ETHIOPIA
IFPRI

Ethiopian Development
Research Institute (EDRI)

STRATEGY SUPPORT PROGRAM | WORKING PAPER 120 | July 2018

Predicting High-Magnitude, Low-Frequency Crop Losses Using Machine Learning

An application to cereal crops in Ethiopia

Michael L. Mann, James M. Warner, and Arun S. Malik

TABLE OF CONTENTS

Abstract 1
1. Introduction 1
2. Background 2
2.1. The 2015 drought and other localized losses 2
2.2. Motivation 5
3. Methods 5
3.1. Data sources 5
3.2. Estimating crop losses 8
4. Results and Discussion 9
4.1. Maize losses 10
4.2. Wheat losses 13
4.3. Sorghum losses 17
4.4. Barley losses 18
4.5. Teff losses 18
5. Conclusions 18
References 20

LIST OF TABLES

Table 3.1. Wheat drought loss – share of observations available at different levels of loss between 2010 and 2015 from the Ethiopia Agricultural Sample Survey 6
Table 3.2. Ethiopia Agricultural Sample Survey variable names and descriptions 7
Table 3.3. Remotely sensed variable names and descriptions 7
Table 4.1. Maize – In-sample confusion matrix for reported losses ≥ 25 percent 10
Table 4.2. Maize – In-sample performance metrics for reported losses ≥ 25 percent 10
Table 4.3. Maize – Out-of-sample confusion matrix for reported losses ≥ 25 percent 10
Table 4.4. Maize – Out-of-sample performance metrics for reported losses ≥ 25 percent 10
Table 4.5. Wheat – Out-of-sample confusion matrix for reported losses ≥ 25 percent 14
Table 4.6. Wheat – Out-of-sample performance metrics for reported losses ≥ 25 percent 14
Table 4.7. Sorghum – Out-of-sample confusion matrix for reported losses ≥ 25 percent 17
Table 4.8. Sorghum – Out-of-sample performance metrics for reported losses ≥ 25 percent 17
Table 4.9. Barley – Out-of-sample confusion matrix for reported losses ≥ 25 percent 18
Table 4.10. Barley – Out-of-sample performance metrics for reported losses ≥ 25 percent 18
Table 4.11. Teff – Out-of-sample confusion matrix for reported losses ≥ 25 percent 18
Table 4.12. Teff – Out-of-sample performance metrics for reported losses ≥ 25 percent 18

LIST OF FIGURES

Figure 2.1. Recorded rainfall in 2015 compared to median rainfall per month, 2010 to 2015 3
Figure 2.2. Planted area – percent of area planted by crop and planted area damaged by cause of loss, by region, 2010 to 2015 4
Figure 2.3. Losses in 2015 reported for wheat, maize, and teff in Tigray region relative to the national average, mean percentage 4
Figure 4.1. Maize loss prediction – plot of importance of ten most important variables 11
Figure 4.2. Maize loss prediction – partial dependence plot for influence of G_mx_dates_diff_md 12
Figure 4.3. Maize loss prediction – partial dependence plot for influence of PPT_G_AUC_leading 13
Figure 4.4. Maize loss prediction – partial dependence plot for influence of G_Qnt_diff 13

Figure 4.5. Wheat loss prediction – plot of importance of ten most important variables.....14
Figure 4.6. Wheat loss prediction – partial dependence plot for influence of G_Qnt.....15
Figure 4.7. Wheat loss prediction – partial dependence plot for influence of G_mx_dates_diff_md.....16
Figure 4.8. Wheat loss prediction – partial dependence plot for influence of PPT_G_AUC_leading.....16
Figure 4.9. Wheat loss prediction – partial dependence plot for influence of G_Qnt_diff_md.....17

ABSTRACT

Timely and accurate agricultural impact assessments for droughts are critical for designing appropriate interventions and policy. These assessments are often ad hoc, late, or spatially imprecise, with reporting at the zonal or regional level. This is problematic as we find substantial variability in losses at the village-level that are missing when reporting even at the zonal level. In this paper we propose a new data fusion method combining remotely-sensed data with agricultural survey data that might address these limitations.

We apply the method to Ethiopia, which is regularly hit by droughts and is a substantial recipient of ad hoc imported food aid. We then utilize remotely-sensed data obtained near mid-season to predict substantial crop losses of greater than or equal to 25 percent due to drought at the village level for five primary cereal crops. We train machine learning models to predict the likelihood of losses and explore the most influential variables. On independent samples, the models identify substantial drought loss cases with up to 70 percent accuracy by mid- to late-September. We believe the proposed models could be used to help monitor and predict yields for disaster response teams and policy makers, particularly with further development of the models and integration of newly available high resolution remotely-sensed data, such as the Harmonized Landsat Sentinel (HLS) data set.

1. INTRODUCTION

The ability to monitor and forecast agricultural outcomes in developing countries is important in several dimensions. Forecasts of agricultural losses from drought can help communities, government agencies, and NGOs put in place measures to mitigate the effects of reduced agricultural output before food insecurity or famine strike. They can also help crop insurers anticipate the magnitudes of payouts. These benefits are especially pronounced if the forecasts can be made at a fine spatial scale and with short lead times. At a more fundamental level, monitoring of losses from droughts can be beneficial in identifying appropriate measures for adapting to changing weather patterns in the form of changing crop varieties or management techniques, such as planting dates, amongst others.

Remote-sensing-based efforts to characterize the extent and productivity of croplands has a long history. Substantial progress has been made in mapping cropland extent, crop types, irrigation status, cropping intensity, and productivity from remotely-sensed imagery, particularly for developed countries. Initial efforts, e.g., LACIE and AgRISTARS, primarily utilized remotely-sensed imagery to characterize the spatial extent and growth stage of crops, relying on models driven largely by meteorological information to predict crop yields (Idso, Jackson, and Reginato 1977; Doraiswamy et al. 2003). Subsequent efforts also exploited the well-established biophysical link between canopy spectral reflectance and net primary production (Tucker and Sellers 1986), demonstrating that satellite measurements can play a role in measuring crop yield directly. Many studies have documented highly explanatory empirical relationships between satellite measures of plant phenology, such as the Normalized Difference Vegetation Index (NDVI), and yields for a variety of crops, particularly at regional scales (Rasmussen 1992; Benedetti and Rossini 1993; Funk and Budde 2009; Becker-Reshef, Vermote, et al. 2010; Becker-Reshef, Justice, et al. 2010; Mkhabela et al. 2011). Because certain crop growth stages are critical for final yield (Butler and Huybers 2015), improved results are often seen when remotely-sensed data are used to characterize crop phenology (Bolton and Friedl 2013).

The ability to monitor agricultural outcomes is still limited in the more complex environments common to many developing countries. As an example, Ethiopia's agriculture is characterized by smallholder farmers dependent on rain-fed crops with only one to two percent of farmers having access to irrigation (Mann and Warner 2015). Improving our ability to monitor and forecast changes in agricultural

outcomes in such environments is especially important given that projections of climate change point to increasing severity and duration of extreme weather events, including drought, in large portions of the developing world (Parry et al. 2007).

The main objective of this paper is to develop scalable machine-learning models that forecast major crop losses from drought in Ethiopia at the village level relying solely on remotely-sensed data from early in the growing season. Such forecasts allow potential shortfalls in agricultural output to be anticipated before the end of the growing season. The models are applied to predict major losses from drought for each of the five principal cereal crops in Ethiopia at the level of a sub-kebele. A sub-kebele corresponds to a village of approximately 200 households, covering an average area of 24 km^2. We define major crop losses to be losses that equal or exceed a village average of 25 percent, as reported by farmers. Our models allow losses to be forecast by the date of peak greenness in any sub-kebele. The dates vary by location, therefore influencing the timing of when predictions can be made. Looking at the data from 2010 to 2015, 67 percent of sub-kebele losses for the main meher growing season can be estimated by 6 September and 93 percent by 30 September.

The data employed to generate predictions of crop loss come from three sources: (1) precipitation data are from the Climate Hazards Group Infrared Precipitation Station (CHIRPS) database; (2) data on hydrological availability of water and available energy for plant growth are from FEWS NET (2017); and (3) NDVI data are derived from two 16-day MODIS products (MOD13Q1, MYD13Q1) from the Aqua and Terra satellites (Didan and Huete 2006). The models we develop provide a framework for utilizing even higher resolution vegetative index imagery, such as the forthcoming Harmonized Landsat Sentinel data, at a national scale. We also augment a suite of algorithms used to extract, summarize, and organize remotely-sensed data and prepare it for spatiotemporal analysis (Mann et al. 2018).

The machine learning models we employ are of the random forest variety. They are trained using data on reported average crop losses at the sub-kebele level for the years 2010 to 2015 collected by Ethiopia's Central Statistical Agency (CSA). The Agricultural Sample Survey (AgSS) captures data from over 45,000 households, making it one of the largest annual household agricultural surveys in the world. We focus on the five principal cereal crops grown in Ethiopia: maize, wheat, sorghum, barley, and teff.

