre for Medium-Range Weather Forecasts (ECMWF) | Provides bias-corrected reconstruction of near-surface meteorological variables derived from fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalyses (ERA5). | We used daily mean temperature and total precipitation data computed from hourly observations at 0.5 degree resolution. | | [67] |
| Global Information and Early Warning System (GIEWS) | FAO's GIEWS contains information and analysis on domestic prices of basic foods mainly in developing countries. | We downloaded monthly maize price data from 99 markets in 20 countries, with price data available as far back as 2000 in some markets. | 2000-2020 | [69] |
| AfricaFertilizer.org | AfricaFertilizer.org provides information on fertilizers to public and private sector with a focus on 39 sub-Saharan African countries. | We downloaded monthly urea price data from 102 markets in 17 countries, with price data available as far back as 2010 in some markets. | 2010-2020 | [70] |
| OANDA | Provides foreign exchange data for corporate and foreign exchange trading platforms. | We used OANDA's historical currency converter to convert local retail urea prices to USD/kg using monthly local currency exchange rates (https://oanda.com). | 2010-2020 | |

**Extended Data Table 3 | Fertilizer response and profitability predictions by country.** The first column shows the number of cells per country for which yield predictions are generated (each cell is approximately 770km$^2$ at the equator). The criteria for generating yield predictions are described in the methods section of the paper. The second column shows the mean yield response for each country at the probability threshold (i.e., the average over the country of the yield response to fertilizer that is expected at least 70% of the time). The third column shows what fertilizer to maize price ratio would be required for fertilizer use to be profitable at the probability threshold, given the predicted yield response at the profitability threshold. A lower fertilizer to maize price ratio indicates fertilizer is more affordable relative to maize. The fourth column shows the share of cells modeled in the country for which price data are also available. The fifth column shows the mean profitability (IRR) across cells within a country, for which price data are available, at the probability threshold. The sixth column shows the share of cells in the country (out of the total cells for which price data are available) in which profitability is robust according to the criteria described in the paper.

| | Number cells modeled | Mean yield response (t/ha) | Mean F:M ratio required for profitability | Share cells with price data | Mean IRR | Share of price data cells robustly profitable |
|---|---|---|---|---|---|---|
| Angola | 1,433 | 1.39 | 7.48 | 0.93 | 1.45 | 0.93 |
| Benin | 149 | 0.76 | 4.12 | 1.00 | 0.50 | 0.70 |
| Botswana | 53 | 0.48 | 2.58 | | | |
| Burkina Faso | 322 | 0.57 | 3.09 | 0.49 | 0.28 | 0.75 |
| Burundi | 35 | 1.27 | 6.88 | 1.00 | 1.96 | 1.00 |
| Cameroon | 459 | 1.46 | 7.90 | 1.00 | 3.53 | 1.00 |
| Central African Repub. | 204 | 1.27 | 6.86 | 1.00 | 2.12 | 0.99 |
| Chad | 588 | 0.84 | 4.54 | 1.00 | 1.28 | 0.84 |
| Cote d'Ivoire | 384 | 0.98 | 5.29 | | | |
| DRC | 2,322 | 1.52 | 8.22 | 1.00 | 2.21 | 0.99 |
| Eritrea | 46 | 0.61 | 3.30 | 1.00 | -0.24 | 0.00 |
| Ethiopia | 723 | 0.94 | 5.08 | 0.85 | -0.10 | 0.21 |
| Gabon | 99 | 1.84 | 9.95 | 1.00 | 4.66 | 1.00 |
| Ghana | 271 | 0.81 | 4.38 | 0.71 | 0.90 | 0.89 |
| Guinea | 226 | 1.27 | 6.86 | | | |
| Guinea-Bissau | 46 | 1.68 | 9.08 | | | |
| Lesotho | 28 | 0.73 | 3.92 | | | |
| Madagascar | 760 | 1.14 | 6.14 | | | |
| Malawi | 132 | 1.36 | 7.32 | 1.00 | -0.41 | 0.00 |
| Mali | 516 | 0.75 | 4.06 | | | |
| Mauritania | 62 | 0.46 | 2.49 | | | |
| Mozambique | 991 | 1.13 | 6.09 | 0.49 | -0.17 | 0.53 |
| Namibia | 166 | 0.64 | 3.48 | | | |
| Niger | 157 | 0.72 | 3.91 | 1.00 | 1.20 | 0.96 |
| Nigeria | 990 | 1.04 | 5.63 | 1.00 | 1.64 | 0.99 |
| Repub. of the Congo | 317 | 1.63 | 8.82 | 1.00 | 3.75 | 1.00 |
| Rwanda | 29 | 1.02 | 5.51 | 1.00 | 1.30 | 0.97 |
| Senegal | 189 | 0.83 | 4.50 | | | |
| Sierra Leone | 94 | 1.46 | 7.87 | | | |
| Somalia | 43 | 0.68 | 3.68 | | | |
| South Africa | 758 | 0.88 | 4.77 | | | |
| South Sudan | 356 | 0.88 | 4.74 | 0.83 | 0.64 | 0.74 |
| Sudan | 586 | 0.82 | 4.42 | 1.00 | -0.23 | 0.10 |
| Swaziland | 25 | 1.07 | 5.79 | | | |
| Tanzania | 681 | 1.15 | 6.22 | 1.00 | 0.79 | 0.76 |
| Togo | 79 | 0.78 | 4.20 | 1.00 | 0.50 | 0.75 |
| Uganda | 231 | 1.29 | 6.97 | 1.00 | 0.59 | 0.70 |
| Zambia | 881 | 1.49 | 8.03 | 0.94 | 0.41 | 0.56 |
| Zimbabwe | 497 | 0.83 | 4.48 | 0.14 | -0.41 | 0.86 |

**Extended Data Table 4 | Predicted yield response by estimation approach.** FGLS indicates the standard parametric production function estimated with a feasible generalized least squares regression, with variables selected from the full 18x18 prediction vector interacted with itself (fertilizer treatment plus 17 predictors). We select the subset of interacted predictors to include in the regression using a variable selection technique in order to minimize out of sample prediction error (root mean square error). The random forest model results are most directly comparable to the FGLS results in that the random forest model predicts yields using the fertilizer treatment variable and the additional 17 predictors. In both cases, we generate an average predicted fertilizer response by predicting yields at each data point with and without fertilizer use, then collapsing those predicted differences to the site level then averaging using site level propensity weights. RMSE is based on the difference between actual and predicted yields for the prediction dataset. The causal forest model directly estimates the yield response to the treatment variable (fertilizer use) conditional on the vector of 17 predictors, using an approach that is very similar to the random forest model. It does not generate yield predictions, nor can we calculate RMSE because we do not observe an "actual" fertilizer treatment effect. We collapse the causal forest model yield predictions also by collapsing them to the site level then averaging using site level propensity weights.

| | FGLS | Random Forest | Causal Forest |
|---|---|---|---|
| Predicted yield (F=0) | 2.45 | 3.01 | |
| | (0.84) | (0.07) | |
| Predicted yield (F=1) | 4.51 | 4.27 | |
| | (0.70) | (0.18) | |
| Predicted fertilizer response | 2.07 | 1.26 | 1.49 |
| | (0.52) | (0.12) | (0.13) |
| RMSE | 5.20 | 2.26 | |