

Modeling Impacts of Drought on all Major Crops in Ethiopia: 2010-2015

Michael L Mann & James Warner

September 29, 2017

Abstract

In this study, although we do not find evidence of extensive drought, we provide evidence of localized clusters of drought in the 2015 growing season. Despite this we find scant evidence of substantial crop losses in the fifteen most productive zones, but instead that they were able to maintain or exceed production levels seen in 2014. That being said, there is a substantial uptick in the number of drought related losses reported by farmers, within localized areas, particularly in 2015. To explore the impact of these local droughts and management techniques on output per hectare, we develop a panel regression utilizing sub-kebele level agricultural survey data, and remotely sensed data on greenness, precipitation and measures of evaporative demand. We create a set of models to provide mid-season estimates of output per hectare for maize, wheat and barley, and obtain R^2 between 0.39 and 0.35 at the sub-kebele level. We also find crop specific responses to a variety of management practices, and climate conditions with R^2 between 0.47 and 0.31 at the sub-kebele level. We believe these models could be used to help monitor and predict yields for disaster response teams and policy makers to use, particularly with further development and integration of the newly available 30m Harmonized Landsat Sentinel (HLS) dataset. We further utilize our agricultural survey and remotely sensed data to develop machine learning models aimed at predicting substantial crop losses (>25

1 Introduction

The ability to monitor and predict crop yields in developing countries is critical to the successful adaptation to changes in our climate. Increased temperatures and variability has already been linked to losses in maize and wheat yields (-3.8 and 5.5% respectively) and crop prices globally (Lobell, Schlenker, and Costa-Roberts 2011). Although much effort has been placed on modeling the spatial distribution of these shifts, less effort has been placed on how yields vary across space and time especially in a longitudinal setting (Ray et al. 2015). Advances in remote sensing provide avenues to monitor agricultural crop health at high spatial and temporal resolution. However, our ability to monitor changes in plant productivity is still limited in the more complex environments common to many developing countries (Mann and Warner 2015).

Remote sensing based efforts to characterize the extent, cultivation practices, and productivity of global 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. Remotely sensed imagery has also been employed to map irrigated areas (Thenkabail et al. 2009, Portmann, Siebert, and Döll (2010)), and cropping frequency/intensity (Biradar and Xiao 2011, Gray et al. (2014), Li et al. (2014)).

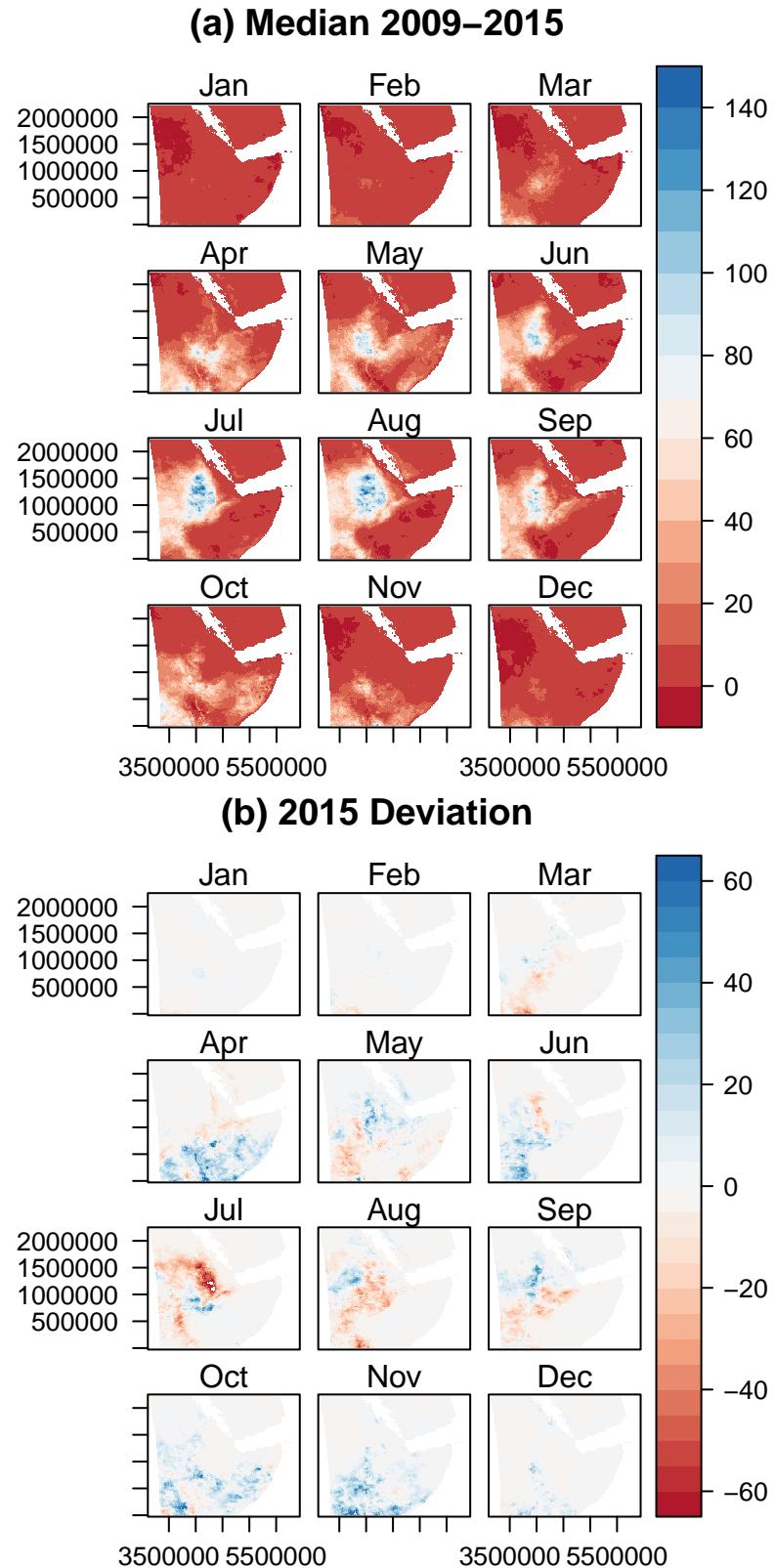
Initial efforts (e.g. LACIE and AgRISTARS) primarily utilized remotely sensed imagery to characterize the spatial extent and growth stage of crops, but relied on models driven chiefly by meteorological information to predict crop yield (Idso, Jackson, and Reginato 1977, Doraiswamy et al. (2003)). However, the biophysical link between canopy spectral reflectance and net primary production has long been established (Tucker and Sellers 1986); indicating that satellite measurements could play a role in determining crop yield directly. A wide number of studies have documented highly explanatory empirical relationships between satellite measures such as NDVI (in many forms: growing season maximum and mean, seasonally integrated, etc.) 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 particularly 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). More recently, methods for forecasting yields with remotely sensed variables at the field scale have been explored (Lobell et al. 2015, Gao et al. (2017)).

The main objectives of this paper is to create a suite of models that can a) provides reliable estimates of yields based on remotely sensed data, b) reveal insights into the impacts of management. We also create a suite of algorithms used to extract, summarize, and organize remotely sensed data and prepare it for spatiotemporal analysis.

1.1 The State of Drought

Spurred 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. Reports of below average rainfalls in 2015-17 across the low lying pastoral communities and South Eastern Ethiopia triggered an international effort to provide food aid to roughly 8.5 million people in the summer of 2017 (“Ethiopia: Drought - 2015-2017” 2017, (“Mid-Year Review, Ethiopia Humanitarian Requirements Document, July 2017” 2017)). The understated response from the Ethiopian government has lead to accusations that the government is downplaying the severity of the drought as early as 2016 (Schemm 2016, Schemm (2017)). Despite the reality of a severe and crippling drought in food insecure low lying areas, there is little sign of food shortages nationally, with a recent study observing no significant change in grain prices throughout the country (Bachewe, Yimer, and Minten 2016). The lack of a price response would indicate that while food insecure areas were hit heavily by water shortages largely effecting livestock, while leaving the highland areas largely unaffected or triggering only minor losses. Despite international concern, conditions up until the 2016-2017 growing season may have been largely handled by domestic resources and trade flows, therefore requiring shifts in distribution rather than a large-scale international intervention.

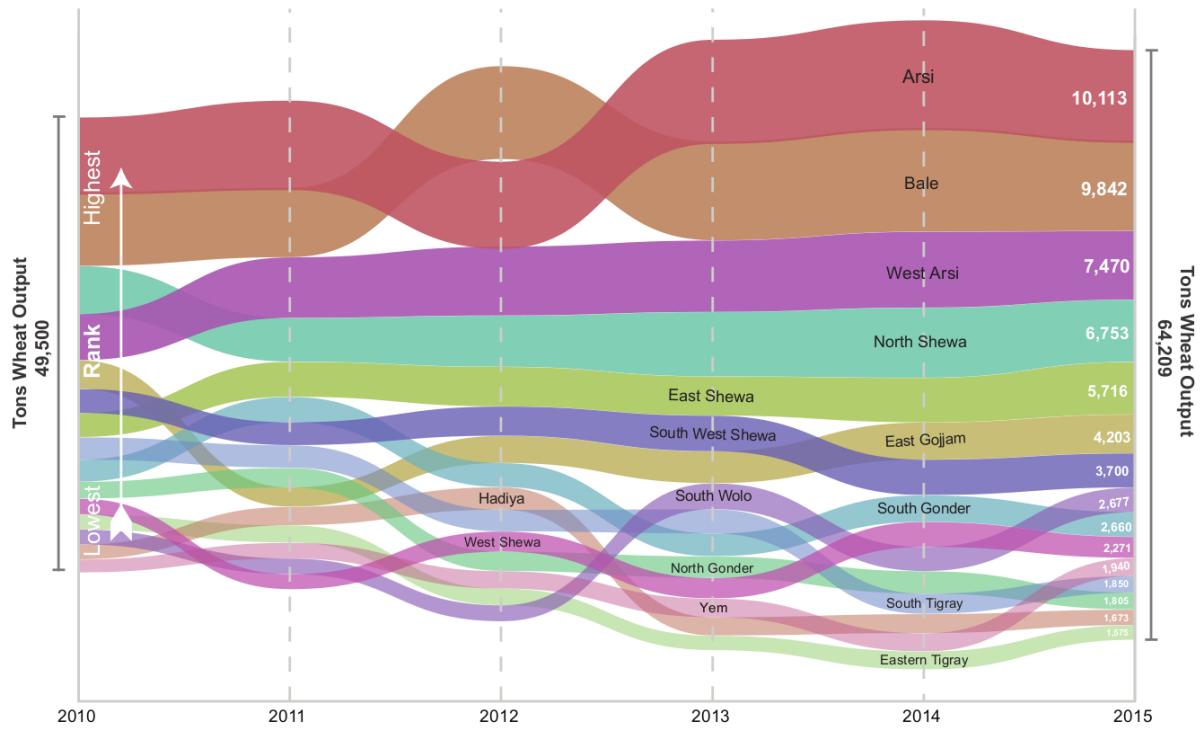
Figure 1: Median rainfall per 10 day period 2010-2015 compared to 2015 actual



Looking at Figure 1 we can see the spatiotemporal patterns of rainfall across Ethiopia. The top panel (a) reports the median total rain fall per 10 day period during each month (in mm). Here we can see the onset

of the heavy rains during the Meher growing season, and the decline near harvest in October. In panels (b-c) we can see the deviation from this norm for the years 2015 and 2016 (in mm). The only striking shortfall is July in 2015 and patchy isolated declines through August that year. 2016 by comparison shows indications of generally above average rainfall. These figures also demonstrate the spatial heterogeneity of rainfall over time. Both years show small areas that interchange between having above or below average rainfall. It should be pointed out that the erratic nature of these rains might point to the importance of the timing of these rains, with rainfall being more important during critical growth stages like germination and flowering. If rains fail during critical periods, even small shortfalls can precipitate significant losses. Therefore Ethiopia, like many rainfall dependent countries has and will continue to be prone to localized and occasionally wide-spread yield losses.

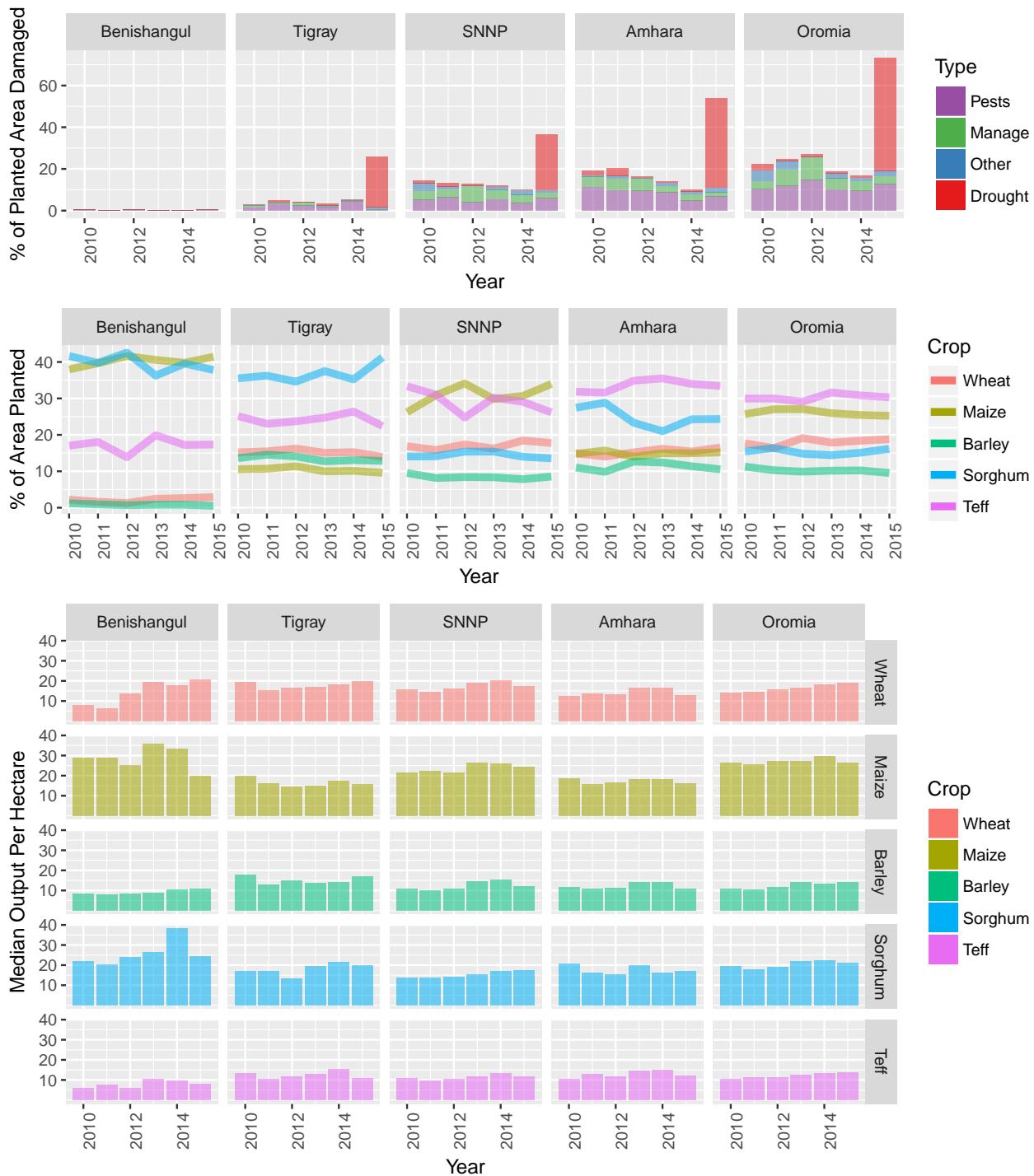
Figure 2: Total wheat output of top 15 producing zones in tons



Looking at the example of wheat production for the top 15 producing zones (Figure: 2), we can see that total output has remained steady or increased in the majority of top performing zones. This can be seen by the increase in total production from 49,500 in 2010 to 64,209 in the 2015 growing season. Additionally, we can see that the total production of each zone (as indicated by the width of each bar) generally increases over time, or remains relatively constant. Moreover we can see that the rank of regions (as indicated by the plot order, bottom being lowest rank, top being highest rank) remains surprisingly steady despite the exogenous shock of drought beginning in the 2015 planting season. We do however see a retraction in total output in 2015 due to the drought, however this only seemed to revert total yields to 2014 levels.