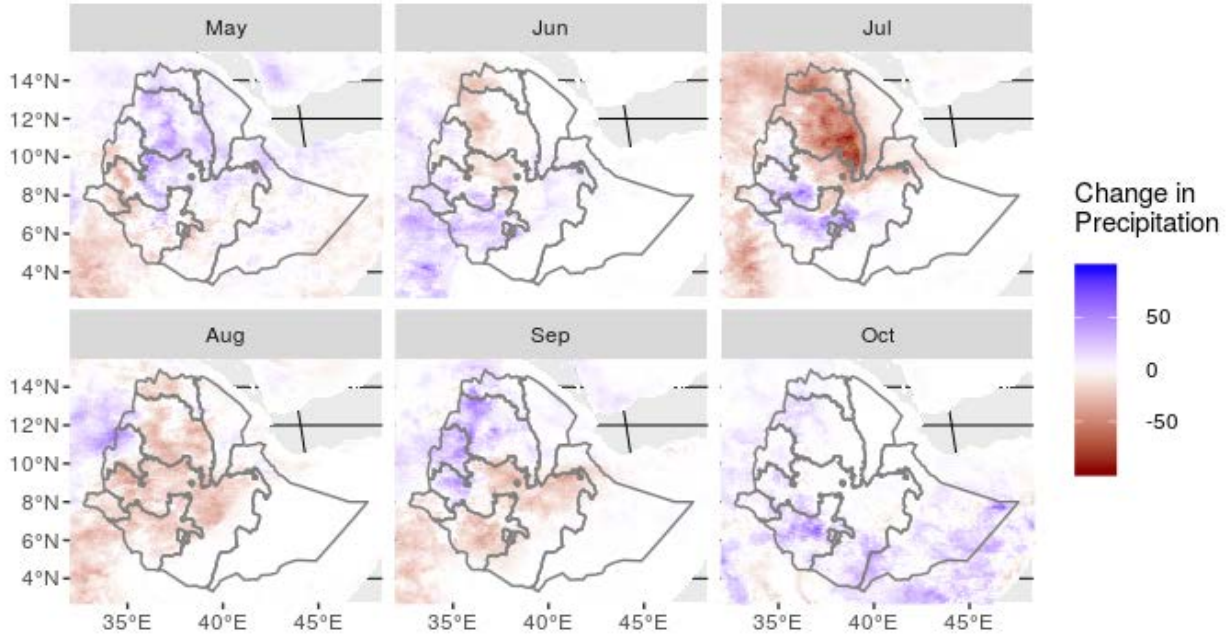
The AgSS data together with the remotely-sensed data from the weather stations and satellites mentioned above provide rich information, permitting the coupling of observations of agricultural outcomes with dense observations of variables, such as plant greenness and the timing of rains. This data fusion allows us to create models that, once trained, can predict crop losses relying only on remotely-sensed observations from satellites and weather stations.

2. BACKGROUND

2.1. The 2015 drought and other localized losses

Motivated by a long history of monsoon failures across the Sahel, forecasters and aid teams anxiously watched eastern Africa at the onset of a significant El Nino event in late 2014. Below average rainfalls between 2015 and 2017, particularly across the low lying pastoral communities of south-eastern Ethiopia, triggered an international effort to provide food aid to roughly 8.5 million people in the summer of 2017 (ReliefWeb 2017; UNOCHA 2017). The understated response from the Ethiopian government had led to accusations, as early as 2016, that the government was downplaying the severity of the drought (Schemm 2016; 2017). However, there are many indications that the drought was less severe than expected and that it primarily affected areas of lesser significance for national grain production.

Figure 2.1. Recorded rainfall in 2015 compared to median rainfall per month, 2010 to 2015



Source: Climate Hazards Group Infrared Precipitation Station (CHIRPS) database

Looking at the deviations of rainfall for 2015 compared to longer term patterns (Figure 2.1), we can see that lower than expected rains began in June, peaked in July, and subsided by September. The largest shortfalls are centered on the northern and eastern areas of the country with much of it in pastoral areas near Mekele and near Dire Dawa or in high elevation upland areas near Debre Birhan. The timing of these shortfalls is critical because planting for the major meher growing season typically starts in May and continues into late June, depending on crop type. Setbacks early in the season could undermine the germination of seed, potentially leading to failed crops. Early crop failure is often followed by farmers attempting to plant drought-resistant or shorter duration crops that would reach peak growth later in the season.

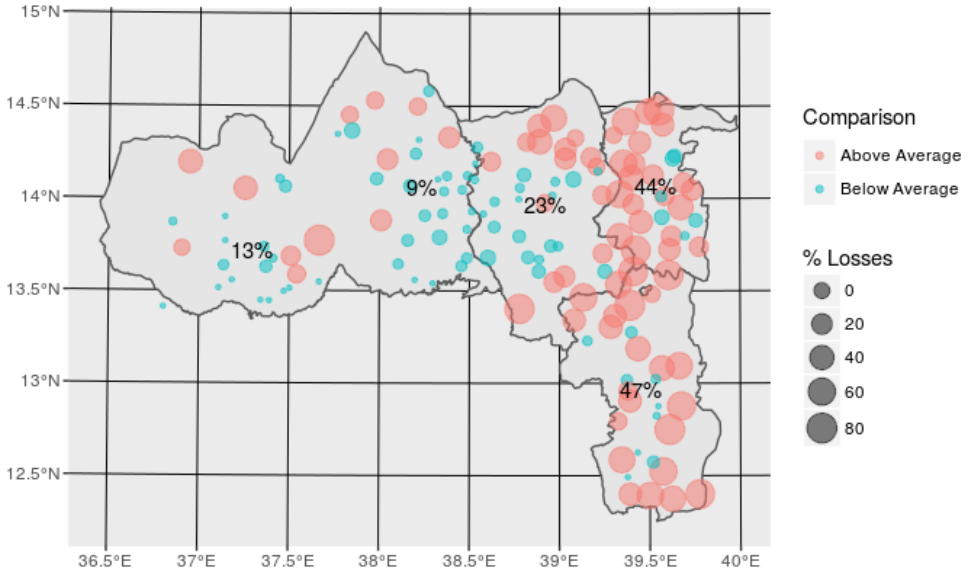
Despite the reality of a severe and crippling drought in food insecure low-lying areas, there was scant evidence of food shortages nationally. A recent study, for example, observed no significant change in grain prices in the country as a whole (Bachewe, Yimer, and Minten 2016). The lack of a price response would indicate that, while food insecure areas were hit hard, largely livestock was affected. The most productive highland areas were mostly unaffected or experienced only minor losses. An early rapid assessment performed by Tufts University researchers indicated that losses of production would exceed 18 percent nationally (AKLDP 2016). However, this was contradicted by the arguably more accurate measure of a 2 percent loss, estimated from AgSS crop cuts, and released later in 2016 (AgSS 2016).

Figure 2.2. Planted area – percent of area planted by crop and planted area damaged by cause of loss, by region, 2010 to 2015

Source: Agricultural Sample Survey, various years

Nonetheless, there was a significant up-tick in farmer reported drought losses in 2015. Looking at the top panel of Figure 2.2, the average planted area damaged due to drought increased fivefold in Tigray, and fourfold in Amhara, albeit in their northern and eastern areas. The diversity and composition of crop types is shown in the graphs of percentage of area planted by crop in the bottom panel. Interestingly, likely in response to the drought, Tigray had a sudden up-tick in 2015 in the planting of sorghum, a naturally drought-resistant plant.

Figure 2.3. Losses in 2015 reported for wheat, maize, and teff in Tigray region relative to the national average, mean percentage



Source: Agricultural Sample Survey

Looking at the spatial pattern of losses in Tigray compared to the national mean, we can see two general patterns: a clear clustering of losses on the eastern border of the province and scattered, but substantial, losses over the eastern zone (Figure 2.3). The latter pattern of losses, displaying substantial intra-zonal variability, presents a considerable challenge for current early warning systems, which aggregate above-average and below-average losses at the zonal level. Looking at the zonal means, which are the value labels presented in Figure 2.3, we can see that Western Tigray zone has pockets of well-above-average losses, but these are averaged out by a larger number of areas with below-average losses.

2.2. Motivation

This paper attempts to develop a replicable methodology for identifying multiple types of crop losses by mid-season at a high level of spatial resolution. This acuity will provide better insights into the effects of droughts at multiple scales: local, zonal, and regional. We believe these tools can be used to better design and target local, national, and international responses.

3. METHODS

3.1. Data sources

Survey data - Agricultural Sample Survey (2010-2015)

Data on crop losses for the years 2010 to 2105 was obtained from Ethiopia's Central Statistical Agency's Agricultural Sample Survey (AgSS) database, which was chosen for its annual collection, spatial coverage, and unique sampling structure. The AgSS is an annual, government-administered survey that measures agricultural production in Ethiopia. The survey interviews approximately 45,000 farmers each year and collects information on household demographics, farm management, and planting, harvesting and sales at the plot level.

Typically, about 20 farm households are randomly sampled from small, village-level areas of approximately 200 households, termed "sub-kebeles". A kebele is the smallest, official administrative division in Ethiopia, with approximately 1,000 households.[1] Each kebele is divided into five sub-kebeles, based roughly on population. There are over 75,000 sub-kebeles in the country.

For this study, we construct a sub-kebele-level panel data set for the meher growing season for five crops for the years 2010 to 2015.[2] We were able to construct the panel data set because approximately 75 percent of all sub-kebeles were sampled by CSA in each of the six meher crop seasons[3]. This effectively comprises a base of five relatively favorable crop seasons, from 2010 to 2014, followed by a drought year, 2015. We did not apply sampling weights because approximately 25 percent of the sub-kebeles sampled by CSA were omitted; moreover, our primary interest is in estimating local effects, not total production.