We can see from Figure 3 three time series for the major regions depicting the drought, in the top panel (a) the percentage of planted area damaged by damage type, the middle panel (b) the percentage of area planted in each major crop, in the bottom panel (c) median output per hectare (OPH) in quintiles for each major crop. Although we see a marked spike in reports of drought damage in the 2015 planting season (covering harvest in 2016), we see a few important responses. First, we see that although slight, farmers increased the percentage of area planted in drought tolerant crops such as maize and sorghum. Second, we see that although some regions saw declines in median output per hectare, these losses are largely offset by gains in other crops or by the steady increase in OPH since 2010.

Figure 3: Drought in one graph 2010-2015



Pulling from later in the study (see “Predicting Drought Damages”), we can see that although large-scale drought may have been elusive in 2015 for major growing regions, localized drought was not. Ethiopia has an extraordinary topography, with jagged mountains, rugged hillsides, and narrow canyons all being graded and tilled for agriculture. This complex terrain in turn leads to complex micro-climates in many ways much like those found in Eastern California. Rain shadows, elevation, strong inversion layers, monsoonal rain, complex wind patterns all lead to an extremely heterogenous growing environment, where conditions on one side of the mountain could be completely different than 1/2 a mile on the other.

2 Methods

2.1 Data Sources

2.1.1 Survey Data - Agricultural Sample Survey Data (2010-2015)

Survey data was obtained from Ethiopia’s Central Statistical Agency’s (CSA) Agricultural Sample Survey (AgSS) and was chosen for its annual collection, spatial coverage, and unique sampling structure through these six crop years. The AgSS is an annually collected government administered large-scale survey tasked with measuring agricultural production in Ethiopia at the zonal level.¹ The survey interviews approximately 45,000 farmers per year on a range of farm management questions covering some basic demographics of the household as well as a range of questions concerning planting, harvesting and selling at the plot level. Typically, about 20 farm households are randomly sampled from small local village-level areas of approximately 200 households (the sub-kebele level). On average, based on population, approximately five of these sample areas exist at the kebele-level administration area. Although the figures vary widely, sub-kebeles are meant to approximate about 200 households per unit and the average kebele has around 1,000 households each. From this sampling frame, a random selection of about 2,200 sub-kebeles are chosen as a representative sub-sample for zonal level production projections of major agricultural production areas. Population weights are then applied to project agricultural production at the zonal level. For this study, we construct a panel data set over the six crop seasons identified because approximately 75% of the same sub-kebeles were sampled by CSA over the six meher crop seasons of interest. This effectively builds a base of five relatively favorable crop seasons, from 2010 to 2014, and allows for the drought effects of 2015 as a study in contrast in terms of this significant weather event. The sample is with replacement and therefore representative of the sub-kebele but not a true panel because the same farmers were not chosen. In addition, no sampling weights are applied because we have omitted approximately 25% of the sample and are generally most interested in the localized effects.

While this study collected data on the five principal Ethiopian field crops (teff, wheat, maize, barley, sorghum), because of their predominant importance in value terms for Ethiopia as well as their geographically wide-spread adoption, the analysis should therefore be placed within this context. This study covers the drought effects of only these five crops and not all agricultural production. In other words, when we outline percentages of crop losses we are referring to losses of only these five crops and not total output in the country. We have no reason to believe that these are not generally representative of all crop production, in terms of the drought, but a more detailed analysis would be needed to make this assertion.

The principal unit of analysis is the sub-kebele and all relevant CSA data is aggregated to this level. The survey data represents an amalgamation of all 20 households as a single representative farmer, we refer to as a “super-farmer.” This was done for a variety of reasons including the fragmented plot farming system common in Ethiopia as well as CSA data collection methodology. More specifically, CSA data collection methodology relies on crop cuts to estimate productivity at the local level. Depending on the actual number of farmers growing the specific crop, CSA collects up to five different individual farmer crop cuts, averages the yields, and projects this figure onto all plot areas for that crop in the sub-kebele. In this way, individual

¹Beyond the nation, Ethiopia has four official levels of administrative areas. These include, in order of geographic size, regions, zones, woredas and kebeles. There are approximately 12 regions, 88 zones, 690 woredas and 15,000 kebeles. For our purposes, we use the sub-kebele, not recognized as a geographic region for administrative purposes, but commonly used for statistical sampling and are roughly based on population. There are over 75,000 sub-kebeles in the country.

plot level yields are not able to be determined and, consequently, household level estimates are not possible. Additionally, while the level of the pixel used for the remote sensing data is relatively granular ($6,250 \text{ m}^2$), cost and logistic considerations make plot level estimations prohibitive in this research. Therefore, the data from 20 farmers is aggregated to a single super-farmer resulting in 1,780 observations sampled from the sub-kebele area (approximately $23,900,000 \text{ m}^2$) over a six-year period. The total number of observations collected is about 10,680 but will vary in sample size depending on the crop choice (i.e.. some areas do not grow all five designated crops). A breakdown of each super-farmer is done with aggregated non-crop related information combined with the specificity of the individual crop chosen. More precisely, several variables are consistent across all five crop choices (elevation, proportional numbers of male and female household heads, total land area, etc.) and others are crop specific (use of fertilizer, seed, etc.). Each crop year was checked for consistency, cleaned and aggregated to the sub-kebele level. The six years were then aggregated and merged with other data to develop the panel data set. Importantly, for our analysis, we chose specific variables to reflect farmer management decisions and other actions affecting yields. For data cleaning, we Winsorize yield estimates capping them at the 90th percentile, and we also drop any sub-kebele unit that has less than 4 years of data from regressions.

2.1.2 Remotely Sensed Data - Greeness, Precipitation and Evaporation

2.1.2.1 Normalized Difference Vegetation Index

Considering the relatively small scale of agriculture in this region we use 250m vegetation products from the MODIS satellites. Vegetation indices are obtained from 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 8-day periods (Doraiswamy, Stern, and Akhmedov 2007). We find that the combination of these two products provides a stable and informative time series.

For the sake of simplicity and replicability, we focus on the Normalized Difference Vegetation Index (NDVI) and examine its predictive power using panel econometric techniques. NDVI is sensitive to the amount of chlorophyll in any given pixel and is commonly used to estimate plant productivity and health in agricultural applications (Mann and Warner 2017,Mann and Warner (2015),Funk and Budde (2009)). After removal of snow, cloud and other flagged low quality cells, we remove all non-agricultural cells through the use of the 500m MODIS land cover product (MCD12Q1) for the appropriate year.

2.1.2.2 Water Availability Variables

In order to estimate the effects of precipitation (PPT) on our crop production models the Climate Hazards Group Infrared Precipitation Station (CHIRPS) data was included. Data is collected as total precipitation by dekad (Funk et al. 2015). A dekad represents three periods each month, the first two periods being 10 days each, and the third being any remaining days of that month.

Hydrological availability of water and available energy for plant growth have been shown in previous research to be important factors in crop models and are therefore included here. For this reason, we include monthly estimates of potential evaporation (PET) and actual evapotranspiration anomaly (ETA). Actual evapotranspiration (AET) is the sum of transpiration of water through plants plus evaporation from soils and water surfaces. AET is a correlated with vascular plant productivity and correlates with the biomass accumulation and regrowth(Major 1967,Mann et al. (2016))#nowcastresults. The ETA variable used in this study is the current AET compared to the 2003-2013 mean. ETA therefore is a proxy for drought as low current values of AET would correspond to reductions either in available energy to move water, or water itself. Potential evapotranspiration 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 light and heat, but also includes the impact of wind speed, pressure, and humidity amongst others. The standard deviation of PET, reflects variability at the local level of these weather conditions. Both ETA and PET are available for download through FEWS NET (2017).

2.1.3 Variable Definitions

Table 1: AgSS variable names & 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
Postfixes and Other	
_diff	First difference
lag(X,1)	One year lag of variable X
AgSS Variables	
OPH_W	Winsorized (95th) values of output per hectare (quintals)
IMSEED	Quantity of improved seeds (kg)
IMSEED_P	Percentage of planted area planted with improved seeds
FERT_CHEMICAL_AREA_P	Percentage of area treated with chemical fertilizers
Fert_Amt_Per_Area	Quantity of chemical fertilizer used per unit area
DAMAGE_AREA_P	Percentage of area reported as having ‘damage’ (all causes)
DAMAGE_DROUGHT_AREA_P	Percentage of area reported as having ‘damage’, (drought)
EXTAREA	Area reported as being covered by extension services
Other Geographic Variables	
dist_pp50k	Distance from nearest populated place with greater than 50k persons (km)
dist_rcap	Distance from nearest regional capital (km)
elevation	Elevation (m)

Table 2: Remotely sensed variable names & descriptions

Name	Description
Prefixes	
G_	Calculated from Meher growing season values
A_	Calculated from all calendar year values
ETA_	Actual evapotranspiration
PET_	Potential evapotranspiration
PPT_	Precipitation
Postfixes and Other	
_diff	First difference
lag(X,1)	One year lag of variable X
NDVI Basic Summary Statistics	
G(A)_mn	Meher (Annual) growing season mean values of NDVI
G(A)_min	Meher (Annual) growing season min values of NDVI
G(A)_max	Meher (Annual) maximum values of NDVI
G(A)_height	Meher (Annual) difference between G(A)_max and G(A)_min
G(A)_sd	Meher (Annual) standard deviation in values of NDVI
G_plant_date	Estimated date of Meher planting based on NDVI phenology
G_harvest_date	Estimated date of Meher harvest based on NDVI phenology
season_length	Difference between G_plant_date and G_harvest_date
Other Basic Summary Statistics	

Name	Description
PET(ETA,PPT)_G_mn	Meher growing season mean values of PET (or ETA, PPT) values for the Meher growing season defined by G_plant_date and G_harvest_date
PET(ETA,PPT)_G_sd	Meher growing season sd of values for PET (or ETA, PPT) values for the Meher growing season defined by G_plant_date and G_harvest_date
Integrated Summary Statistics	
G_AUC	Rabi(Kharif) area under the curve of NDVI
G_AUC_v2	Rabi(Kharif) area under the curve of NDVI estimated by splines
G_AUC_leading	Area under the curve of NDVI for the ascending part of the curve during the Rabi(Kharif) season
G_AUC_trailing	Area under the curve of NDVI for the decreasing part of the curve during the Rabi(Kharif) season
Comparison to Norms	
G(PET,PPT)_AUC_diff_mn	Difference between current mean Meher NDVI (or PET, PET) and mean of historic values
G(PET,PPT)_AUC_diff_90th	Difference between current 90th percentile Meher NDVI (or PET, PET) and 90th percentile of historic values

2.2 Spatial Considerations of the Drought

One issue of consequence in drought conditions, such as that in 2015, is to determine to what degree the drought is extensive or limited to a small number of areas. To aid in this process we utilize measures of spatial autocorrelation, one global and another local. Global measures of spatial autocorrelation evaluate whether or not observations of drought impacts are clustered (similar values are near each other), or whether they are spatially dispersed (dissimilar values are closer to each other) for the entire dataset. Local measures of spatial autocorrelation can pick up on local concentrations of high (“Hotspots”) or low values (“Coldspots”).

2.2.1 Global Spatial Autocorrelation

Global Moran’s I depicts relative spatial autocorrelation for the entire area, where positive values convey clustering and negative values convey dispersion, and can be mathematically depicted by formula 1: below.

(1)

$$I = \frac{n \sum_{i=1}^N \sum_{j=1}^N w_{ij} z_i z_j}{S_0 \sum_{j=1}^N z_i^2}$$

where z_i is the difference between a location i and its mean ($x_i - \bar{X}$), w_{ij} is the spatial weight for feature i and its neighbors j , n is the number of features, and S_0 is the aggregate of all spatial weights in formula 2: below:

(2)

$$S_0 = \sum_{i=1}^N \sum_{j=1}^N w_{ij}$$

Here we use neighborhood definition of a 25-km radius with a low decay rate of 0.05 power. This decay function suggests that about 85% of the total value of a neighbor 25 kms away.

2.2.2 Hotspot Analysis

A typical method for identifying local spatial clustering is the Getis-Ord analysis that identifies spatial clustering of values that exist at the statistical tails of a distribution. For our research, we identified “hot” neighborhoods that exist in the top 5% of a normal distributed of damage value and “cold” neighborhoods that exist in the bottom 5%. In other words, does a group neighborhood value statistically deviate from the global mean? In our case we defined a neighborhood as a 25-km radius from the centroid of the sub-kebele with a slight decay rate of 0.05 power².

Using Getis-Ord generates a Z-score and P-value for each sub-kebele in their respective neighborhood. Equation (3) demonstrates the spatial clustering with a 25-kilometer neighborhood and a small distance decay parameter as a weight. A high Z score and small P value for a feature indicates a significant hot spot. A low negative Z score and small P value indicates a significant cold spot. The higher (or lower) the Z score, the more intense the clustering. A Z score near zero means no spatial clustering. See formula 3 below.

(3)

$$G_i^* = \frac{\sum_{j=1}^N w_{ij}(d)x_j}{\sum_{j=1}^N x_j}$$

$$(d) = \begin{cases} 1, & \text{if } d_{ij} < d \text{ for all } i,j \\ 0, & \text{otherwise} \end{cases}$$

2.3 Estimating Crop Yields

2.3.1 Preprocessing

2.3.1.1 Data Aggregation

Because historical crop yields are only available at the level of certain political units (EAs in this case), pixel-level NDVI data must be aggregated at this same level. In this study, we calculate the mean across the raw NDVI values for all available agricultural pixels at the EA-level and further extract statistics of interest from this resulting time-series (at a frequency of 8-days).

2.3.1.2 Summarizing Temporal Data

One of the primary challenges in utilizing high-frequency time series (8-day NDVI) to estimate low-frequency events (seasonal yields) is in reconciling this temporal mismatch when formalizing the relationship between sources of data. In general, low-frequency properties must be extracted from the high-frequency time series that may be relevant to characterize and identify important aspects of plant phenology across the growing season that affect crop yields. In this paper we aim to capture relevant phenological features of wheat through 41 metrics summarizing 8-day NDVI data. These metrics cover Meher growing season statistics, spanning the estimated planting date until harvest date. These plant and harvest dates are estimated on a EA by EA basis, therefore differences in timing due to elevation, crop choice or other management considerations should be captured.

We estimate three classes of statistics across each of these two distinct periods: summary statistics (e.g. mean, max, variance); integrated summary statistics (e.g. area under the curve for the first half of the growing season); and deviation from the norm statistics (e.g. deviations of a given statistic from its 90th historical percentile).

²The issue of what defines a neighbor is somewhat controversial in the spatial literature and the choice has important impacts on the analysis. For our purposes, 25 kilometers seemed reasonable from both expert opinion and fit the actual distribution of the data well.

2.3.2 “Now Casting” - Estimation Strategy

One primary use for these data is for early season estimation of crop yields. Although AgSS provides invaluable estimates of agricultural productivity on an annual basis, lag between data collection and release is substantial. This gap leaves substantial need to timely first-order estimates for crop yields. Motivated by preliminary panel regression results reported in the Appendix, we find that reasonable estimates of output per hectare can be created around mid-season. Where ‘mid-season’ is defined by reaching the maximum NDVI greenness value.

2.3.2.1 Panel Regression for “Now-Casting”

Compared to cross-sectional approaches, panel analysis can substantially increase the degree of observed variance over both space and time. For instance, in our current application, the integration of EAS (with $n = 1752$) over the 2010-2015 sample period ($t = 6$), results in a substantially larger data ($N \sim 6900$), allowing for a much greater flexibility and degrees of freedom in the modelling of the issue of interest. Here we include a series of variables that can be calculated in the first half of the growing season including contemporaneous and lagged values, Zonal fixed effects, as well as a lagged value of output per hectare. Here we estimate a series of random effect models of the following form 4:

(4)

$$Y_{itc} = \alpha_i + \sum_{k=1}^K \beta_k X_{it} + \mu_{it} + \epsilon_{it}$$

where Y_{itc} is the output per hectare (OPH) for crop c for sub-kebele i for year t , α is the average output per hectare for sample population (all major growing regions), $\sum_{k=1}^K \beta_k X_{it}$ is the sum of all K coefficient estimates β_k for all variables X , where X comprises all variables that are available a mid growing season, μ_i is the sub-kebele specific random effect that measures the difference between the average OPH at sub-kebele i and the average OPH in all major growing regions. The term ϵ_{it} is the individual-specific effect, which can be described as the deviation of the i th OPH from the average for that year.

We present preliminary “Now-Casting” results from panel regressions on output per hectare for Ethiopia’s major crops [here](#).

2.4 Predicting Crop Losses from Drought

Another critical emphasis of this project is to see if remotely sensed data on greenness, precipitation, or evapotranspiration and the zone code can be used in conjunction with household surveys to predict the location and/or extent of damages before the next years AgSS survey results are released. In this vein we aim to utilize advances from machine learning to accurately predict substantial crop damages for all major crops. We define our dependent variable, ‘substantial’ losses due to drought, as farmer reported losses over 25%. Two types of models are then run for each crop, one, a “national-level” model is trained on all observations for the major growing regions, and a second group of “region-level” models where each model is trained from data in a specific region. Specifically our model takes the following form:

(5)

$$Y_{cn} = f(X)$$

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

Where Y_{ci} is a binary variable where losses for crop c due to drought in sub-kebele i are greater than 25%, and it is estimated a non-linear function of all remotely sensed variables of interest X . We present results for damage predictions [here](#)

2.4.1 Estimation Strategy

2.4.1.1 Random Forests Description

Random forests (RF) are a flexible (used for classification and regression) 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 method where by each binary split (e.g. NDVI < 0.1) is identified as the most informative, as defined by the greatest reduction in Gini. 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 perform consistently out-of-sample and with noisy data sets.

2.4.1.2 Training vs Testing Data

Before we run any models we break the data into two groups, with 85% of all sub-kebele units being retained for the training and tuning of the models, and 15% of these held aside to provide independent measures of out-of-sample performance.

2.4.1.3 Training Data: Cross Validation & Tuning

To enhance the performance of the models out-of-sample, we undertake two procedures on the training data; tuning and cross validation. Tuning in this case is limited to choosing the number of features (variables) randomly selected for use in each tree. We do a grid search independently for each crop finding the parameter tuning that provides the best performance. To ensure that these parameter tunings work well, we utilize k-fold cross validation whereby the training data is split further into 3 sets of sub-kebele data, the model is trained on two thirds of the data and its performance, using the Kappa coefficient, is estimated using the omitted third sub-sample. Performance for each parameter value is therefore evaluated out-of-sample three times and averaged. The optimal parameter then is that with the highest Kappa value and is specific for each crop model. Kappa[[^]kappa] is a performance indicator similar to the percentage of observations that are correctly classified but that controls for the fact that categories with large numbers of observations are easier to guess by random chance. [[^]kappa]: see: https://en.m.wikipedia.org/wiki/Cohen%27s_kappa

2.4.1.4 Testing Data: Out-Of-Sample Performance

Independent of the training data we initially withheld 15% of sub-kebele units to provide ‘testing data’ for accurate measures of truly out-of-sample performance for our tuned RF models. These tests should therefore be good indicators of performance for sub-kebeles excluded from the AgSS survey. To test accuracy we make predictions of the damage category on this testing data and compare them to each observations observed damage category. We provide three major performance indicators, a) overall accuracy, 2) Kappa coefficient, and 3) the recall rate³. The same metrics are applied to both the national-level (all major growing regions), and the regional-level models (specific to each region).

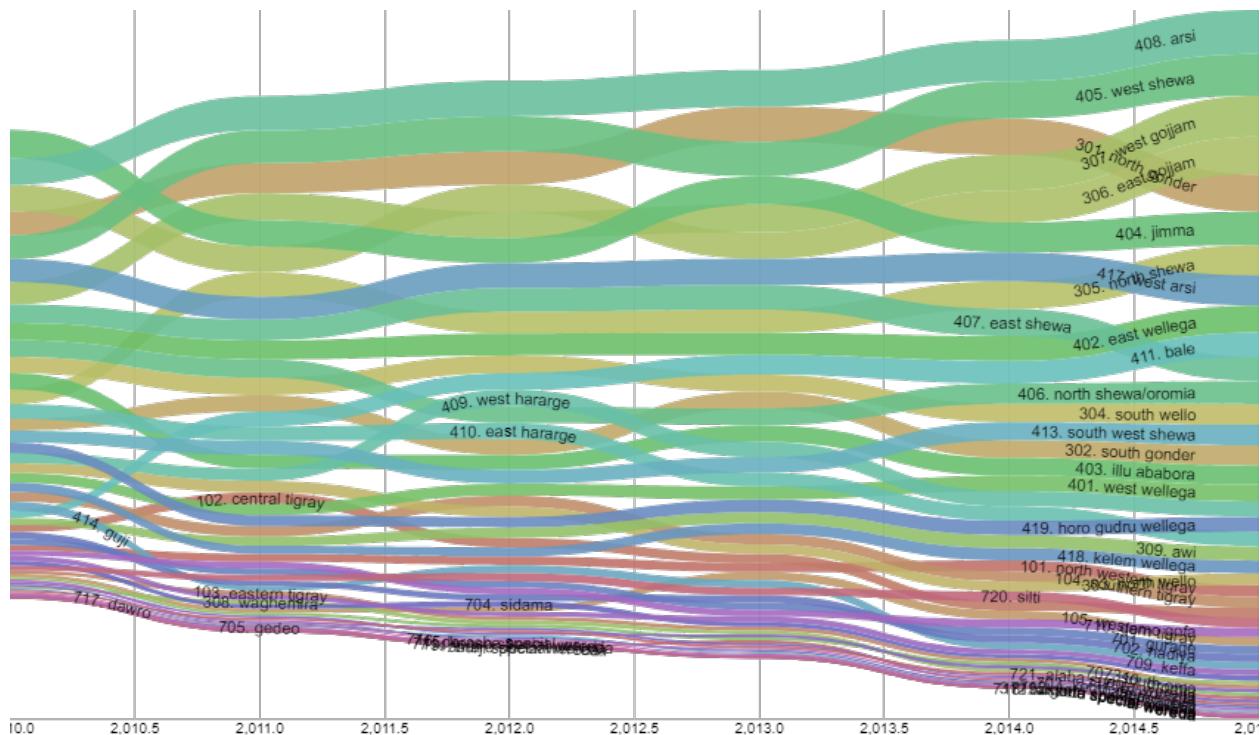
³the percentage of sub-kebeles with ‘substantial drought damage’ that are correctly classified by our models.

3 Results & Discussion

3.1 Spatial Considerations of the 2015 Drought

The 2015 drought had varying spatial impacts on Ethiopia and this section seeks to review how the drought impacted cereal crop production at a variety of spatial scales. We find that impacts generally occurred in local neighborhoods of relatively less productive areas. This creates “pockets” of severely affected areas that has limited impacts on overall national production figures. Figure 5 reveals that only two of the top ten cereal production zones, aggregated by weight, were significantly affected by the drought. More specifically, North Shewa of Amhara and East Shewa of Oromia. Importantly, they are ranked eighth and ninth with this ranking scale. Furthermore, higher impact areas were located in the north and northeast of the country. Of critical importance is the fact that while larger administrative areas had greater impact there was significant variation at the local level. This is important because many designations of drought areas are determined at the woreda or higher administrative area level. With the high variance, this suggests that areas designated as drought areas may, in fact, not be experiencing these negative shocks. We suggest a review of this research with an eye to using smaller level administrative areas such as the kebele for assistance. In addition, time series trends of damage are fairly consistent suggesting longer term policy actions in these higher damage areas.

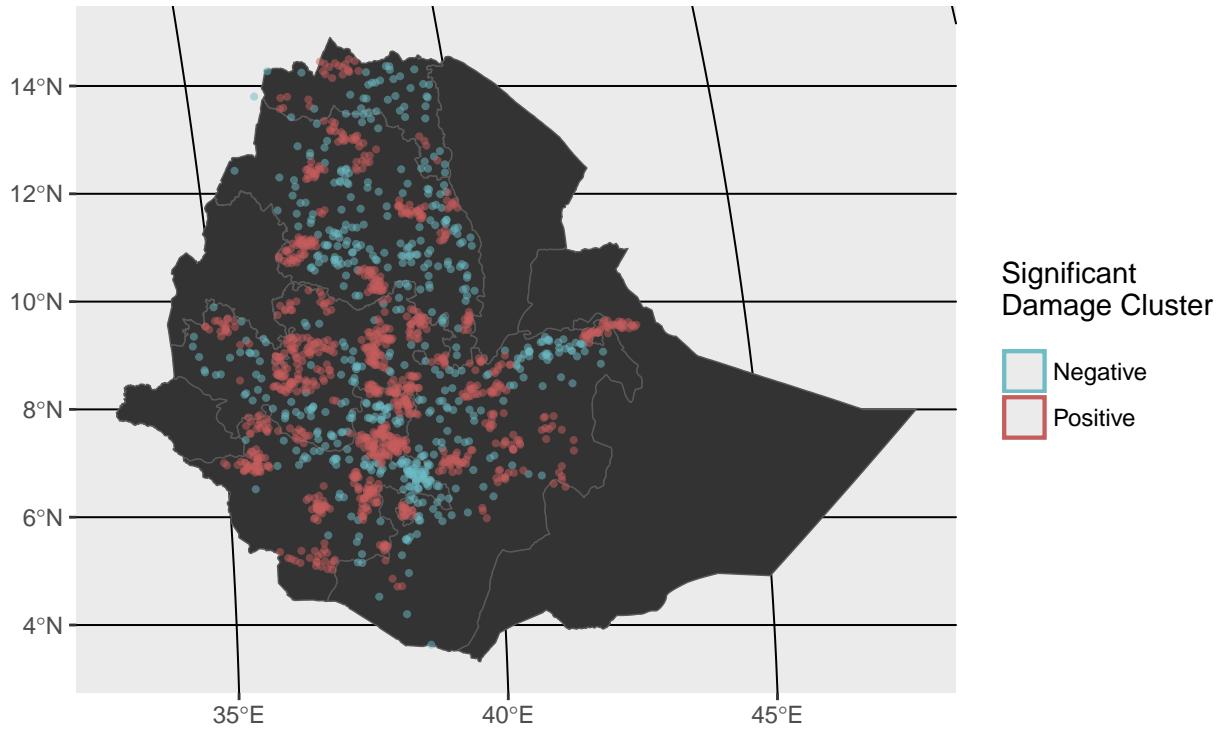
Figure 5: Total Major Cereal Crop Production (by weight 2010-2015)



Aggregated zonal cereal production is presented in Figure 5⁴. The graph demonstrates relative stability of rankings and significant expansion of production for 2010 to 2015. Amhara and Oromia are the important zones, in terms of cereal production, and have the top 20 zones represented. Of course, this has to do with population densities.

Figure 6: 2010-2015 Local clusters for the sum of drought related damages - All major crops

⁴While aggregating the five crops by weight has some important methodological concerns, weight was used to measure relative loss in production rather than value lost. Incorporating a price index would speak to monetary value but the price changes might reflect market movements over drought effects. For this reason, simple weight was used.



To illustrate the localized nature of drought damages in Ethiopia from 2010-2015 we pull from the “Predicting Drought Damages” section of the paper. Figure 6 above presents the results from a local Moran’s I test on the sum of *drought related* damages for all major crops. The “Positive” class indicates a significant spatial cluster of high damage values based on a neighborhood of 25km, and “Negative” class indicates a significant spatial clustering of low damage values.