For each household surveyed and for each crop, CSA collects reports on the percentage of total crop production lost due to a range of causes, including drought, management error, pests, and weather, e.g., winds or flooding. We focus on crop losses due to drought because such losses can be most readily assessed from satellite imagery by characterizing crop phenology. We choose drought losses of 25 percent or more as a cutoff value given the decreasing number of observations at higher loss levels. Table 3.1 shows

[1] Ethiopia has four official administrative divisions. These are, in order of decreasing geographic size: regions, zones, woredas and kebeles. There are approximately 12 regions, 88 zones, 690 woredas and 15,000 kebeles.
[2] The two main annual crop seasons in Ethiopia are the meher and belg seasons. The meher season is the main season and produces 90 to 95 percent of the nation's total cereals output. (AgSS 2016)
[3] For each sub-kebele, CSA samples households with replacement. Therefore, the households surveyed in a given sub-kebele can vary from one year to the next. Thus, our dataset does not constitute a panel at the household level, though it does at the sub-kebele level.

that moving the cutoff above 25 percent significantly reduces the number of observations available for training for wheat, for example.

Table 3.1. Wheat drought loss – share of observations available at different levels of loss between 2010 and 2015 from the Ethiopia Agricultural Sample Survey

	Level of loss, percent			
	10	**25**	**50**	**75**
Share of observations	0.39	0.18	0.05	0.02

Source: Analysis by authors

We study the five principal Ethiopian field crops (maize, wheat, sorghum, barley, teff), because of their importance in value terms and their geographically widespread adoption. Accordingly, our estimates of crop losses are for these five crops alone and not for total agricultural output. We have no reason to believe that our estimates for these five crops are not generally representative of all crop production in terms of the effects of droughts, but a more detailed analysis would be needed to substantiate this presumption.

For each crop, data on crop losses was aggregated to the sub-kebele level by averaging losses across the 20 households in a sub-kebele. This yielded approximately 1,780 observations at the sub-kebele level for each year, and a total of 10,680 observations for all six years. The number of observations varies by crop since not all sub-kebeles grow all five crops. Data for each crop year was checked for consistency and cleaned. In the course of data cleaning, we dropped any sub-kebeles for which crop loss data was not available for at least four years.

Remotely Sensed Data - Greenness, Precipitation and Evaporation

Normalized Difference Vegetation Index. For the sake of simplicity and replicability, we focus on the NDVI from among a large list of potential spectral indices, since it is commonly used for agriculture applications (Lessio, Fissore, and Borgogno-Mondino 2017; Mann and Warner 2017; Mann and Warner 2015; Funk and Budde 2009). NDVI is sensitive to the amount of chlorophyll in a given pixel and is commonly used to estimate plant productivity and health in agricultural applications. Given the relatively small scale of agriculture in Ethiopia and the lack of high frequency observations from other satellites (such as HLS), we derive the NDVI using the 250m vegetation products from the MODIS satellites, specifically the two 16-day MODIS products (MOD13Q1, MYD13Q1) from the Aqua and Terra satellites (Didan and Huete 2006). Due to the staggered nature of acquisition, these products are treated as partially overlapping windows representing eight-day periods (Doraiswamy, Stern, and Akhmedov 2007). We find that the combination of these two products provides a stable and informative time series. After removal of cloudy and other flagged low-quality cells, we remove all non-agricultural cells through the use of a purpose-built, machine-learning-based land classification for Ethiopia for the period of interest.

Water Availability Variables. Data on precipitation (PPT) was obtained from the Climate Hazards Group Infrared Precipitation Station (CHIRPS) database. Data is collected as precipitation by dekad (Funk et al. 2015). There are three dekads in a month, the first two being 10 days each, and the third being the remaining days of the month.

Previous research has shown hydrological availability of water and available energy for plant growth to be important factors in crop growth. Accordingly, we include monthly estimates of potential evaporation (PET) and actual evapotranspiration anomaly (ETA). Actual evapotranspiration (AET) is the sum of transpirated water through plants plus evaporation from soils and water surfaces. In other words, it is the total amount of water moved through a system. AET is a correlated with vascular plant productivity, biomass accumulation, and regrowth (Major 1967; Mann et al. 2016). The ETA variable used in this study is the current AET compared to the mean for 2003 to 2013. Thus, ETA is a proxy for drought, as low current

values of AET would correspond to reductions either in available energy to move water, or in water itself. Potential evapotranspiration (PET) is an estimate of the total amount of water that could be moved through the system if water was not a limiting factor. As such it is an ideal indicator of available (or excessive) energy in the form of heat, but commonly includes the impact of wind speed, pressure, and humidity, among others. More simply, PET can be viewed as a proxy for temperature. Both ETA and PET are available for download through FEWS NET (2017).

Variable Definitions

Tables 3.2 and 3.3 provide a key to the variables obtained and derived from the AgSS, GIS, and remotely-sensed data.

Table 3.2. Ethiopia Agricultural Sample Survey variable names and descriptions

Name	Description
Prefixes	
*MAIZE_*	Variables based on maize plots
*WHEAT_*	Variables based on wheat plots
*BARLEY_*	Variables based on barley plots
*SORGHUM_*	Variables based on sorghum plots
*TEFF_*	Variables based on teff plots
Suffixes	
_diff	First difference
_diff_md	Difference from the median of all observations
Variables from AgSS	
DAMAGE_DROUGHT_AREA_P	Percentage of crop production reported as having experienced drought damage
Other Geographic Variables	
elevation	Elevation (m)
z_code	Zonal codes

Source: Prepared by authors

Table 3.3. Remotely sensed variable names and descriptions

Name	Description
Prefixes	
*G_*	Calculated from meher growing season values
*ETA_*	Actual evapotranspiration
*PET_*	Potential evapotranspiration
*PPT_*	Precipitation
Suffixes	
_diff	First difference
_diff_md	Difference from the median value for all observations
NDVI basic summary statistics	
G_max	Meher maximum values of NDVI
G_Qnt	Meher 90th percentile value of NDVI
G_plant_date	Estimated date of meher planting based on NDVI phenology
G_mx_dates	Estimated date of maximum NDVI during meher season
Integrated summary statistics	
G (PET, ETA, PPT) _AUC_leading	Area under the curve of NDVI (or PET, ETA, PPT) for the ascending part of the NDVI curve during the meher season
Comparison norms	
G (PET, ETA, PPT) _diff_md	Difference between the median meher NDVI (or PET, PET) and observed value

Source: Prepared by authors

3.2. Estimating crop losses

Preprocessing

Data aggregation. Because our unit of analysis is the sub-kebele, remotely-sensed data must be spatially aggregated to match the sub-kebeles of the country. We calculate the mean of the raw remotely-sensed data values for all available agricultural pixels at the sub-kebele-level and further extract statistics of interest from the resulting eight-day or dekadal time series.

Summarizing Temporal Data. One of the primary challenges in utilizing high-frequency time series data, such as eight-day NDVI, to estimate low-frequency events (seasonal crop losses from drought) is reconciling this temporal mismatch when formalizing the relationship between sources of data. Low-frequency properties must be extracted from the high-frequency time series. In our setting, the low-frequency properties extracted must be relevant to identifying and characterizing the aspects of plant phenology that affect crop yields until the middle of the growing season. We start by capturing the relevant phenological features of each crop through 41 metrics summarizing eight-day NDVI data as well as ten-day PPT, PET, and AET. These measures span the portion of the meher growing season from the estimated planting date to the harvest date. The mean planting date is estimated at the sub-kebele level so as to capture any differences in timing due to elevation, crop choice, or other management considerations. It is worth emphasizing that given our interest in forecasting losses early in the growing season, we limit the variables in our models to those that rely only on data up until the middle of the growing season.

We utilized phenological timing observed from NDVI to summarize PPT, PET, and AET for the relevant portion of the growing season based on estimated planting date and the date of the maximum observed NDVI value. For instance, the variable that includes 'PPT_leading' in its name is calculated using the PPT time series only for the period between the estimated planting date and the date of the maximum observed NDVI value.

We estimate three classes of statistics: summary statistics, e.g., mean, maximum, variance; integrated summary statistics, e.g., area under the curve for the leading half of the growing season; and deviation-from-the-norm statistics, e.g., deviations of a given statistic from its 90th historical percentile. Deviations from the norm provide a comparison between expected values and the current year and, thus, should provide insight into low frequency events like drought.

Predicting Crop Losses from Drought

As noted, our objective is to determine if crop losses due to drought at the sub-kebele level can be predicted using remotely-sensed data alone at mid-season. To do so we make use of machine-learning methods in the form of random forests. We define our dependent variable to be a binary variable that reflects whether losses due to drought for a given crop were 25 percent or higher. Given the binary nature of our dependent variable, our prediction problem is one of classification: Can we predict whether crop losses due to drought will be 25 percent or higher for a given sub-kebele? Our model takes the following form:

$$Y_{cn} = f\Big(\sum(X + X\_diff + X\_diff\_md), \sum(\rho\_AUC\_leading), \sum(\sigma\_dates + \sigma\_dates\_diff\_md), elevation, zone\_code\Big)$$

$$Y_{ci} = \begin{cases} 1, \text{ if drought loss} >= 0.25 \\ 0, \text{ otherwise} \end{cases} \quad (1)$$

Where Y_{ci} is a binary variable where losses for crop *c* due to drought in sub-kebele *i* are greater than or equal to 25 percent. The model is estimated using a non-linear function of two measures of NDVI, G_mx and G_qnt in *X*; the sum of G; PET, ETA, and PPT up until peak greenness in ρ; the dates of initial green-up (planting date) and of peak NDVI, and their difference from the median date in σ;[4] elevation; and zonal codes.

Estimation Strategy

Random Forests Description. Random forests (RF) are a flexible ensemble learning method that aggregates across multiple classification or regression trees, each based on a randomized subset of variables. Each tree is formed through hierarchical splitting, whereby each binary split (e.g. NDVI < 0.1) is identified as the most informative, as defined by the greatest reduction in a Gini impurity.[5] This ensemble method improves performance by finding the mode or median of 'weak learners'. This process in turn allows RFs to avoid over-fitting and, therefore, to perform consistently out-of-sample and with noisy data sets.