Table 3: Average Cereal Crop Damage by Year (%)

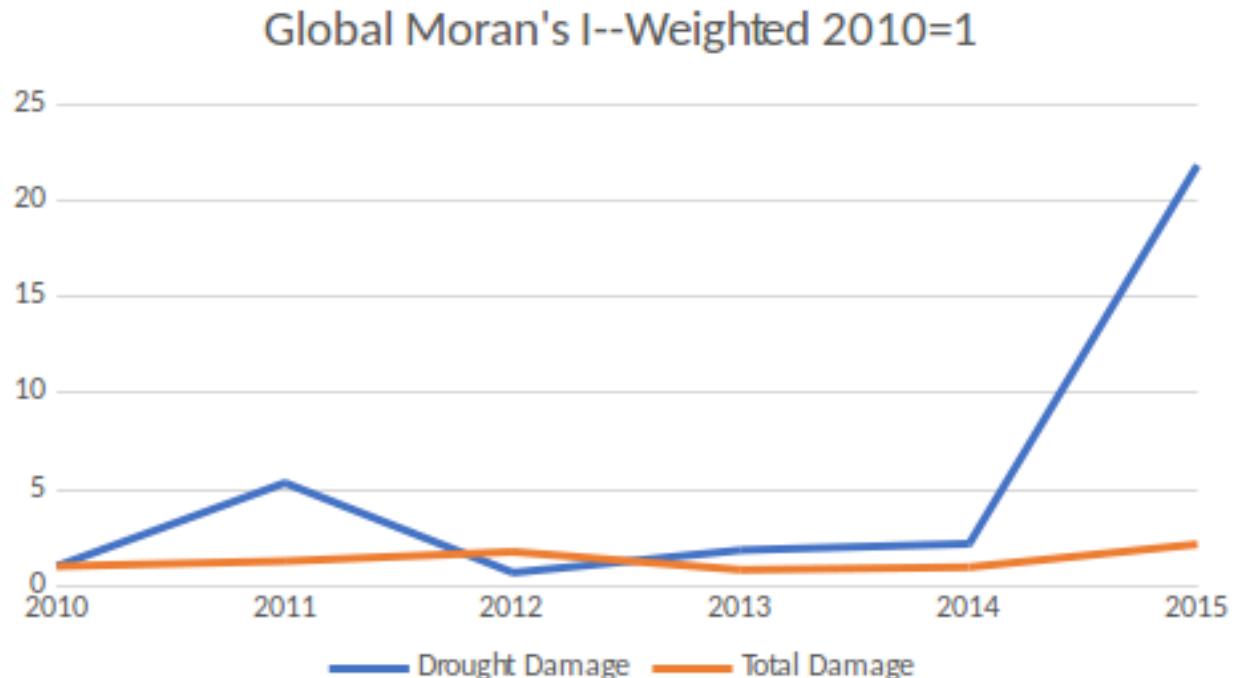
Year	Mean	Standard.Deviation	Minimum	Maximum
2010	11.6	10.9	0	83.2
2011	9.8	9.8	0	76.1
2012	11.6	11.6	0	79.1
2013	8.1	8.1	0	60.5
2014	8.6	8.6	0	94
2015	17.9	17.9	0	95.8

N=1752 for all years

Table 3 depicts average cereal crop loss across the selected sub-kebeles for the six-year period. Overall average rates of cereal crop damage are relatively high for the six years sampled. Typical damage reported was between 8 to 11% per sampled sub-kebele. This suggests that about 10% of all cereal crops are lost due to various damages in a typical crop year. As this is an average across 20 sampled farmers in the sub-kebele this figure does not account for variation within the sample. In other words, individual farmers, or particular plots, may also experience catastrophic losses while others experience minor damage. Policy to address this form of crop loss should be explored and could go a long way to addressing crop loss. As expected, 2015 average crop loss rises dramatically. However, it is important to point out that the 17.9% figure is not evenly applied across equally productive areas and so the total effects on nationally produced cereal crops is different than this figure indicates.

Global Moran's I depicts relative spatial autocorrelation for the entire area, where positive values convey clustering and negative values convey dispersion.

Figure 7: Indexed Global Moran's I—Major Cereal Damage 2010-16



While there is some statistically significant global clustering of damage of major cereal crops it is relatively low and demonstrates no observable trend (See Fig. 7). This is an indication that there are some areas of consistent damage but, as demonstrated later, local measures of spatial clustering provide a more nuanced

view. The AgSS allows for about 20 different types of damage (including different types of weather events (hail, floods drought), pest, birds and bad seed) but the clustering of drought impacted cereal crop damage jumped over 7-fold in 2015. The 2011 increase may reflect a drought outlined in this period. Overall, while there is some statistically significant global clustering, it is the 2015 drought that exhibits relatively dramatic differences.

Figure 8: Drought Effects by Region

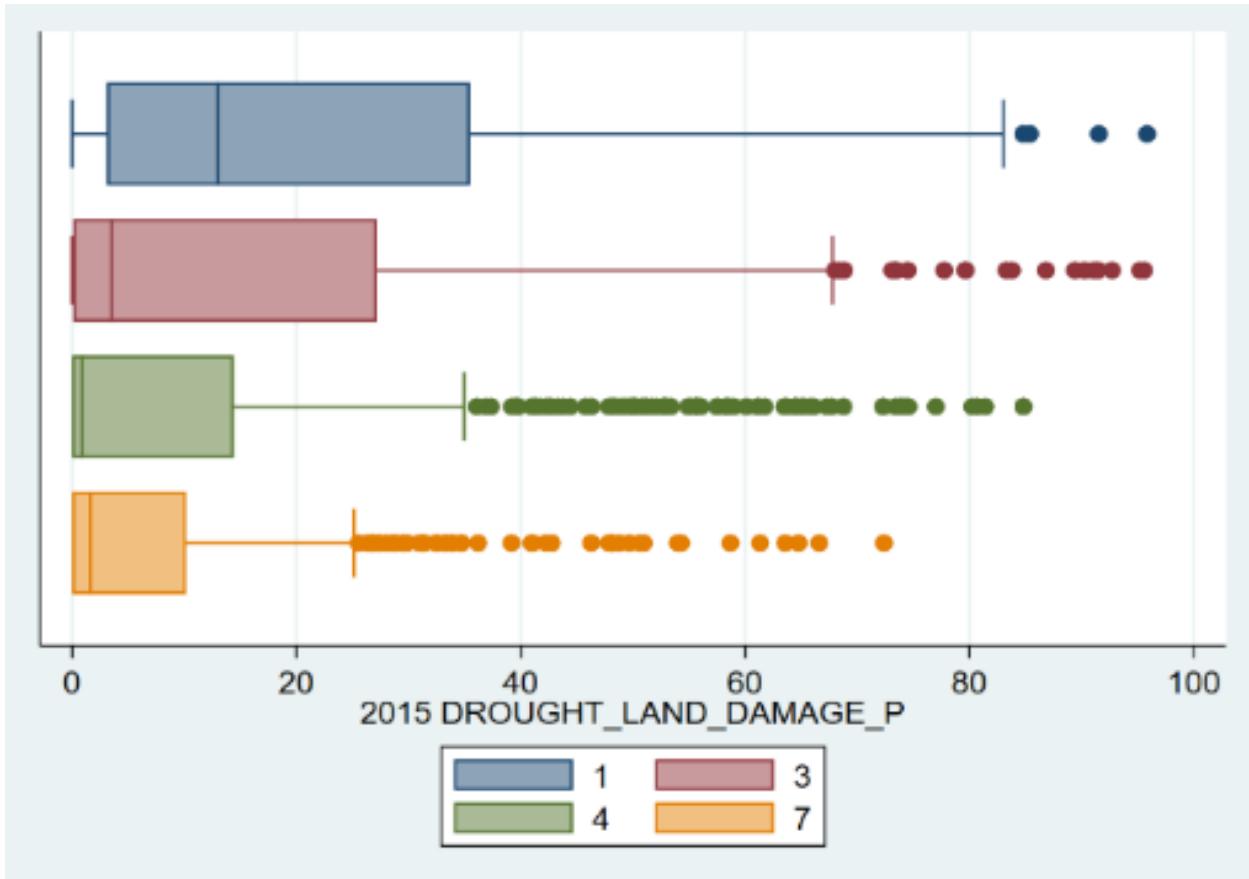
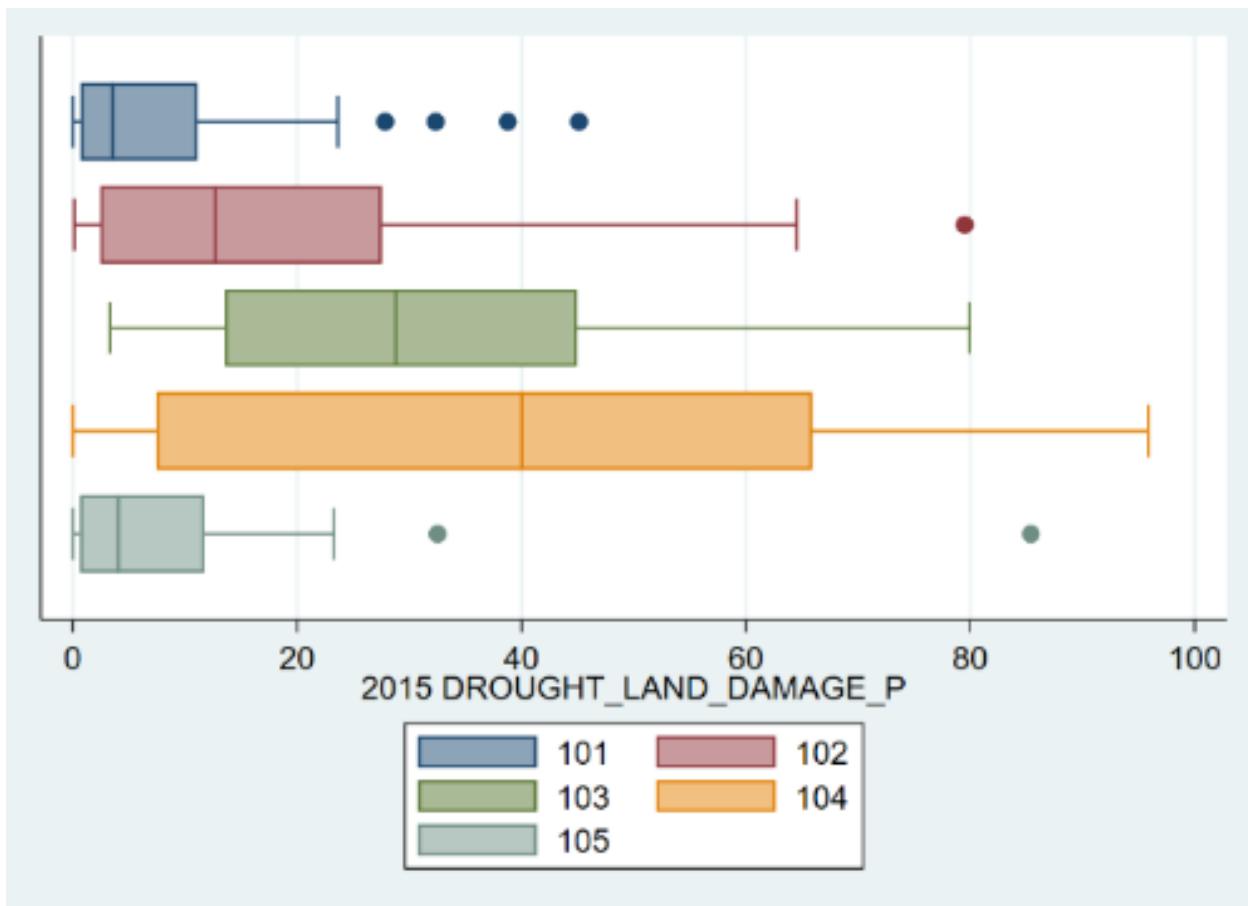


Figure 9: Drought Effects by Zone: Tigray Region



Boxplots are presented of the four major regions in Figure 8. What is evident from the Figure is that Tigray suffered the most overall damage, has the highest median damage and the widest distribution. The 50% quartiles are more compact, indicating a tighter distribution, moving from northern Tigray to southern SNNP. In addition, the median is skewed far left for all regions suggesting non-normal distributions that large number of observations group towards zero with outliers of higher damage much farther to the left. For example, approximately 50% of damage values are less than (2%) in both Oromia and SNNPR. Figure 9 splits Tigray region into smaller zonal drought effects. The five zonal distributions are markedly with median values ranging from 5% to as high as 40%. Both figures suggest wide impacts at both the region and intra-regional level. The following figures drill down into smaller spatial scales and suggest this variation exists at the intra-zonal level as well.

Figure 10: Drought Effects by Woreda: Zone 104

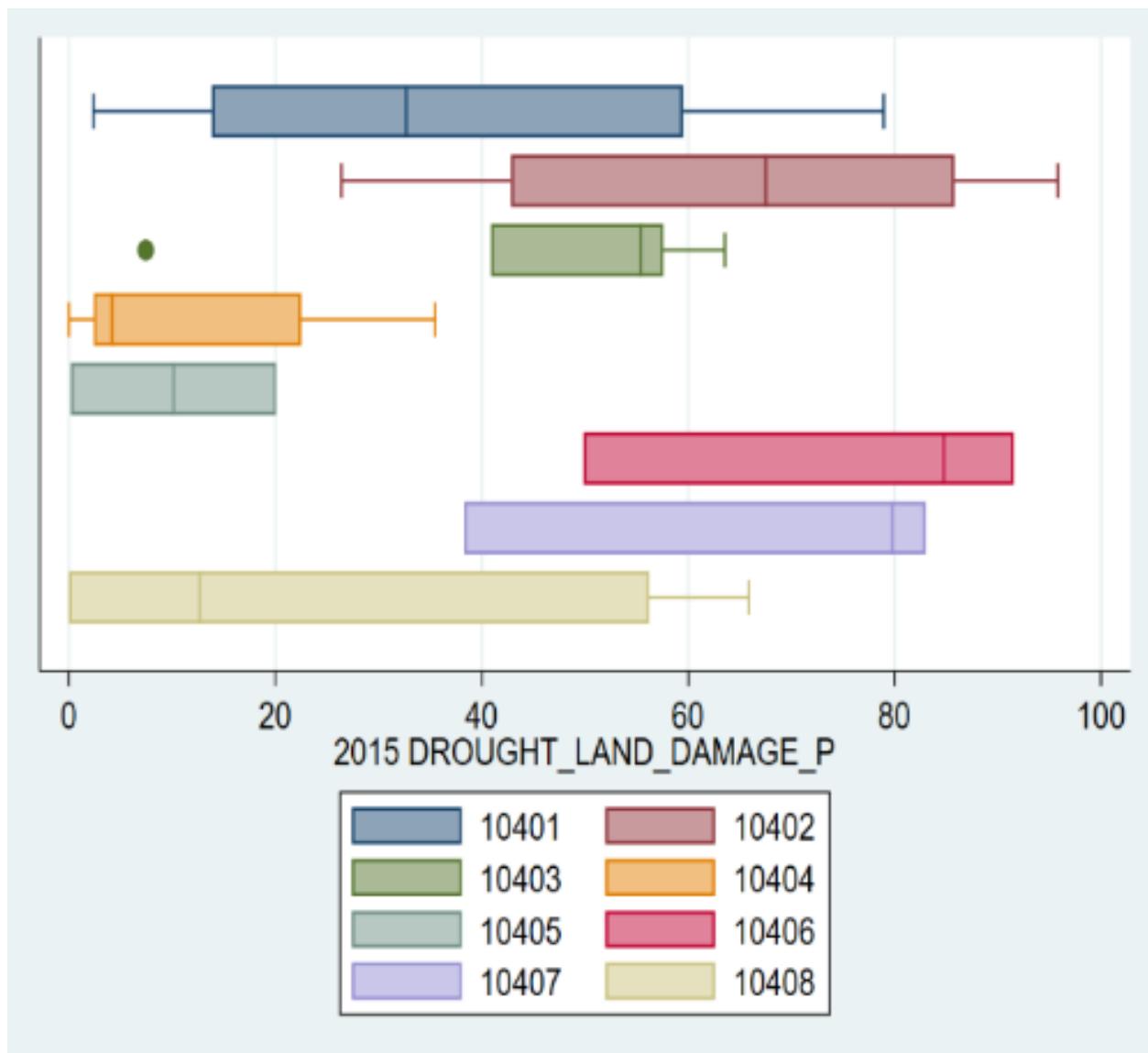


Figure 11: Mao of Drought Effects by Zone: Tigray Region

2015 % Cereal Crop Loss by Drought Southern Tigray Zone (104)

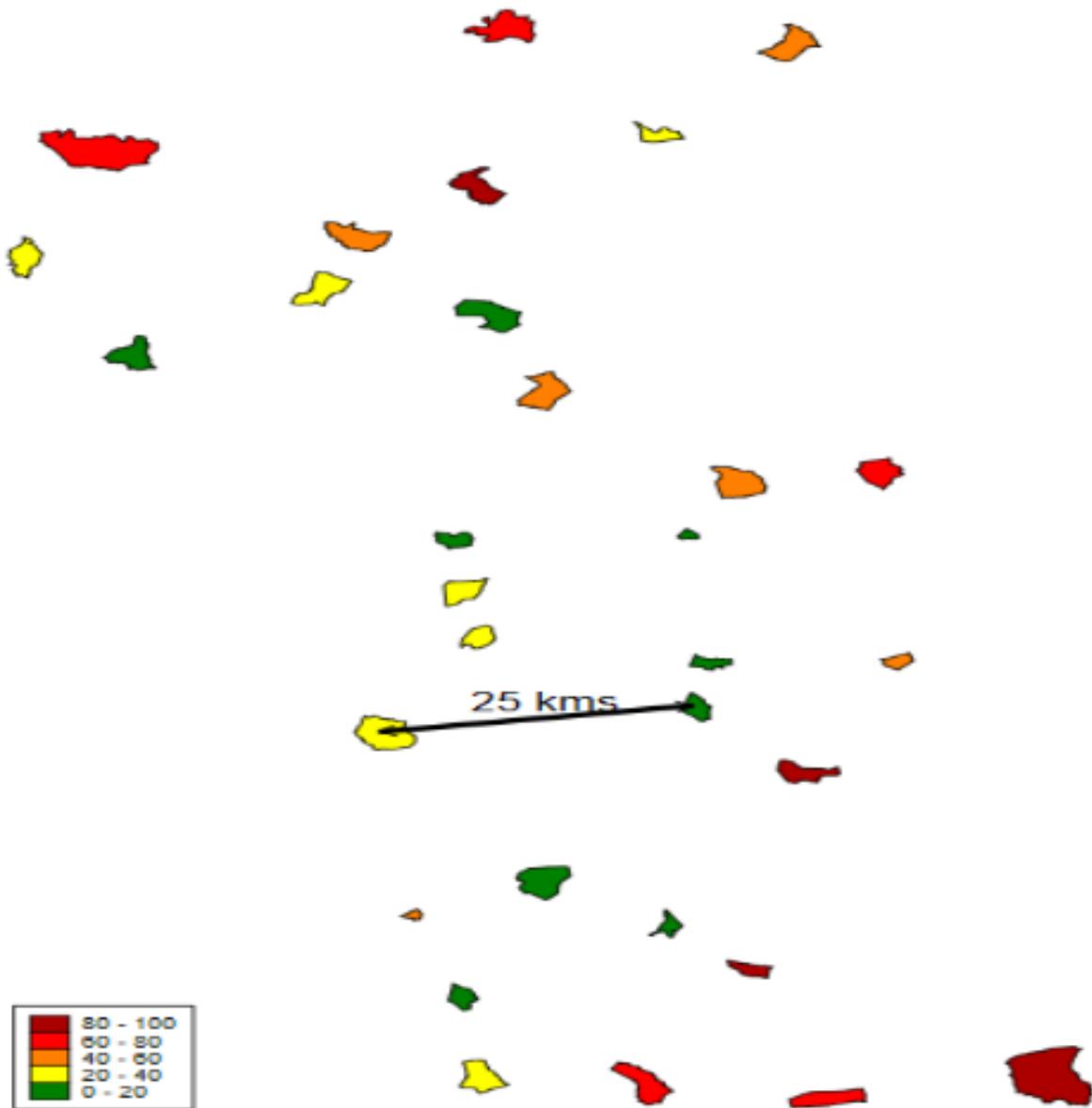


Figure 10 demonstrates that the variability continues down to the woreda level. While the data is not as robust, given that the number of observations are fewer as higher levels of administrative units, values are wide-ranging and indicative of the localized effects of the drought. When placed on a map, the sub-kebeles aggregations, potentially form neighborhoods of similar values that don't necessarily respect administrative boundaries. Broken into 20% categorical damage values, groupings of similar colors seem to present themselves in Figure 11. The 25 kilometer arrow indicates what will later be used as a neighborhood and is placed there for spatial context. The next section attempts to demonstrate group clustering of "hot" (above average) and "cold" (below average) clustering of drought damage and investigates statistical patterns related to the drought.

3.1.1 Hotspot Analysis

In order to determine hot and cold spatial damage time trends, Getis-Ord analysis was performed on total damage per crop year. This analysis identified local neighborhoods that deviated from the mean value of that year. Individual sub-kebeles were assigned a value of plus one, zero, or minus one if they were in the top 5%, middle 90%, or bottom 5% of neighborhood damage values, respectively for each year. These values were summed for each of the six crop years. For example, a value of -2 indicates that for two crop seasons this particular sub-kebele was in statistically significant low damage neighborhoods. A summation of all values is provided in Table 4.

Table 4: Cold/Hot Damage (per area)

Category	0	1	2	3	4	5	6	Total
0	46.4% (813)	13.6% (239)	7.6% (133)	4.0% (70)	3.5% (62)	1.4% (25)	0.0% (0)	76.6% (1342)
-1	7.9% (139)	1.0% (17)	0.3% (5)	0	0	0		9.2% (161)
-2	5.4% (95)	0% (1)	0	0	0			5.5% (96)
-3	3.0% (53)	0	0	0				3.0% (53)
-4	3.4% (60)	0	0					3.4% (60)
-5	1.3% (23)	0						1.3% (23)
-6	1.0% (17)							1.0% (17)
Total	68.5% (1200)	14.7% (257)	7.9% (138)	4.0% (70)	3.5% (62)	1.4% (25)	0	100% (1752)

Several important issues emerge from the hot/cold spot summation table⁵. Most notably, is the lack of sub-kebeles that had both hot and cold values. More specifically, only 23, or 1.3%, of all sub-kebeles had both positive and negative values. This can be contrasted to 916, or 52%, of sub-kebeles that had either hot or cold values but not both. This suggests that there are certain areas prone to both above and below damage over time. Consistent areas of above and below damage which is explored later.

Table 5: Cold/Hot Damage (per area)

Sub-Kebele Group	Obs	Hot/Cold Spots	Crop Damage	Elevation (m)	NDVI-AOC	PSNP (% of Total)
Control	1459	-0.21	8.80%	1982	225426	30.00%
Drought	293	1.42**	15.5%**	1,855**	162,396**	76.5%**

Access to Extension	Access to Extension (not Tigray)	
	0.12	-0.06
0.52**		-0.01

⁵ All possible values have been inserted into the table and zeros are entered where no value exists but would be theoretically possible. In other words, since there was only six crop seasons included, a value of 5, -2 would not be possible.

** Statistically significant at the 99% level (using relevant normal or non-parametric tests)

The values of Table 5 were divided into base years (2010-2014) and the specific 2015 hot drought year. The base years were averaged to determine “typical values” for each sub-kebele. The goal was to explore predictable patterns for the 2015 drought versus non-drought sub-kebeles. Importantly, many statistically significant differences emerge and suggest that base-line high damage areas were most effected by the drought and suggest a predictability to the impacts.

In total, 293 (16.7%) sub-kebeles experienced statistically significant, above average, neighborhood drought effects. For a variety of variables, statistical differences were found between the drought sub-set and the overall control group. At the very least, this suggests a non-random pattern to the 2015 drought.

The control and drought groups are divided areas that either experienced 2015 neighborhood drought effects or not. The control group represents 1,459 or 83.3% of the entire sample and the remaining are the 2015 drought sub-group. Averages of all variables indicated in Table 5 are taken for 2010-2014 crop seasons and separated by categorical. The results are striking and confirm that the drought effects were not randomly distributed. For example, on average the control sub-kebeles experienced 0.2 number of cold spots across all years and essentially suggest no neighborhood effects of damage. On the other hand, each drought affected sub-kebele had an average of 1.4 damage hotspots. This indicates that areas that suffered from the drought, already where in high damage neighborhoods. This is confirmed by the percentage average damage from 2010-14 being about 76% greater on average (15.5% vs. 8.8%) BEFORE the drought year effects. Other variables such as elevation and area of NDVI further suggest significant differences between the groups. The final four columns seek to identify relative policy actions existing in these two groups.

Ethiopia's Poverty Safety Net Program (PSNP) is an internationally recognized safety net program that provides a variety of social programs to assist poorer households. There are approximately 200 woredas designated as PSNP. For our sample, 660 or 37.7% are in PSNP designated woredas. However, when separated by drought effects, about 30% are in non-drought areas and an astonishing 76.5% of PSNP samples are in drought neighborhoods. In other words, ¾ of all of the drought woredas are in PSNP designated areas, despite the fact that PSNP sub-kebeles make up less than 40% of the entire sample. This suggests that this government project is relatively well positioned, at least relative to crop damage effects. Average production losses, over the entire six-year period, in PSNP areas, is 17.9% versus 10.9% in non-PSNP areas. Another potential variable is extension services, although one could expect higher services in high production areas. Contrary to this expectation, extension services are disproportionately in drought areas but this is because the Tigray region experiences disproportionately high extension services. If Tigray is removed, there is no statistical difference.

On the other hand, Agricultural Growth Program (AGP) woredas were selected in 2010 to represent high growth potential areas and are targeted for increased production. As PSNP is social protection and AGP is high production we would expect little overlap of these two groups and actual comparisons confirm this to be the case. In our sample, contrasting the 190 woredas listed as PSNP and 81 listed as AGP, only 6 woredas are listed in both (5 from Tigray and 1 from SNNP). At the sub-kebele level, only 23 of the 1752 observations (1.3%) are listed in both. Unsurprisingly, these AGP woredas have far less damage to crops and suffered less in the 2015 drought. While the AGP sample represents 18.6% of all sampled sub-kebeles (326/1752), 21% of total sub-kebeles were not in high damage neighborhoods in 2015 and only 6.1%, or one-third the overall average, were in hot spots. Further, the non-AGP sub-kebeles had an average of 0.21 hotspots over the 2010-2015 period while the AGP woredas had an average of 0.47 cold spots over the six crop seasons. Finally, in terms of average crop losses over the six crop seasons, AGP sub-kebeles only had an average of 10.8% damage as contrasted to non-AGP value of 14.2%. Overall, Table 5 reveal several statistically significant predictor variables that indicate the non-random effects of the drought.

3.2 Now Casting - Mid-Season OPH Estimates

Table 6: Now-Casting: Panel RE regression on output per hectare for maize in quintals

Panel RE regression

	Dependent variable:	
	MAIZEOPH_W	
	Full (1)	Final (2)
MAIZEOPH_W, 1)	0.2782*** (0.0142)	0.2828*** (0.0137)
G_mx	0.0031*** (0.0006)	0.0029*** (0.0005)
G_mx_Qnt	-0.0018*** (0.0005)	-0.0018*** (0.0005)
G_AUC_leading	-0.00001 (0.00001)	
G_height	0.0010*** (0.0004)	0.0012*** (0.0003)
PET_G_AUC_leading	0.0007*** (0.0001)	0.0006*** (0.0001)
ETA_G_AUC_leading	-0.0002 (0.0005)	
PPT_G_AUC_leading	-0.0017* (0.0009)	-0.0024*** (0.0007)
lag(G_height, 1)	0.0001 (0.0003)	
lag(season_length, 1)	-0.0494** (0.0204)	-0.0386** (0.0177)
lag(PET_G_mn, 1)	-0.0788*** (0.0225)	-0.0732*** (0.0210)
lag(PET_G_mx, 1)	0.0208* (0.0107)	0.0243*** (0.0081)
lag(PET_G_AUC, 1)	0.0011** (0.0005)	0.0009** (0.0004)
lag(PET_G_sd, 1)	0.0098 (0.0175)	
lag(PPT_G_mn, 1)	0.5977*** (0.2072)	0.5751*** (0.1946)
lag(PPT_G_mx, 1)	0.0100 (0.0230)	
lag(PPT_G_AUC, 1)	-0.0100* (0.0051)	-0.0100** (0.0046)
lag(PPT_G_sd, 1)	-0.0644 (0.0846)	
lag(MAIZEIMSEED, 1)	0.0093* (0.0051)	0.0179*** (0.0039)
lag(MAIZEDAMAGEAREA_P, 1)	-0.6969 (0.7407)	
lag(MAIZEFERT_CHEMICAL_AREA_P, 1)	2.4212*** (0.6323)	2.8591*** (0.6091)
lag(Fert_Amt_Per_Area, 1)	0.0007 (0.0004)	0.0008* (0.0004)
lag(MAIZEEXTAREA, 1)	0.2640**	

	(0.1044)	
lag(MAIZEIMSEED_P, 1)	0.0388** (0.0165)	0.0273* (0.0160)
dist_pp50k	0.000001 (0.00001)	0.000001 (0.00001)
bs(elevation, 3)1	7.9240* (4.4083)	8.1318* (4.4075)
bs(elevation, 3)2	-5.9525* (3.2616)	-6.3998** (3.2556)
bs(elevation, 3)3	2.9830 (6.0532)	3.2109 (6.0528)
Constant	14.8922* (7.9471)	11.2996 (6.8916)
<hr/>		
Zone Dummy	Yes	Yes
<hr/>		
Observations	5,051	5,051
R2	0.4047	0.4028
Adjusted R2	0.3953	0.3943
F Statistic	42.7821***	47.3026***
<hr/>		
Note:	*p<0.1; **p<0.05; ***p<0.01	

Looking at the results from Table 6 with regression results from Maize OPH (quintals per hectare) we can see the results from two model outputs. The first, “Full”, presents the results based on the full set of early season variables and includes zonal fixed effects (FE). The second, “Final”, narrows the variable selection.

In Table 6 we can see a decent adjusted R² of 0.39. Residuals are also relatively small with 1st and 3rd quintiles spanning -7.44 and 6.34 quintals per hectare. Although R²s are reasonably high for real-world data, we believe these models could be significantly improved with the use of 30m harmonized landsat sentinel data. This data should be released in 2018 and will have 5-day coverage. Implying one image is captured every 5-days. This would likely provide a high enough frequency to avoid clouds, and provide reasonable estimates of NDVI curves, but at a much higher resolution (compared to 250m cells used here).