Training vs. Testing Data. Before running any models, we divide the data into two sets. A random sample of 85 percent of all sub-kebeles are retained for the training and tuning of models, while the remaining 15 percent are set aside for out-of-sample testing of our models. Because some missing values were present, a K-nearest-neighbor (KNN) algorithm was used to impute these values. The KNN algorithm used here centers all data by default, therefore all variables have been centered around their mean values. Training and testing data were imputed separately to avoid information leakage.

Training Data: Cross Validation and Tuning. To enhance the out-of-sample performance of our models, we employ parameter tuning and cross validation when training our models. We limit tuning to choosing the number of features (variables) randomly selected for use in each tree. For each crop, we do a grid search and find the parameter tuning that provides the best performance. To ensure that these parameter tunings work well, we utilize out-of-bag (OOB) error. OOB errors can provide unbiased estimates of error on out-of-sample data because random forests already use bootstrapping for fitting individual trees. As such, it can take the place of the validation or test error and is cheaper to compute than using cross validation. The model is tuned to optimize performance of the Kappa coefficient.[6] The optimal parameter value is the one with the highest Kappa value and is specific to each crop model.

Testing Data: Out-of-Sample Performance. As noted above, we withheld 15 percent of sub-kebeles from our training sample and used the data for these "test" sub-kebeles to assess the true out-of-sample performance of our tuned RF models. Specifically, we test our tuned RF models by predicting whether losses were 25 percent or higher, by crop, for each of the test sub-kebeles using remotely-sensed data alone. We then compare our predictions to actual losses reported in the AgSS survey. We employ three common performance metrics: a) overall accuracy, 2) Kappa coefficient, and 3) the recall rate.[7]

4. RESULTS AND DISCUSSION

In this section we present the results from our machine learning models. Results for both in-sample and out-of-sample predictions are presented. For each crop, the models predict whether a sub-kebele experienced crop losses of 25 percent or more due to drought in a given year, with the predictions based solely on remotely-sensed vegetation and weather data obtained by mid-season.

[4] Datetime variables are measured in days since some arbitrary origin date, in this case, days since Jan 1, 1960.
[5] Gini impurity measures how often a randomly chosen observation from a set would be incorrectly labeled if it was randomly labeled according to the distribution of labels in the subset. Gini impurity reaches zero when all cases fall into a single category.
[6] Kappa is a performance indicator similar to the percentage of observations that are correctly classified but controls for the fact that categories with large numbers of observations are easier to guess by random chance.
[7] The recall rate is the percentage of sub-kebeles with 'substantial drought losses' that are correctly classified by our models.

For the sake of brevity, from hereon we will refer to crop losses of 25 percent or more due to drought as "substantial losses." We convey the accuracy of our predictions by presenting confusion matrices as well as the values of the three measures mentioned earlier.

4.1. Maize losses

The confusion matrix for our in-sample predictions of substantial maize losses is presented in Table 4.1 and the values of the three performance metrics are presented in Table 4.2. The corresponding results for our out-of-sample predictions are presented in Tables 4.3 and 4.4.

Table 4.1. Maize – In-sample confusion matrix for reported losses ≥ 25 percent

	FALSE	TRUE
FALSE	6,680	52
TRUE	1	572

Source: Analysis by authors

Table 4.2. Maize – In-sample performance metrics for reported losses ≥ 25 percent

	Accuracy	Kappa	Recall
Value	0.993	0.952	0.917

Source: Analysis by authors

We see high overall accuracy for our in-sample predictions of 99.3 percent. The recall rate is 0.917, indicating that 91.7 percent of the 572 sub-kebeles that experienced substantial losses in a given year were correctly predicted using the remote sensing data alone as having done so. No cases were falsely predicted as having experienced substantial losses. In-sample performance for the other crops is similar to that for maize and for the sake of brevity will not be reported for the remaining crops.

Table 4.3. Maize – Out-of-sample confusion matrix for reported losses ≥ 25 percent

	FALSE	TRUE
FALSE	1,165	30
TRUE	28	65

Source: Analysis by authors

Table 4.4. Maize – Out-of-sample performance metrics for reported losses ≥ 25 percent

	Accuracy	Kappa	Recall
Value	0.95	0.67	0.68

Source: Analysis by authors

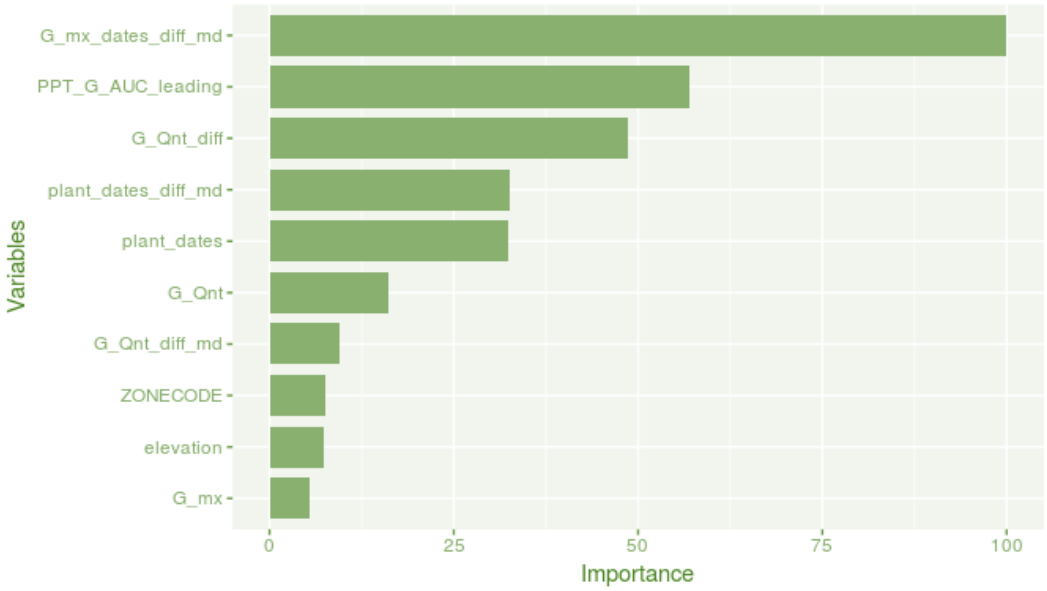
Encouragingly, we also see a high level of accuracy for our out-of-sample predictions, with an overall accuracy of 95 percent. We see reasonable identification of reported substantial losses with a recall rate of 0.68, indicating that 68 percent of all substantial loss cases were correctly predicted, or 65 out of 95 such cases. Twenty-eight cases were falsely predicted as substantial loss cases. In addition to the high recall rate, we were able to predict 1,165 out of 1,193 healthy maize harvests, or 98 percent of such harvests.

Clearly, further work could help better identify low-probability, substantial loss events. However, the importance of this research should not be understated. Early identification of localized crop losses could be very useful for policy action, and adoption of the more accurate 30m HLS satellite data could markedly improve results.

We can also examine the relative importance of each variable to the model, as well as its estimated non-linear relationship with the dependent variable through partial dependence plots. We first examine

the relative importance of each variable in the random forest through the role it played in Gini coefficient loss.

Figure 4.1. Maize loss prediction – plot of importance of ten most important variables



Source: Analysis by authors

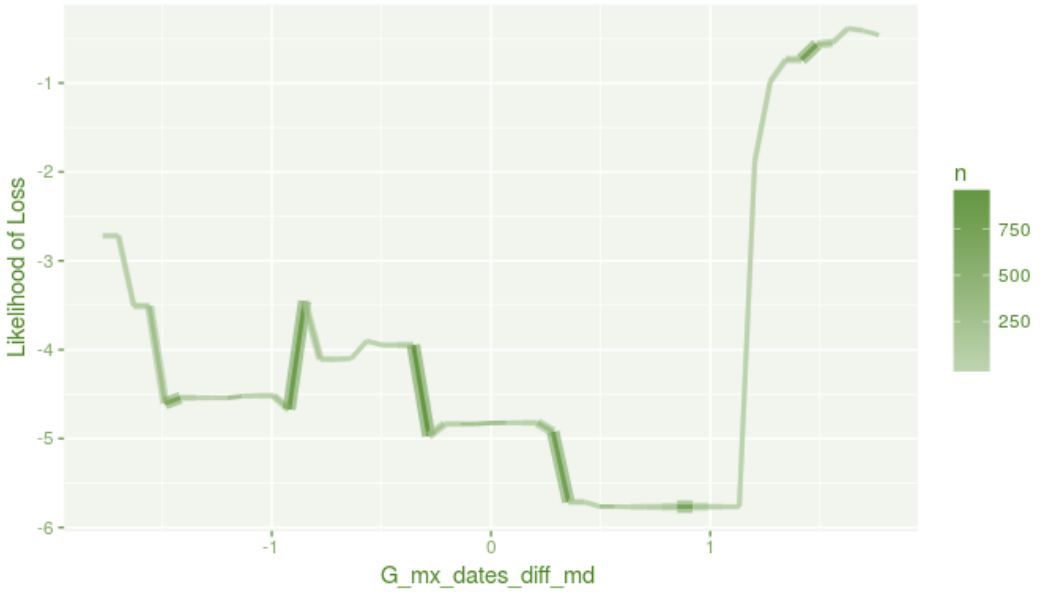
From Figure 4.1, we can see that the most important variables are primarily date variables. A large weight is placed on date of maximum greenness (G_mx), as measured by NDVI, the date of initial green-up (plant_dates), and their differences from their median values (G_mx_dates_diff_md and plant_dates_diff_md). The importance of these variables can be explained as follows. If the July rains fail, farmers are forced to adjust their planting dates, which in turn delays when maximum greenness is reached.