Although interpretation is complex for this prediction-oriented model, most variables have the expected sign and significance. First we can look at contemporaneous variables measured in the current growing season. We can see that increased maximum greenness during the growing season (*G_mx*) corresponds to significantly higher yields (p<7.17e-06. A greater NDVI ‘height’, the difference between growing season minimum and maximum values (*G_height*), corresponds to significantly higher yields (p<1.12e-05. Next variables with the postfix “_leading” indicates that the variable summarizes data between the estimated planting date and maximum NDVI level in the Meher growing season. A higher total potential evapotranspiration⁶ values during ‘leading’ period of the growing season (*PET_G_AUC_leading*) corresponds to higher yields (p<1.68e-08. Interestingly, increased precipitation totals in the leading period of the season (*PPT_G_AUC_leading*) correspond to lower maize yields (p<1.68e-08. Next we look at one period lagged variables that give some indication of expected values for any given sub-kebele. The first (*MAIZEOPH_W, 1*) is a one-period lag of maize outputs per hectare, unsurprisingly this variable is both significant and major determinant (p<1.98e-89. We also see an inverse relationship between last year’s season length and yields (*lag(season_length, 1)*) (p<0.0173), where season length is the difference between estimated planting date and harvest date. This inverse relationship may be related to management practices, where farmers respond to weather or a failing season by trying to either plant early or harvest late. We also see the impact of PET in the previous season higher maximum temperatures and area under the curve (AUC) of PET corresponds to slightly higher yields (p<0.00275). Similarly, a higher mean precipitation last year (*lag(PPT_G_mn, 1)*) corresponds to significantly higher yields (p<0.00184), while higher total precip last year corresponds to slightly lower yields (p<0.0219), although this is likely a

⁶Potential evapotranspiration is a measure of evaporative demand. It is the amount of water that could move through a system if water was not a limiting factor and is therefore related to higher temperatures and low humidity.

non-linear response. We can also look at lagged values of management practices from the AgSS, with lagged improved seed use ($lag(MAIZEIMSEED, 1)$), the percentage ($lag(MAIZEFERT_CHEMICAL_AREA_P, 1)$), and intensity of chemical fertilizers ($lag(Fert_Amt_Per_Area, 1)$) all significantly related to yields now ($p < 2.08e-06$, $p < 5.99e-06$, $p < 0.0785$, respectively). We can also see the impacts of market access the variable $dist_pp50k$ with increasing distances to population centers of 50k+ people decreasing yields ($p < 0.97$).

Table 7: Now-Casting: Panel RE regression on output per hectare for wheat in quintals

Panel RE regression

	Dependent variable:	
	WHEATOPH_W	
	Full (1)	Final (2)
G_mx	0.0028*** (0.0003)	0.0030*** (0.0002)
G_AUC_leading	-0.00005*** (0.00001)	-0.0001*** (0.00001)
G_height	0.0002 (0.0003)	
PET_G_AUC_leading	0.0008*** (0.0001)	0.0008*** (0.0001)
ETA_G_AUC_leading	0.0002 (0.0004)	
PPT_G_AUC_leading	-0.0009 (0.0007)	
WHEATOPH_W, 1)	0.2559*** (0.0141)	0.2559*** (0.0136)
lag(G_height, 1)	-0.00002 (0.0002)	
lag(season_length, 1)	0.0312** (0.0156)	0.0353*** (0.0112)
lag(PET_G_mn, 1)	-0.0120 (0.0177)	
lag(PET_G_mx, 1)	-0.0144* (0.0086)	-0.0216*** (0.0039)
lag(PET_G_AUC, 1)	0.0001 (0.0004)	
lag(PET_G_sd, 1)	0.0277** (0.0141)	0.0333*** (0.0121)
lag(PPT_G_mn, 1)	0.5598*** (0.1600)	0.5465*** (0.0790)
lag(PPT_G_mx, 1)	0.0200 (0.0179)	0.0229*** (0.0080)
lag(PPT_G_AUC, 1)	-0.0133*** (0.0040)	-0.0134*** (0.0022)
lag(PPT_G_sd, 1)	0.0072 (0.0662)	
lag(WHEATIMSEED, 1)	0.0019 (0.0012)	0.0030*** (0.0009)
lag(WHEATDAMAGEAREA_P, 1)	0.1109 (0.7415)	

lag(WHEATFERT_CHEMICAL_AREA_P, 1)	1.6010*** (0.3873)	1.7386*** (0.3747)
lag(Fert_Amt_Per_Area, 1)	0.0005 (0.0006)	
lag(WHEATEXTAREA, 1)	0.0082 (0.0377)	
lag(WHEATIMSEED_P, 1)	0.0070 (0.0048)	
dist_pp50k	-0.00001 (0.00001)	
bs(elevation, 3)1	13.9993*** (3.5715)	14.2694*** (3.5285)
bs(elevation, 3)2	8.8182*** (1.4739)	8.8200*** (1.4484)
bs(elevation, 3)3	9.7192*** (2.5937)	9.8421*** (2.5647)
Constant	-22.0545*** (7.3098)	-23.7424*** (5.9855)

Zone Dummy	Yes	Yes
Observations	5,114	5,115
R2	0.3543	0.3514
Adjusted R2	0.3443	0.3429
F Statistic	35.4135***	41.4393***

Note:

*p<0.1; **p<0.05; ***p<0.01

In Table 7 we can see a decent adjusted R² of 0.34. Residuals are also relatively small with 1st and 3rd quintiles spanning -5.65 and 4.66 quintals per hectare.

Comparing the results for maize with those here for wheat, we can see interesting differences related to plant physiology, demands, and management. First we can look at contemporaneous variables measured in the current growing season. Like maize we can see that increased maximum greeness during the growing season (*G_mx*) corresponds to significantly higher yields (p<1.1e-36), but the sum of NDVI area (*G_AUC_leading*) corresponds to slightly lower yields (p<1.23e-14). This may reflect higher yields in areas with a more rapid greenup (from minimum to peak greeness). Similarly, we see higher ‘leading’ period potential evapotranspiration (*PET_G_AUC_leading*) increases yields (p<2.65e-16). Looking at lagged variables we can see that again lagged OPH value are a good predictor of the current year (p<6.82e-77). As opposed to maize, longer growing seasons in the previous year (*lag(season_length, 1)*) corresponds to higher yields (p<0.00165). Also opposite of maize, higher evaporative demands in the previous year (*lag(PET_G_mx, 1)*) correspond to lower yields (p<0.00165), reflecting wheat’s preference for cooler zones. For precipitation, similar to maize, last years mean (*lag(PPT_G_mn, 1)*) and maximum precipitation rates (*lag(PPT_G_max, 1)*) correspond with higher yields (p<5.17e-12, p<5.17e-12 respectively), while higher previous year precipitation (*lag(PPT_G_AUC, 1)*) decreases them (p<6.95e-10). For AgSS data, we can see that increasing the quantity of improved seeds (*lag(WHEATIMSEED, 1)*) in the previous year, corresponds to higher yields in this one (p<0.00113). We also see benefits from increasing the percentage covered by chemical fertilizers (*lag(WHEATFERT_CHEMICAL_AREA_P, 1)*) (p<3.58e-06).

Table 8: Now-Casting: Panel RE regression on output per hectare for barley in quintals

Panel RE regression

Dependent variable:

	BARLEYOPH_W	
	Full (1)	Final (2)
G_mx	0.0020*** (0.0002)	0.0018*** (0.0002)
G_AUC_leading	-0.00001 (0.00001)	
G_height	-0.0001 (0.0002)	
PET_G_AUC_leading	0.0004*** (0.0001)	0.0003*** (0.0001)
ETA_G_AUC_leading	0.00002 (0.0003)	
PPT_G_AUC_leading	-0.0021*** (0.0005)	-0.0024*** (0.0004)
BARLEYOPH_W, 1)	0.2512*** (0.0130)	0.2529*** (0.0126)
lag(G_height, 1)	-0.0003** (0.0002)	-0.0003** (0.0001)
lag(season_length, 1)	-0.0172 (0.0122)	
lag(PET_G_mn, 1)	-0.0015 (0.0138)	
lag(PET_G_mx, 1)	-0.0254*** (0.0066)	-0.0222*** (0.0056)
lag(PET_G_AUC, 1)	0.0004 (0.0003)	0.0003*** (0.0001)
lag(PET_G_sd, 1)	0.0267** (0.0107)	0.0254** (0.0104)
lag(PPT_G_mn, 1)	0.3245*** (0.1241)	0.3609*** (0.0814)
lag(PPT_G_mx, 1)	0.0059 (0.0139)	
lag(PPT_G_AUC, 1)	-0.0059* (0.0031)	-0.0070*** (0.0020)
lag(PPT_G_sd, 1)	-0.0298 (0.0504)	
lag(BARLEYIMSEED, 1)	0.00002 (0.0022)	
lag(BARLEYDAMAGEAREA_P, 1)	-0.2724 (0.5207)	
lag(BARLEYFERT_CHEMICAL_AREA_P, 1)	1.4540*** (0.3199)	1.4035*** (0.3109)
lag(Fert_Amt_Per_Area, 1)	0.0001 (0.0005)	
lag(BARLEYEXTAREA, 1)	0.0113 (0.0725)	
lag(BARLEYIMSEED_P, 1)	0.0037 (0.0104)	
dist_pp50k	0.000004 (0.000004)	
bs(elevation, 3)1	3.0097 (2.1364)	2.7110 (2.1068)

bs(elevation, 3)2	6.1522*** (0.9764)	6.1983*** (0.9617)
bs(elevation, 3)3	7.3919*** (1.5324)	7.1424*** (1.5093)
Constant	3.9554 (4.8000)	-1.3144 (2.2091)
-----	-----	-----
Zone Dummy	Yes	Yes
-----	-----	-----
Observations	5,854	5,854
R2	0.3721	0.3728
Adjusted R2	0.3637	0.3657
F Statistic	43.8820***	52.9167***
=====	=====	=====

Note: *p<0.1; **p<0.05; ***p<0.01

In Table 8 we can see a decent adjusted R² of 0.37. Residuals are also relatively small with 1st and 3rd quintiles spanning -4.93 and 4.03 quintals per hectare.

Again we can see similarities with other crop models. Like maize, we can see that increased maximum greeness during the growing season (*G_mx*) corresponds to significantly higher yields (p<3.9e-24), but see no impact from other contemporaneous measures of NDVI. Similar to all models, we also see *PET_G_AUC_leading* increasing yields (p<1.62e-07), but no other ‘leading’ variables. Lagged PET variables are significant however, with greater PET variability and AUC increasing yields (p<0.0132, p<0.00794 respectively), and declining yields with higher last year maximum PET (p<7.16e-05). Again higher lagged mean precipitation *PPT_G_mn* increases yields, and declining with lagged precip AUC (p<0.000384). For AgSS variables, we only see improvements from last years increase in chemical fertilizer coverage (p<6.49e-06).

3.3 Predicting Drought Damages

Here we present results described in the earlier ‘Predicting Crop Losses from Drought’ section in methods. We utilize a random forest to predict losses greater than 25% due to drought using remotely sensed vegetation and weather data, as well as provide measures of test out-of-sample accuracy.

3.3.1 Maize Damages

We can evaluate accuracy with a number of metrics on our independent testing data set described in the “Testing Data: Out-of-sample Performance” section above. Here we compare both in-sample and out-of-sample performance. The first, the confusion matrix compares predicted values against the observed (reference). We also report the a) overall accuracy, 2) Kappa coefficient, and 3) the recall rate as seen in table 9 below:

Table 9: Confusion matrix for reported damages > 0.25% for Maize Out-Of-Sample
FALSE = Crop damages greater than 0.25% for Maize

	FALSE	TRUE
FALSE	41	12
TRUE	30	1205

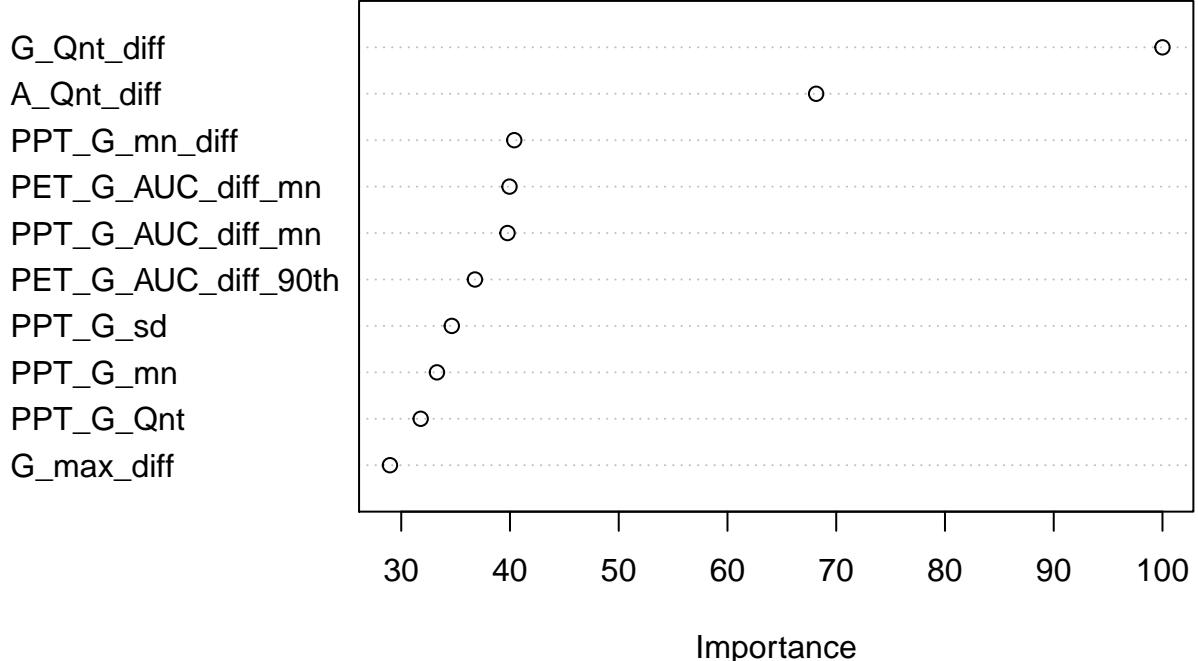
Table 10: Performance metrics for reported damages > 0.25% for Maize

	Accuracy	Kappa	Recall
Value	0.97	0.64	0.58

Overall, we find a high level of accuracy in our completely independent *testing* data with overall accuracy of 97 percent. We see reasonable identification of reported maize crop losses with a recall rate of 0.58, indicating that 58% of all substantial loss cases were predicted. In all, 41 out of 71 reported cases were predicted, and 12 cases were falsely predicted as crop failure areas. Although the recall rate is relatively low, we can still say that we can identify healthy maize harvest 1205 out of 1217 times, or 99% of the time.

We can also examine the relative importance of each variable to the national-level model, as well as its estimated non-linear relationship with the dependent variable through partial dependency plots. First, let's examine the relative importance of each variable in the random forest through the role it played in Gini coefficient loss.

Figure 12: Top 10 variable importance plot for maize damage prediction

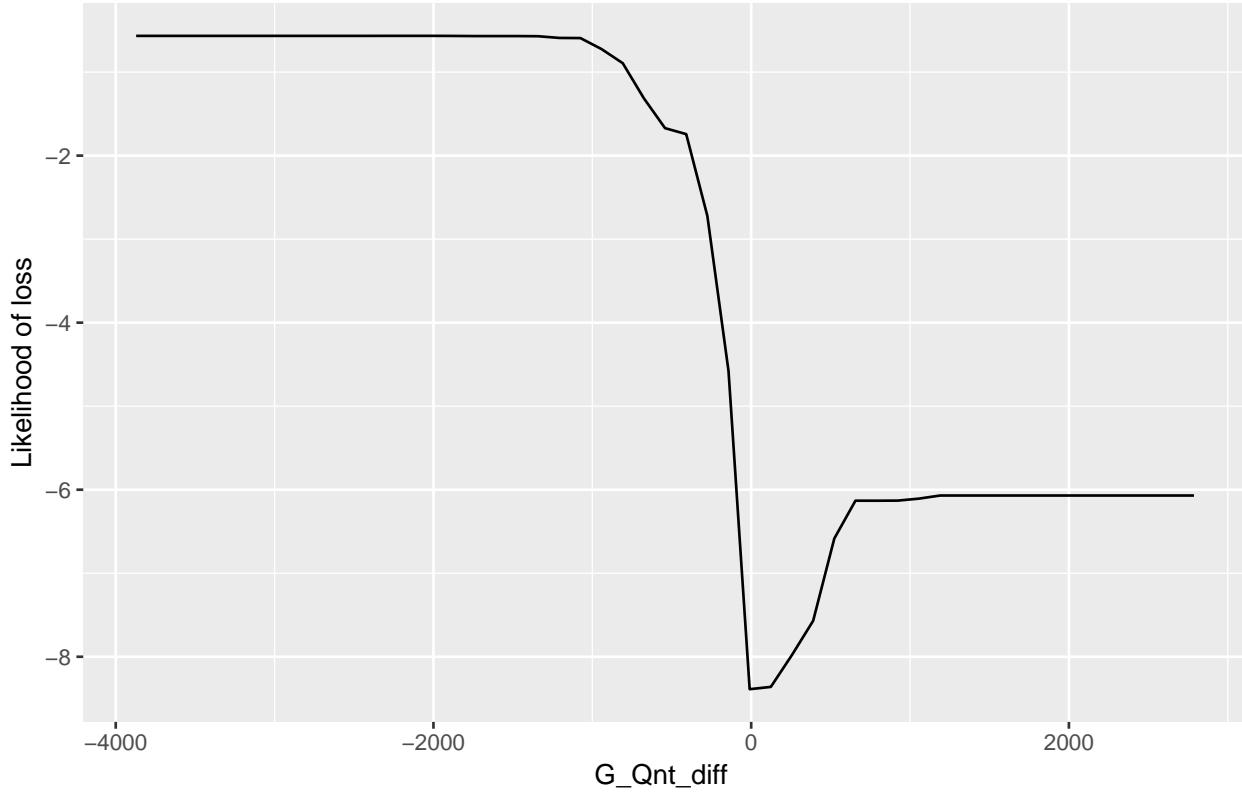


From the figure above we can see that the most important variables are comprised of measures of changes in NDVI (`G_Qnt_diff`, `A_Qnt_diff`, `G_max_diff`) as well as measures of change (`PPT_G_mn_diff`, `PPT_G_AUC_diff_mn`), mean (`PPT_G_mn`), 90th percentile values (`PPT_G_Qnt`) and standard deviation (`PPT_G_sd`) in precipitation. Changes in total potential evapotranspiration over the growing season, a measure of evaporative demand (largely controlled by temperature and humidity), also plays a key role (`PET_G_AUC_diff_mn`, `PET_G_AUC_diff_90th`). We can see that the most important variables for maize are largely comprised of measures of change in the mean and upper level of the distribution of greenness, precipitation, and potential evapotranspiration. This points to one of the advantages of machine learning, since these algorithms can exploit even minor difference between otherwise colinear variables in order to provide a reliable model fit.

We can see that `G_Qnt_diff` has the highest relative importance in the random forest model. `G_Qnt_diff` is measured as the first difference in the 90th percentile NDVI value for the growing season. As such the change in the highest NDVI values likely reflects a measure of change in plant health, with positive `G_Qnt_diff` values corresponding to increased plant health relative to the previous growing season. We can also see that the most important variables are largely comprised of variables summarizing NDVI, with the exception of `PPT_G_mn` which is the mean value of precipitation for each 10 day period across the growing season, and `PPT_G_AUC_Qnt`, the .

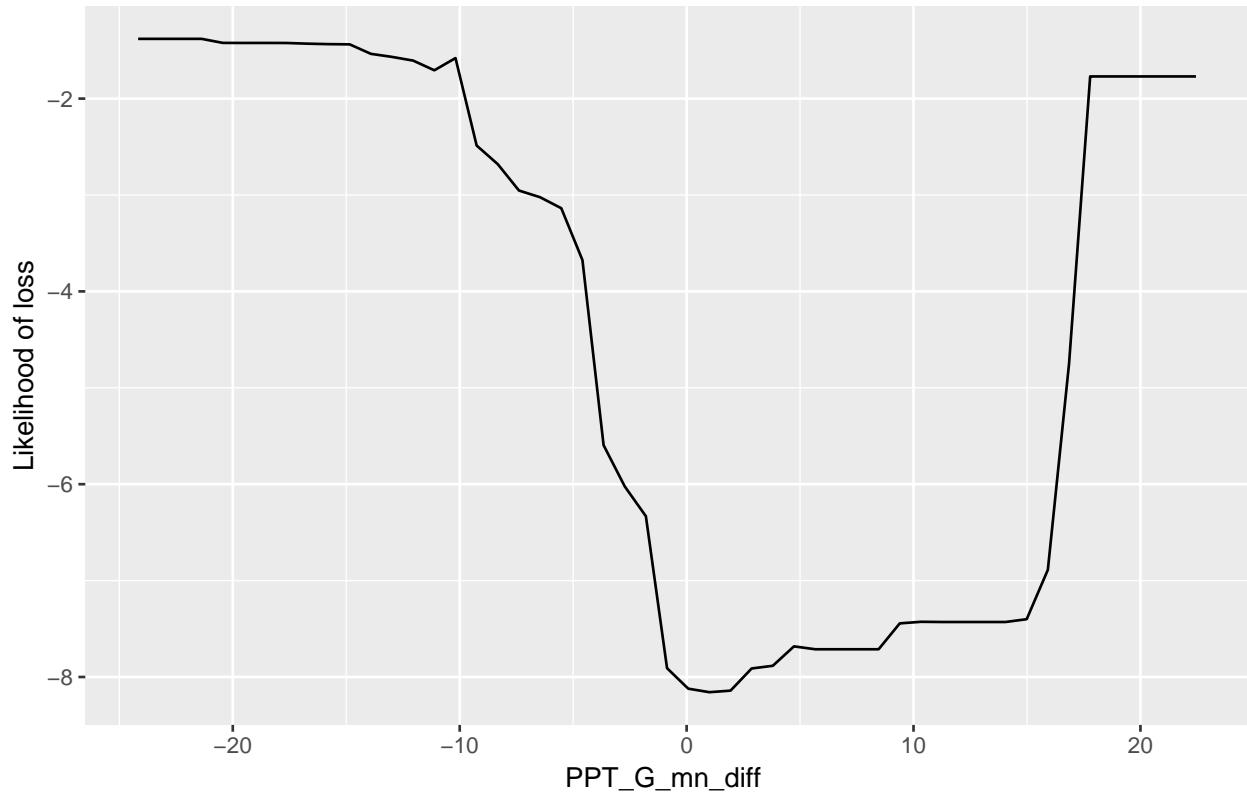
We can now look at the partial dependence plots to help understand the non-linear response of reported damages to variables of interest, such as `G_Qnt_diff`.

Figure 13: Partial dependency plot for influence of G_Qnt_diff on maize damage prediction



In the figure above, we can see how the likelihood of reporting maize losses decreases rapidly as the first difference of G_Qnt (G_Qnt_diff) approaches zero. G_Qnt, itself, measures the 90th percentile of NDVI values over the Meher growing season. Negative values of the first difference therefore indicate this the current year's 90th percentile NDVI is lower than the previous, and positive, higher. Therefore G_Qnt_diff picks up on the changes in peak values of NDVI. Similarly, we can look at the partial dependence plot for other variables of interest such as PPT_G_mn_diff.

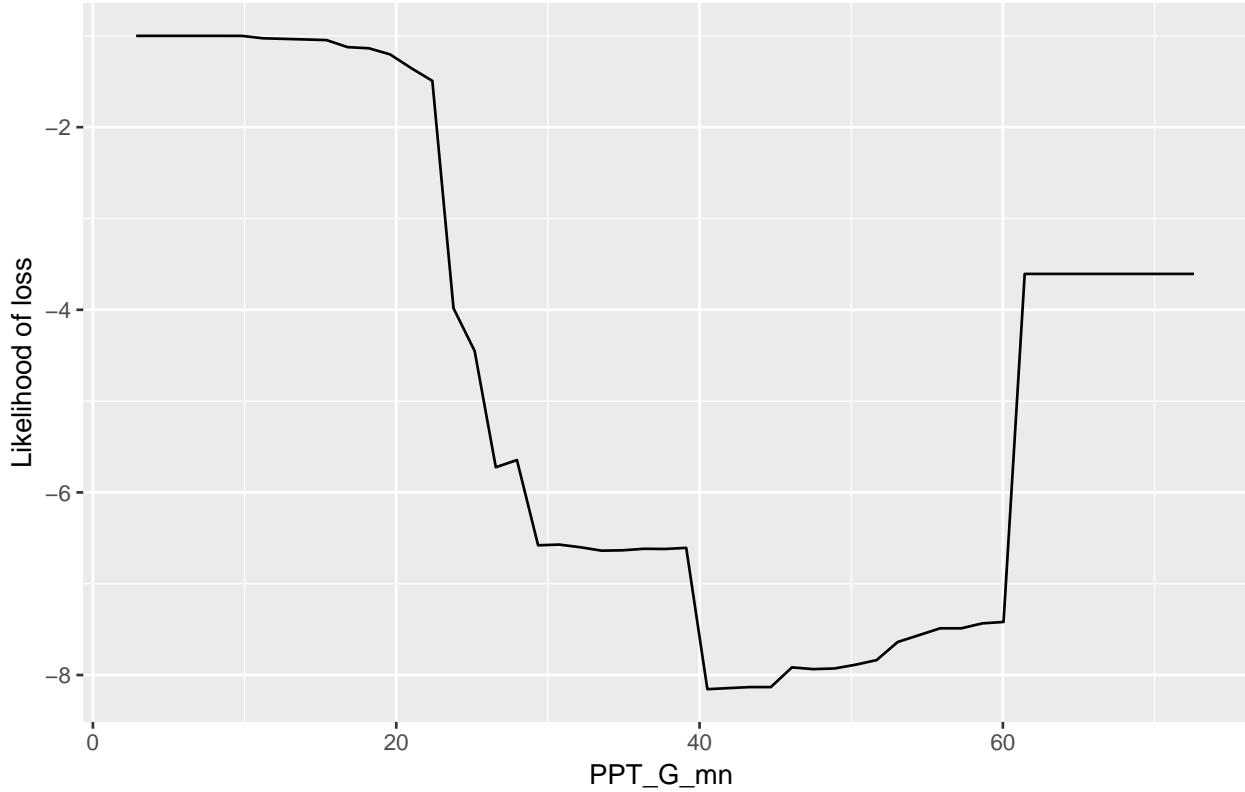
Figure 14: Partial dependency plot for influence of PPT_G_mn_diff on maize damage prediction



PPT_G_mn_diff tracks the effects of the change in average 10-day precipitation across the growing season. We can see that year-over-year declines (negative numbers) correspond to a much higher likelihood of crop losses due to drought. The likelihood of failure declines rapidly around a value of -10mm, and reaches its lowest likelihood just above 0mm. Past this, we again see an increase the likelihood of damage, but at a slower rate. Finally, years with extremely large increases in precipitation also correspond to increased damage due to excess precipitation or damaging storms.

We can see how average precipitation corresponds to drought related crop failures with variables such as PPT_G_mn.

Figure 15: Partial dependency plot for influence of PPT_G_mn on maize damage prediction



Similarly, we can see how the average 10-day precipitation effects the likelihood of loss. Losses decline rapidly above a 20mm 10-day average, reaching a minima around 40mm, but then increasing again after 60mm. Interestingly, the large uptick in reported losses due to drought at high levels of precipitation is unexpected. That being said this may have to do with issues of farmers confusing general weather related losses with drought related losses. From a policy perspective, this difference is likely trivial since interventions should target losses more generally.

3.3.2 Wheat Damages

The confusion matrix compares predicted values against the observed (reference). We also report the a) overall accuracy, 2) Kappa coefficient, and 3) the recall rate as seen in table 11 below:

*Table 11: Confusion matrix for reported damages > 0.25% for Wheat
 FALSE = Crop damages greater than 0.25% for Wheat*

	FALSE	TRUE
FALSE	18	3
TRUE	17	856

Table 12: Performance metrics for reported damages > 0.25% for Wheat

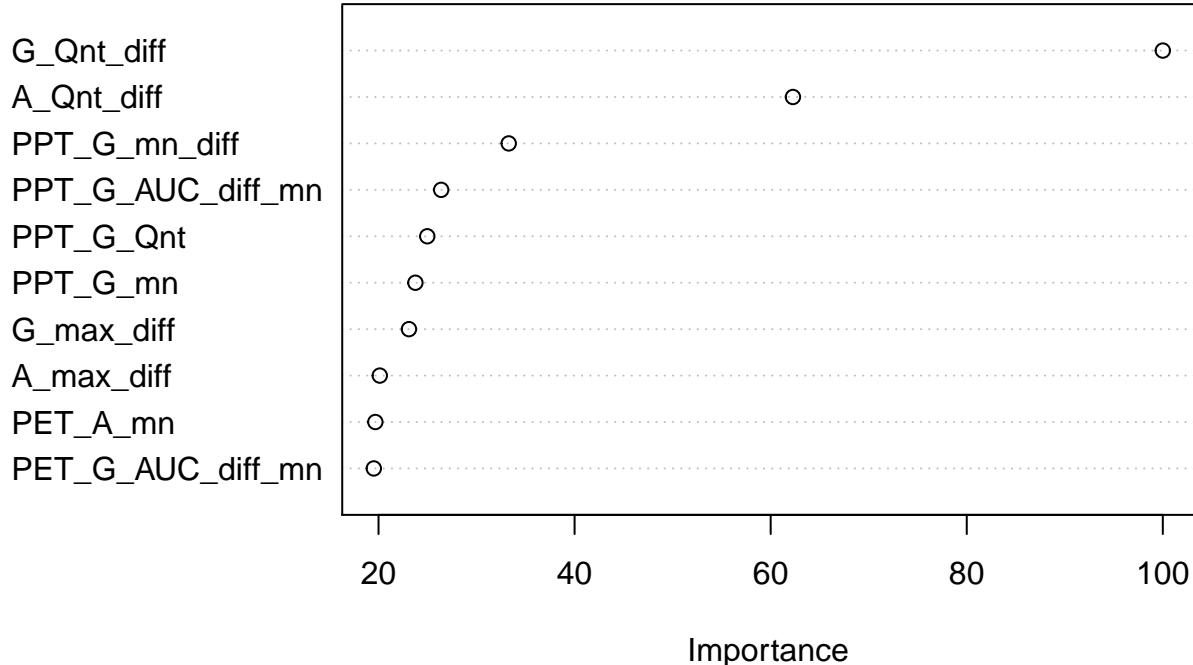
	Accuracy	Kappa	Recall
Value	0.98	0.63	0.51

Overall, we find a high level of accuracy in our completely independent *testing* data with overall accuracy of 98 percent. For the case of wheat, we see a decent identification of reported wheat crop losses with a recall

rate of 0.51, indicating that 51% of all substantial loss cases were predicted. In all, 18 out of 35 reported cases were predicted, and 3 cases were falsely predicted as crop failure areas. Although the recall rate is relatively low, we can still say that we can identify healthy maize harvest 856 out of 859 times, or 100% of the time.