Other variables of interest include the sum of precipitation from the planting date up until the date of maximal greenness (PPT_G_AUC_leading). Precipitation during this early period in the growing season has clear implications for successful germination. Moreover, this variable likely provides a good indication, after controlling for the date of initial green-up, of whether the planting was successful. G_Qnt_diff, the first-difference in 90th percentile greenness and the difference between G_Qnt and its median value are also of importance. Note that "_diff" indicates a 'first-difference' or the difference between the current year's value of a variable and the previous year's value. Drops in near maximal greenness likely provide some evidence that plant health suffered throughout the growing season. Each one of these variables will have a specific non-linear response, with curves being estimated by actual plant and weather patterns observed remotely. Moreover, the emergent properties of some farmer management practices, such as delayed planting, can also be examined.

We can now examine partial dependence plots[8] to help understand the response of reported losses to variables of importance. Figure 4.2 shows the plot for G_mx_dates_diff_md, which is the difference between the date of the current year's maximum NDVI value and the median value of the date. As mentioned above, this variable would strongly track the date of successful planting and germination, as well as plant vigor following successful germination. Note that higher values on the vertical axis indicate a greater likelihood of substantial crop losses.

[8] These plots are graphical visualizations of the marginal effect of a given variable (or multiple variables) on an outcome.

Figure 4.2. Maize loss prediction – partial dependence plot for influence of G_mx_dates_diff_md

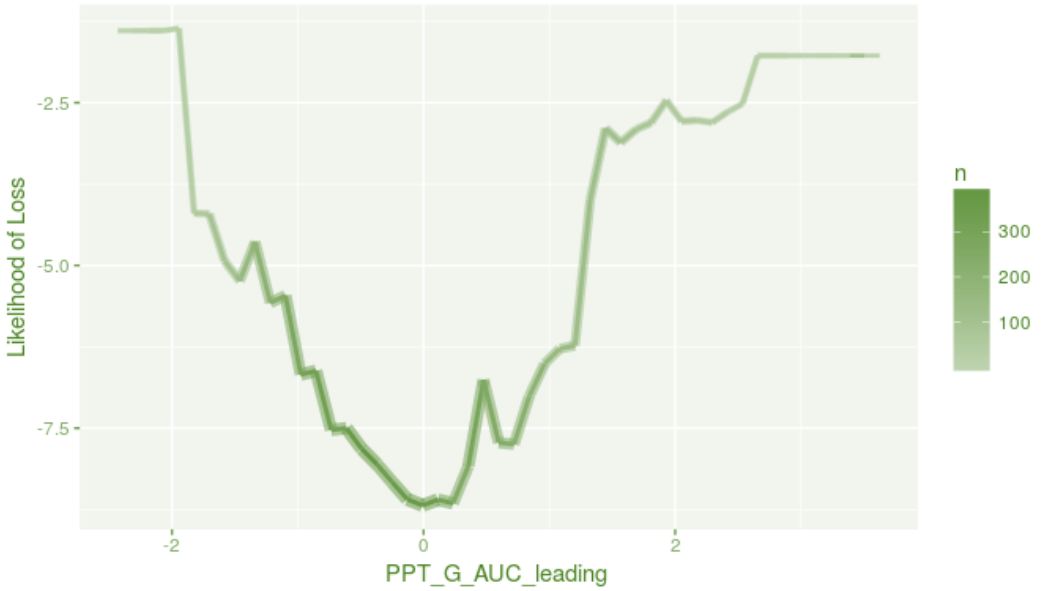


Source: Analysis by authors

Figure 4.2 shows that the likelihood of reporting maize losses decreases slowly as G_mx_dates_diff_md approaches zero and increases dramatically after the value of 1. Looking at the number of observations, indicated by a darker color and thicker line, we can surmise that the erratic step-like function on the left is due to the NDVI observations being collected every 8 days. Generally, sub-kebeles with earlier than normal maximum greenness dates have mixed success, likely depending on geography and weather, with some being successful and others not. Interestingly, according to regional experts, having been given advanced warning of potential failed rains, the Ethiopian government encouraged farmers to plant earlier than normal. From the curve above, we see indications of a mixed outcome, with those planting the earliest seeing a high likelihood of substantial loss, and those planting a few weeks later seeing more mixed outcomes. As would be expected, the lowest likelihood of losses was reached around the median date. This indicates that many areas of the country experienced normal rains and successfully kept to normal or slightly later than average planting dates. On the other hand, those sub-kebeles with much later than expected maximum greenness dates experienced the greatest losses. As discussed earlier, late maximum dates point to areas with one or more failed plantings. In these areas farmers would have likely tried to replant later in the season.

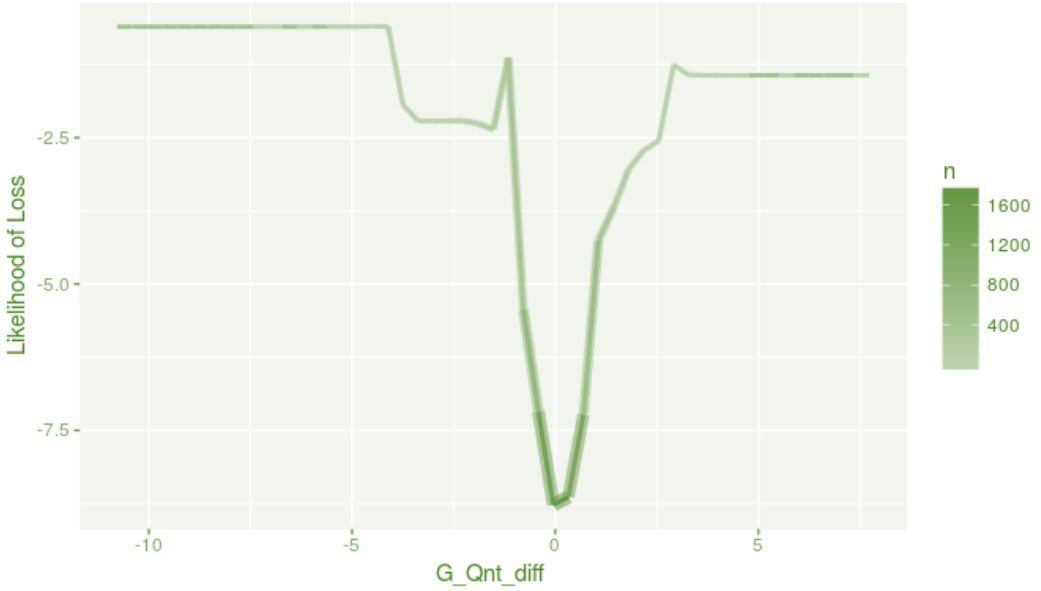
Figure 4.3 shows the partial dependence plot for PPT_G_AUC_leading. This variable tracks the effects of total precipitation from initial green-up until the maximal NDVI value. Note that this variable is automatically centered around its mean. We can see from the figure that the likelihood of crop losses declines dramatically up until its centered mean value (0). This values clearly correspond to below average, or drought, conditions. Additionally, years with above average levels of precipitation also correspond to increased losses due to excess precipitation or damaging storms, although this is likely the result of a few observations. The fact that above average precipitation would correspond to ‘drought losses’ likely relates to the fact that farmer response to questionnaires is imperfect. Farmers might reasonably conflate losses related to drought with losses related to weather or rain more generally.

Figure 4.3. Maize loss prediction – partial dependence plot for influence of PPT_G_AUC_leading



Source: Analysis by authors

Figure 4.4. Maize loss prediction – partial dependence plot for influence of G_Qnt_diff



Source: Analysis by authors

Figure 4.4 shows how the first-difference in G_Qnt, the 90th percentile value of NDVI, influences substantial losses of maize crops. The highest likelihood of losses is observed with large declines in maximal values of NDVI relative to the previous year. This change likely reflects declines in plant health, due to water stress or other climatic stressors, such as high heat. Positive changes in maximal values might reflect poor management, such as failure to control weeds or narrow row spacing due to broadcast seeding.

4.2. Wheat losses

We turn from maize to wheat. As noted earlier, we will focus on the performance of our out-of-sample predictions since the in-sample predictions were very similar to those for maize. Table 4.5 presents the confusion matrix, and Table 4.6 presents our three performance metrics for predictions of substantial losses in wheat.

Table 4.5. Wheat – Out-of-sample confusion matrix for reported losses ≥ 25 percent

	FALSE	TRUE
FALSE	832	19
TRUE	8	35

Source: Analysis by authors

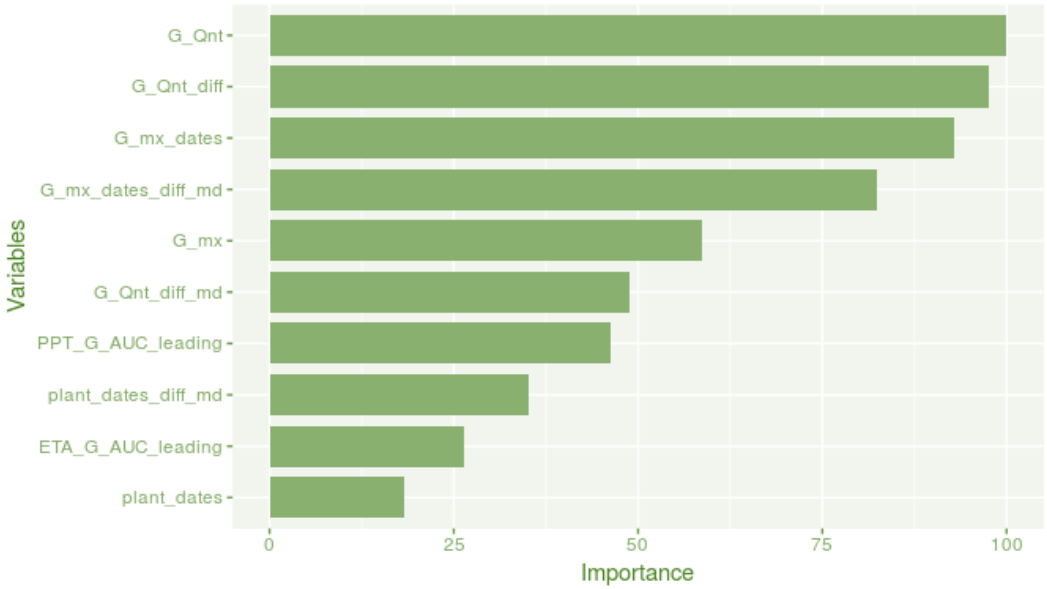
Table 4.6. Wheat – Out-of-sample performance metrics for reported losses ≥ 25 percent

	Accuracy	Kappa	Recall
Value	0.97	0.71	0.65

Source: Analysis by authors

Overall, we find a high level of accuracy for our out-of-sample predictions, with overall accuracy of 97 percent. Compared to maize, we obtain somewhat weaker prediction of substantial wheat losses, with a recall rate of 0.65 instead of 0.68, i.e., 65 percent of all substantial wheat loss cases, or 35 out of 54 such cases, were correctly predicted. Eight cases were falsely predicted as substantial crop losses. Though the recall rate is not very high, we are able to predict healthy wheat harvests well, with 832 out of 840, or 99 percent of such harvests correctly predicted.

Figure 4.5. Wheat loss prediction – plot of importance of ten most important variables



Source: Analysis by authors

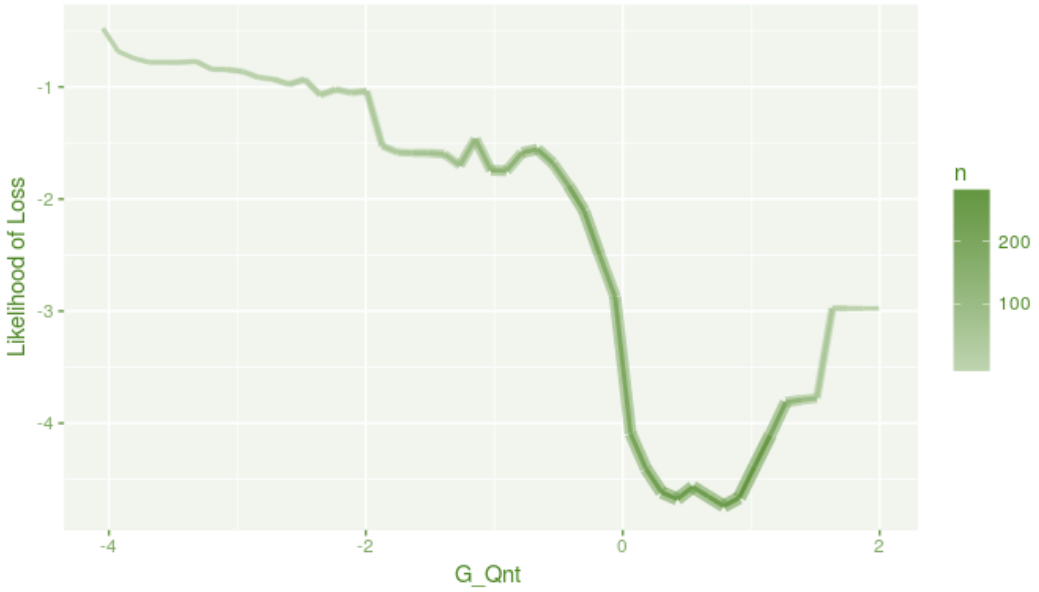
Figure 4.5 shows the 10 most important variables in our random forests, as measured by the role each variable played in Gini coefficient loss. The set of variables is similar to that for maize. However, the absolute and relative importance of the variables differ. This suggests that our models are able to utilize the same broad set of phenological and weather variables to capture differences between crop types and the factors that influence losses for each crop. The variables G_Qnt_diff and G_mx_dates are once again important, suggesting a robustness in their predictive ability.

We can see that G_Qnt has the highest relative importance. G_Qnt is the 90th percentile of NDVI values in the meher season. High NDVI values reflect a measure of plant health, with higher G_Qnt values corresponding to better plant health (such as larger leaf area and lower water stress) at some point in the early growing season. We can also see that many of the most important variables are variables summarizing NDVI, apart from the dates of these phenological indictors. Again, the model has identified and emphasized variables that help distinguish losses related to wheat plantings.

Interestingly, comparing Figures 4.5 and 4.1, we can see that planting and maximal NDVI dates are of less important for predicting substantial wheat losses than maize losses. This might relate to the fact that many of the most productive wheat-growing areas of Ethiopia were largely unaffected by the failure of the July rains during the 2015 drought. Instead, these areas were more affected by later declines in rainfall in August (Warner, Mann, drought paper draft). As a result, many wheat growing sub-kebeles did not need to change their planting dates. Instead plant health was more affected by water stress after a successful germination.

We now turn to the partial dependence plots for the prediction of substantial losses in wheat and start by examining the plot for G_Qnt in G_Qnt.

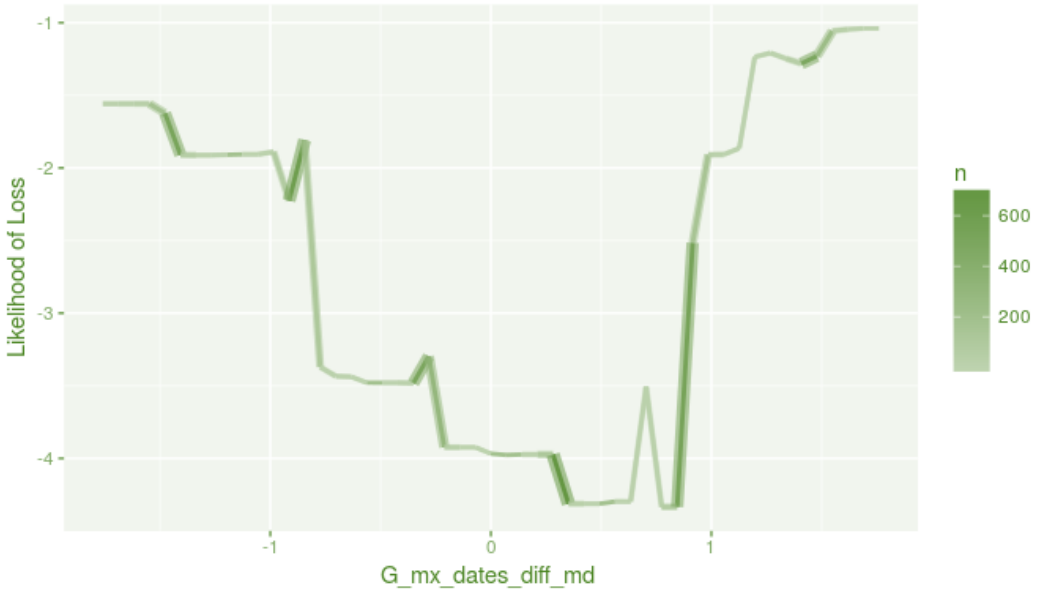
Figure 4.6. Wheat loss prediction – partial dependence plot for influence of G_Qnt



Source: Analysis by authors

Figure 4.6 shows the likelihood of substantial maize losses decreases rapidly as maximal NDVI values approach the mean of zero, with a minimum around 0.5. The likelihood of losses increases again at the upper end of the range of values for G_Qnt. These losses at the upper end might reflect poor management, such as failure to control weeds or narrow row spacing. The consistency across the two different crop models (maize and wheat) is an indication that these models are picking up meaningful nonlinear relationships between plant phenology and the likelihood of loss.

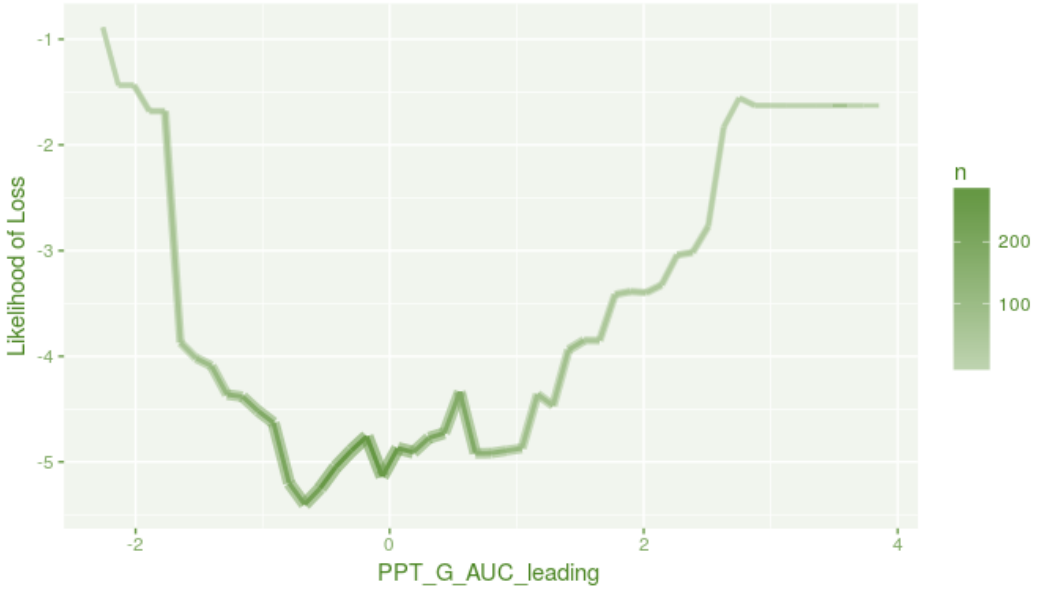
Figure 4.7. Wheat loss prediction – partial dependence plot for influence of G_mx_dates_diff_md



Source: Analysis by authors

Figure 4.7 shows the partial dependence plot for G_mx_dates_diff_md. Comparing Figures 4.7 and 4.2 reveals similarities and differences between the likelihood of losses for wheat and maize. We see that the likelihood of losses at very negative levels of G_mx_dates_diff_md is higher for wheat (note the differences in the y-axis intercept values) and declines more rapidly as the change in maximal greenness date increases. However, for both crops, the curve reaches a minimum above 0. Losses increase again with much later than average planting and therefore later maximal NDVI values.

Figure 4.8. Wheat loss prediction – partial dependence plot for influence of PPT_G_AUC_leading

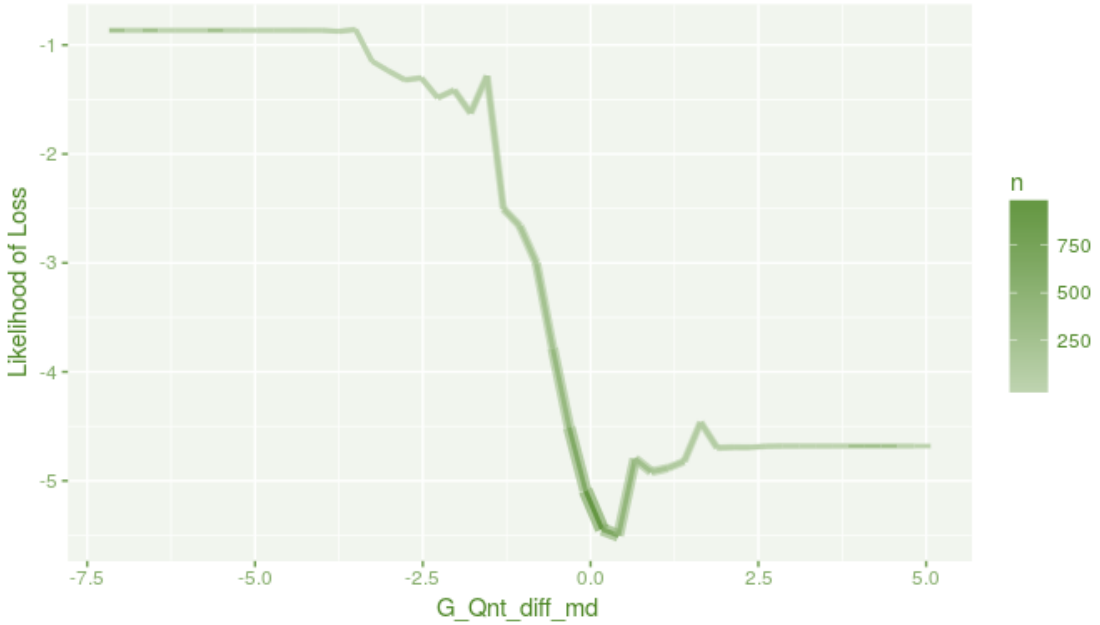


Source: Analysis by authors

The partial dependence plot for PPT_G_AUC_leading for wheat is presented in Figure 4.8. This variable captures total rainfall until maximum greenness. As for maize (Figure 4.3), we see a U-shaped response, with the highest drought losses observed with below average (centered around 0) precipitation. Higher than average precipitation is also linked to reported drought losses. Again, this may reflect

confusion in survey responses, or it might point to the fact that the timing of rains, rather than the quantity of rainfall, is the key driver of substantial losses. The importance of this timing is reflected across models by the plant_date and G_mx_date variables.

Figure 4.9. Wheat loss prediction – partial dependence plot for influence of G_Qnt_diff_md



Source: Analysis by authors

Although the variables we focused on have a common U-shaped function, there are other functional forms. For instance, the difference between the current 90th percentile value and the median 90th percentile of previous years, G_Qnt_diff_md (Figure 4.9), shows a strong downward trend. The likelihood of significant loss is extremely high with 90th percentile greenness values significantly lower than normal. The likelihood of loss drops rapidly as 90th percentile greenness approaches zero and tapers up slightly after.

4.3. Sorghum losses

Turning to sorghum losses, the confusion matrix and performance metrics for our out-of-sample predictions of substantial losses for this crop are presented in Tables 4.7 and 4.8.

Table 4.7. Sorghum – Out-of-sample confusion matrix for reported losses ≥ 25 percent

	FALSE	TRUE
FALSE	878	20
TRUE	6	59

Source: Analysis by authors

Table 4.8. Sorghum – Out-of-sample performance metrics for reported losses ≥ 25 percent

	Accuracy	Kappa	Recall
Value	0.97	0.81	0.75

Source: Analysis by authors

Overall, we find a high level of accuracy in our out-of-sample predictions, with overall accuracy of 97 percent. Prediction of substantial sorghum crop losses was good, with a recall rate of 0.75, indicating that 75 percent, or 59 out of 79, of substantial loss cases were predicted correctly. Six cases were falsely

predicted as substantial crop losses. Healthy sorghum harvests were correctly predicted in 878 out of 884 times, or 99 percent of the time.

4.4. Barley losses

We now examine the ability of our models to predict substantial out-of-sample crop losses for barley. The confusion matrix is presented in Table 4.9 and our three performance metrics are presented in Table 4.10

Table 4.9. Barley – Out-of-sample confusion matrix for reported losses ≥ 25 percent

	FALSE	TRUE
FALSE	829	35
TRUE	17	49

Source: Analysis by authors

Table 4.10. Barley – Out-of-sample performance metrics for reported losses ≥ 25 percent

	Accuracy	Kappa	Recall
Value	0.94	0.62	0.58

Source: Analysis by authors

We obtain a high overall accuracy of 94 percent. Prediction of substantial losses is less accurate than for maize, wheat, and sorghum: the recall rate for barley is 0.58, indicating that 58 percent, or 49 out of 84, of all substantial loss cases were predicted correctly. Seventeen cases were falsely predicted as substantial loss cases. Though the recall rate is relatively low, we were able to correctly predict healthy barley harvests in 829 out of 846 cases, or 98 percent of the time.

4.5. Teff losses

Finally, we turn to the prediction of teff crop losses out-of-sample. The confusion matrix is presented in Table 4.11 and the three performance metrics in Table 4.12

Table 4.11. Teff – Out-of-sample confusion matrix for reported losses ≥ 25 percent

	FALSE	TRUE
FALSE	1,051	33
TRUE	20	44

Source: Analysis by authors

Table 4.12. Teff – Out-of-sample performance metrics for reported losses ≥ 25 percent

	Accuracy	Kappa	Recall
Value	0.95	0.60	0.57

Source: Analysis by authors

As for all the other cereals examined, we obtain a high overall accuracy level of 95 percent. The recall rate is similar to that for barley, with a recall rate of only 0.57, indicating that 57 percent, or 44 out of 77, of all substantial loss cases were predicted correctly. Twenty cases were falsely predicted as substantial loss cases. Although the recall rate is relatively low, we are able to identify healthy teff harvests in 1,051 out of 1,071 times, or 98 percent of the time.

5. CONCLUSIONS

In this study we examine the feasibility of using remotely-sensed data and machine learning models to predict by the middle of the growing season which sub-kebeles will suffer substantial crop losses at the time of harvest. We develop a custom set of algorithms to summarize changes in plant phenology, as

measured by NDVI, precipitation, and potential and actual evapotranspiration up until the date of maximal greenness, which corresponds approximately to the middle of the growing season. The algorithms to generate these statistics have all been made freely available through Github.[9]

Considering the difficulty of the task and available data, our results are extremely encouraging. At the sub-kebele level, the out-of-sample recall rate, or the ability to correctly predict substantial crop losses out-of-sample, ranged from 57 to 75 percent, with a median value of 65 percent. The overall accuracy of the predictions ranged from 94 to 97 percent.

Our models allow crop losses to be predicted sooner than can be done with existing models and at a higher level of spatial resolution. Specifically, we can predict losses by crop in any sub-kebele by the date of peak greenness. The date varies by location, therefore influencing precisely when predictions can be made. Looking at data from between 2010 and 2015, we can predict 67 percent of sub-kebele crop losses by 6 September or 93 percent by 30 September.

We also provide evidence that our models pick up meaningful nonlinear relationships that respond to both crop loss and the expected differences in crop types. With more development, models like these could be used to provide accurate assessments in early in the growing season of agricultural losses at the end of the season. Moreover, these predictions made in the first half of the growing season provide a critical opportunity for identifying the need for and planning early national and local interventions. We believe efforts such as these can augment and enhance current early warning systems and help guide field-based rapid assessments.

Though the predictive accuracy of our models is reasonably high, we believe that these models could be significantly improved with the use of 30m Harmonized Landsat Sentinel (HLS) data. This data should be released in late 2018 and will have five-day coverage. This likely would provide a high enough frequency to avoid clouds and provide accurate estimates of NDVI curves at a much higher resolution than the 250m cells used in the research here. The methods developed in this study have been specifically designed to transfer to the new HLS data stream, thus greatly improving our model accuracy while minimizing further development time.

[9] https://github.com/mmann1123

REFERENCES

AgSS (Agricultural Sample Survey). 2016. "Agricultural Sample Survey 2016/2017 (2009 E.C.) Report on Area and Production of Major Crops." *Feed the Future*. The Federal Democratic Republic of Ethiopia, Central Statistical Agency.

AKLDP (Agriculture Knowledge, Learning, Documentation and Policy Project). 2016. *Food Security in Ethiopia in 2016: Analysing crop production and market function after the main Meher agricultural season*. USAID/Ethiopia, Addis Ababa: AKLDP.

Bachewe, F., F. Yimer, and B. Minten. 2016. *Agricultural Prices During Drought in Ethiopia: An Updated Assessment Using National Producer Data (January 2014 to June 2016)*. (Vol. 88). Intl Food Policy Res Inst.

Becker-Reshef, I., C. Justice, M. Sullivan, E. Vermote, C. Tucker, A. Anyamba, J. Small, E. Pak, E. Masuoka, J. Schmaltz, M. Hansen, K. Pittman, C. Birkett, D. Williams, C. Reynolds, and B. Doorn. 2010. "Monitoring Global Croplands with Coarse Resolution Earth Observations: The Global Agriculture Monitoring (GLAM) Project." *Remote Sensing* 2 (6): 1589–1609.

Becker-Reshef, I., E. Vermote, M. Lindeman, and C. Justice. 2010. "A Generalized Regression-Based Model for Forecasting Winter Wheat Yields in Kansas and Ukraine Using Modis Data." *Remote Sensing of Environment* 114 (6): 1312–23.

Benedetti, R., and P. Rossini. 1993. "On the Use of NDVI Profiles as a Tool for Agricultural Statistics: The Case Study of Wheat Yield Estimate and Forecast in Emilia Romagna." *Remote Sensing of Environment* 45 (3): 311–26.

Bolton, D. K., and M. A. Friedl. 2013. "Forecasting Crop Yield Using Remotely Sensed Vegetation Indices and Crop Phenology Metrics." *Agricultural and Forest Meteorology* 173: 74–84.

Butler, E. E., and P. Huybers. 2015. "Variations in the Sensitivity of Us Maize Yield to Extreme Temperatures by Region and Growth Phase." *Environmental Research Letters* 10 (3): 034009.

Didan, K., and A. Huete. 2006. *MODIS Vegetation Index Product Series Collection 5 Change Summary*. TBRS Lab, the University of Arizona.

Doraiswamy, P. C., A. J. Stern, and B. Akhmedov. 2007. "Crop Classification in the Us Corn Belt Using Modis Imagery." In *Geoscience and Remote Sensing Symposium,* 2007. IEEE International, 809–12.

Doraiswamy, P. C., S. Moulin, P. W. Cook, and A. Stern. 2003. "Crop Yield Assessment from Remote Sensing." *Photogrammetric Engineering & Remote Sensing* 69 (6): 665-674.

FEWS NET. 2017. *Data Portals*. Accessed August 12, 2017. http://fews.net/data

Funk, C., and M.E. Budde. 2009. "Phenologically-Tuned MODIS NDVI-Based Production Anomaly Estimates for Zimbabwe." *Remote Sensing of Environment* 113 (1): 115–25.

Funk, C., P. Peterson, M. Landsfeld, D. Pedreros, J. Verdin, S. Shukla, G. Husak, J. Rowland, L. Harrison, A. Hoell, and J. Michaelsen. 2015. "The Climate Hazards Infrared Precipitation with Stations—a New Environmental Record for Monitoring Extremes." *Scientific Data,* 2, 150066.

Idso, S. B., R. D. Jackson, and R. J. Reginato. 1977. "Remote-Sensing of Crop Yields." *Science* 196 (4285): 19-25.

Lessio, A., V. Fissore, and E. Borgogno-Mondino. 2017. "Preliminary Tests and Results Concerning Integration of Sentinel-2 and Landsat-8 Oli for Crop Monitoring." *Journal of Imaging* 3 (4): 49.

Major, J. 1967. "Potential Evapotranspiration and Plant Distribution in Western States with Emphasis on California." *American Association of Advanced Science* 86.

Mann, M. L., and J. M. Warner. 2017. "Ethiopian wheat yield and yield gap estimation: A spatially explicit small area integrated data approach." *Field Crops Research* 201: 60–74.

Mann, M. L., and J. Warner. 2015. *Ethiopian Wheat Yield and Yield Gap Estimation: A Small Area Integrated Data Approach.* Addis Ababa, Ethiopia: International Food Policy Research Institute.

Mann, M. L., E. Batllori, M. A. Moritz, E. K. Waller, P. Berck, A. L. Flint, L. E. Flint, and E. Dolfi. 2016. "Incorporating Anthropogenic Influences into Fire Probability Models: Effects of Human Activity and Climate Change on Fire Activity in California." *PLoS One* 11 (4): e0153589.

Mann, M.L., B. Kramer, F. Ceballos, M. Robles, J. Gray, and E. Melaas. 2018. *Modeling Wheat Yields in India: Describing & Exploiting Spatiotemporal Variability with Panel Regression*. Unpublished, Washington, DC: Intl Food Policy Res Inst.

Mkhabela, M. S., P. Bullock, S. Raj, S. Wang, and Y. Yang. 2011. "Crop Yield Forecasting on the Canadian Prairies Using Modis NDVI Data." *Agricultural and Forest Meteorology* 151 (3): 385–93.

Parry, M., O. Canziani, J. Palutikof, P. J. van der Linden, C. E. Hanson, eds. 2007. *Climate Change 2007: Impacts, Adaptation and Vulnerability*. Vol. 4. Cambridge University Press, Cambridge.

Rasmussen, M. S. 1992. "Assessment of Millet Yields and Production in Northern Burkina Faso Using Integrated NDVI from the AVHRR." *International Journal of Remote Sensing* 13 (18): 3431–42.

ReliefWeb. 2017. “Ethiopia: Drought - 2015-2017.” Accessed November 12, 2017. https://reliefweb.int/disaster/dr-2015-000109-eth

Schemm, P. 2016. “Ethiopia Is Facing a Devastating Drought, and Food Aid May Soon Run Out.” *The Washington Post*. Accessed November 28, 2017. https://www.washingtonpost.com/sf/world/2016/02/22/history-repeats-itself-in-ethiopia/?noredirect=on&utm_term=.659494979a81

Schemm, P. 2017. “Ethiopia Is Facing a Killer Drought. but It’s Going Almost Unnoticed.” *The Washington Post*. Accessed November 12, 2017. https://www.washingtonpost.com/news/worldviews/wp/2017/05/01/ethiopia-is-facing-a-killer-drought-but-its-going-almost-unnoticed/?utm_term=.926109370d62

Tucker, C. J., and P. J. Sellers. 1986. “Satellite Remote Sensing of Primary Production.” *International Journal of Remote Sensing* 7 (11): 1395–1416.

UNOCHA (United Nations Office for the Coordination of Humanitarian Affairs). 2017. “Mid-Year Review, Ethiopia Humanitarian Requirements Document”, July 2017. https://reliefweb.int/sites/reliefweb.int/files/resources/ethiopia_humanitarian_requirements_document_mid-year_review_2017.pdf.

About the Author(s)

Michael L. Mann is an Assistant Professor in the Department of Geography at George Washington University, Washington, DC, USA; **James M. Warner** is a Research Coordinator and office head of Eastern and Southern Africa Office at the International Food Policy Research Institute, based in Addis Ababa, and **Arun S. Malik** is a Professor in the Department of Economics at George Washington University.

About ESSP

The Ethiopia Strategy Support Program is an initiative to strengthen evidence-based policymaking in Ethiopia in the areas of rural and agricultural development. Facilitated by the International Food Policy Research Institute (IFPRI), ESSP works closely with the government of Ethiopia, the Ethiopian Development Research Institute (EDRI), and other development partners to provide information relevant for the design and implementation of Ethiopia's agricultural and rural development strategies. For more information, see http://www.ifpri.org/book-757/ourwork/program/ethiopia-strategy-support-program; http://essp.ifpri.info/; or http://www.edri-eth.org/.

The ESSP Working Papers contain preliminary material and research results from IFPRI and/or its partners in Ethiopia. The papers are not subject to a formal peer review. They are circulated to stimulate discussion and critical comment.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
1201 Eye Street, NW | Washington, DC 20005-3915 USA
T: +1.202.862.5600 | F: +1.202.862.5606
Email: ifpri@cgiar.org | www.ifpri.org

ETHIOPIAN DEVELOPMENT RESEARCH INSTITUTE
P.O. Box 2479, Addis Ababa, Ethiopia
T: +251.11.550.6066; +251.11.553.8633 | F: +251.11.550.5588
Email: info@edri-eth.org | www.edri-eth.org

IFPRI–ESSP ADDIS ABABA
P.O. Box 5689, Addis Ababa, Ethiopia
T: +251.11.617.2000 | F: +251.11.646.2318
Email: ifpri-essp@cgiar.org | http://essp.ifpri.info

The Ethiopia Strategy Support Program (ESSP) is managed by the International Food Policy Research Institute (IFPRI) and is financially supported by the United States Agency for International Development (USAID), the Department for International Development (DFID) of the government of the United Kingdom, and the European Union. The research presented here was conducted as part of the CGIAR Research Program on Policies, Institutions, and Markets (PIM), which is led by IFPRI. This publication has been prepared as an output of ESSP and has not been independently peer reviewed. Any opinions expressed here belong to the author(s) and do not necessarily reflect those of IFPRI, the Ethiopian Development Research Institute, USAID, DFID, the European Union, PIM, or CGIAR.