We can also examine the relative importance of each variable to the national-level model, as well as its estimated non-linear relationship with the dependent variable through partial dependency plots. First, let's examine the relative importance of each variable in the random forest through the role it played in Gini coefficient loss.

Figure 16: Top 10 variable importance plot for maize damage prediction



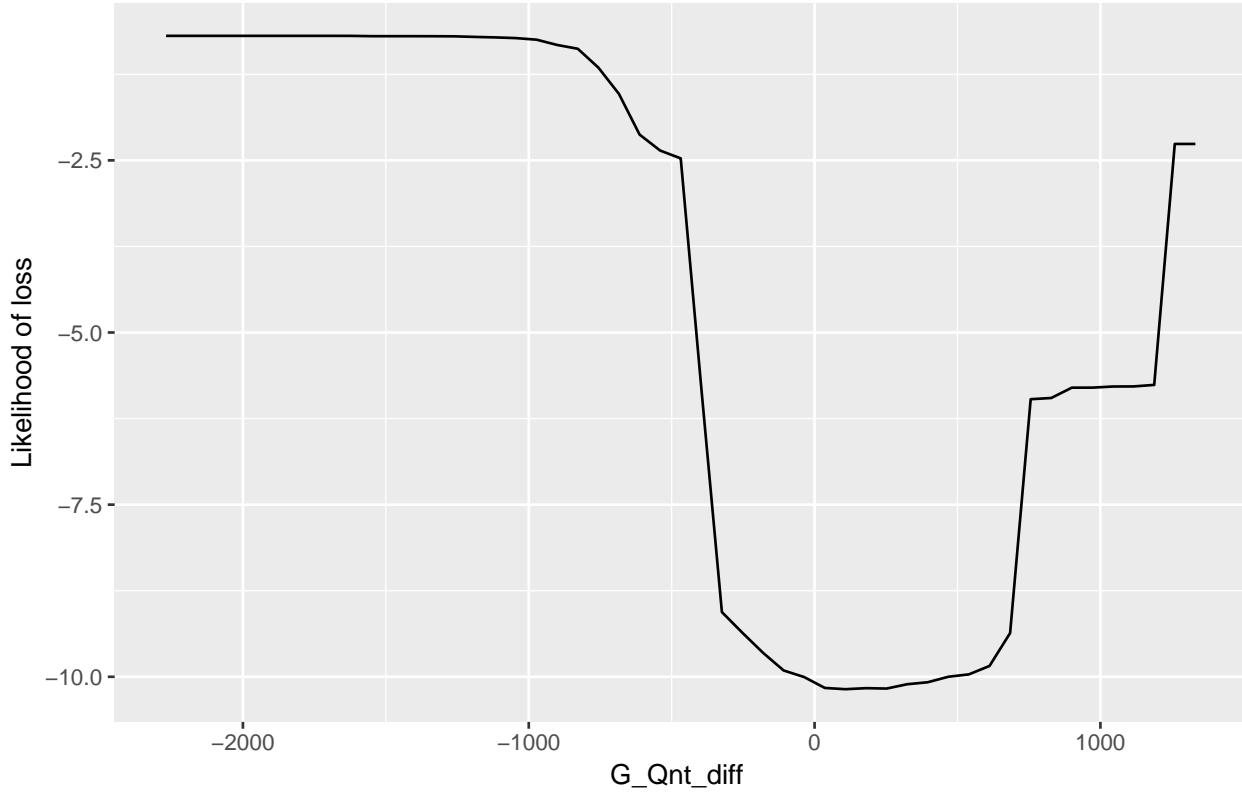
From the figure above we see a similar set of variables relative to maize, while the order and level of importance however has changed. This is an indication that our models are able to utilize the same broad set of phenological and weather variables to identify differences between crop types and how their losses can be observed.

We can see that G_Qnt_diff has the highest relative importance in the random forest model. G_Qnt_diff is the first difference of G_Qnt (G_Qnt_diff), where G_Qnt is the 90th percentile of NDVI values in the Meher season. As such the high NDVI values reflect a measure of plant health, with higher G_Qnt_diff values corresponding to strong plant health (higher leaf area, low water stress) at some point in the growing season. We can also see that the most important variables are largely comprised of variables summarizing NDVI, with the exception of PPT_G_mn which is the mean value of precipitation for each 10 day period across the growing season.

Interestingly, comparing figures 16 and 12, we see that measures of potential evapotranspiration (temperature and humidity) play a smaller role in losses for wheat compared to maize. This might reflect physiological differences between plants, where variability in temperatures more greatly effects yields for Maize. The differences between models validates our suite of variables as well as the need to run plant specific models.

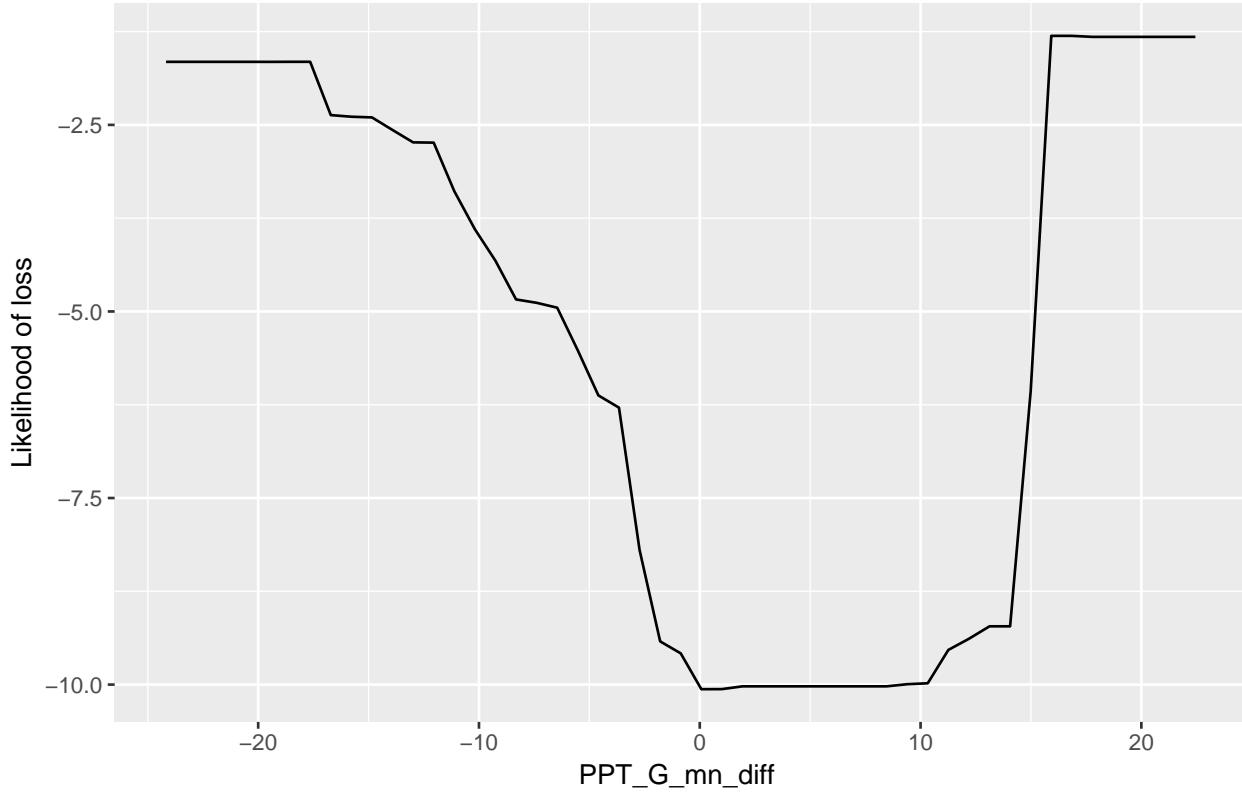
We can now look at the partial dependence plots to help understand the non-linear response of reported drought damages to G_Qnt_diff.

Figure 17: Partial dependency plot for influence of G_Qnt_diff on wheat damage prediction



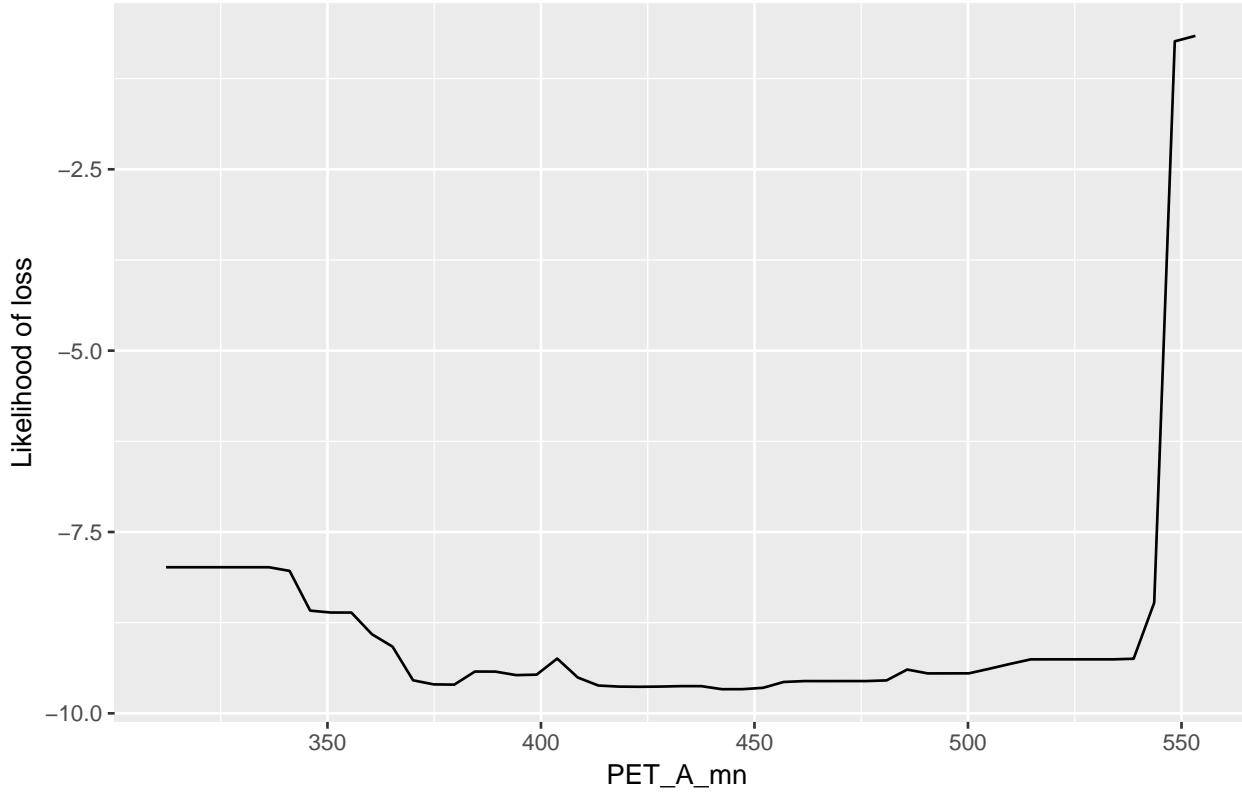
In the figure above, we can see with G_{Qnt_diff} , similar to maize, how the likelihood of reporting maize losses decreases rapidly after -500, with a minima above zero. The likelihood of losses again increases at the upper end of the distribution. These losses at upper end of the distribution might reflect poor management such as failure to control weeds, or narrow row spacing. The consistency across independent crop models (maize and wheat) is a strong indication that these models are picking up on meaningful non-linear relations. Similarly, we can look at the partial dependence plot for other variables of interest such as $PPT_G_mn_diff$.

Figure 18: Partial dependency plot for influence of $PPT_G_mn_diff$ on wheat damage prediction



Looking again at Figure 18, we can see similarities and differences between the likelihood of losses for wheat compared to maize. For wheat, we see that the likelihood of losses at very negative levels of PPT_G_mn_diff is both higher (-3 vs -2.25), and declines more slowly as the change in average precipitation increases. They both however hit a minima above 0. However wheat losses do not increase again, at all, until extremely high increases in precip year-over-year. Now we can look at the partial dependence plot for other variables of interest such as PET_A_mn.

Figure 19: Partial dependency plot for influence of PET_A_mn on wheat damage prediction



In the figure above, we can see the relationship between average annual potential evapotranspiration (PET_A_mn) and drought related losses. As expected losses are flat/stable until extremely high levels of PET, indicating high temperatures and low humidity. This plot again confirms both the critical importance of non-linear relationships and validates the approach used in this study.

3.3.3 Sorghum Damages

The confusion matrix compares predicted values against the observed (reference). We also report the a) overall accuracy, 2) Kappa coefficient, and 3) the recall rate as seen in table 13 below:

*Table 13: Confusion matrix for reported damages > 0.25% for Sorghum
FALSE = Crop damages greater than 0.25% for Sorghum*

	FALSE	TRUE
FALSE	32	9
TRUE	18	904

Table 14: Performance metrics for reported damages > 0.25% for Sorghum

	Accuracy	Kappa	Recall
Value	0.97	0.69	0.64

Overall, we find a high level of accuracy in our completely independent *testing* data with overall accuracy of 97 percent. For the case of Sorghum, we see a decent identification of reported Sorghum crop losses with a recall rate of 0.64, indicating that 64% of all substantial loss cases were predicted in our completely independent *testing* data. In all, 32 out of 50 reported cases were predicted, and 9 cases were falsely predicted as crop

failure areas. Although the recall rate is relatively low, we can still say that we can identify healthy maize harvest 904 out of 913 times, or 99% of the time.

3.3.4 Barley Damages

The confusion matrix compares predicted values against the observed (reference). We also report the a) overall accuracy, 2) Kappa coefficient, and 3) the recall rate as seen in table 15 below:

*Table 15: Confusion matrix for reported damages > 0.25% for barley
FALSE = Crop damages greater than 0.25% for barley*

	FALSE	TRUE
FALSE	28	7
TRUE	23	872

Table 16: Performance metrics for reported damages > 0.25% for barley

	Accuracy	Kappa	Recall
Value	0.97	0.63	0.55

Overall, we find a high level of accuracy in our completely independent *testing* data with overall accuracy of 97 percent. For the case of barley, we see a reasonable identification of reported losses with a recall rate of only 0.55, indicating that 55% of all substantial loss cases were predicted in our independent *testing* data. In all, 28 out of 51 reported cases were predicted, and 7 cases were falsely predicted as crop failure areas. Although the recall rate is relatively low, we can still say that we can identify healthy maize harvest 872 out of 879 times, or 99% of the time.

3.3.5 Teff Damages

The confusion matrix compares predicted values against the observed (reference). We also report the a) overall accuracy, 2) Kappa coefficient, and 3) the recall rate as seen in table 17 below:

*Table 17: Confusion matrix for reported damages > 0.25% for teff
FALSE = Crop damages greater than 0.25% for teff*

	FALSE	TRUE
FALSE	26	2
TRUE	26	1094

Table 18: Performance metrics for reported damages > 0.25% for teff

	Accuracy	Kappa	Recall
Value	0.98	0.64	0.5

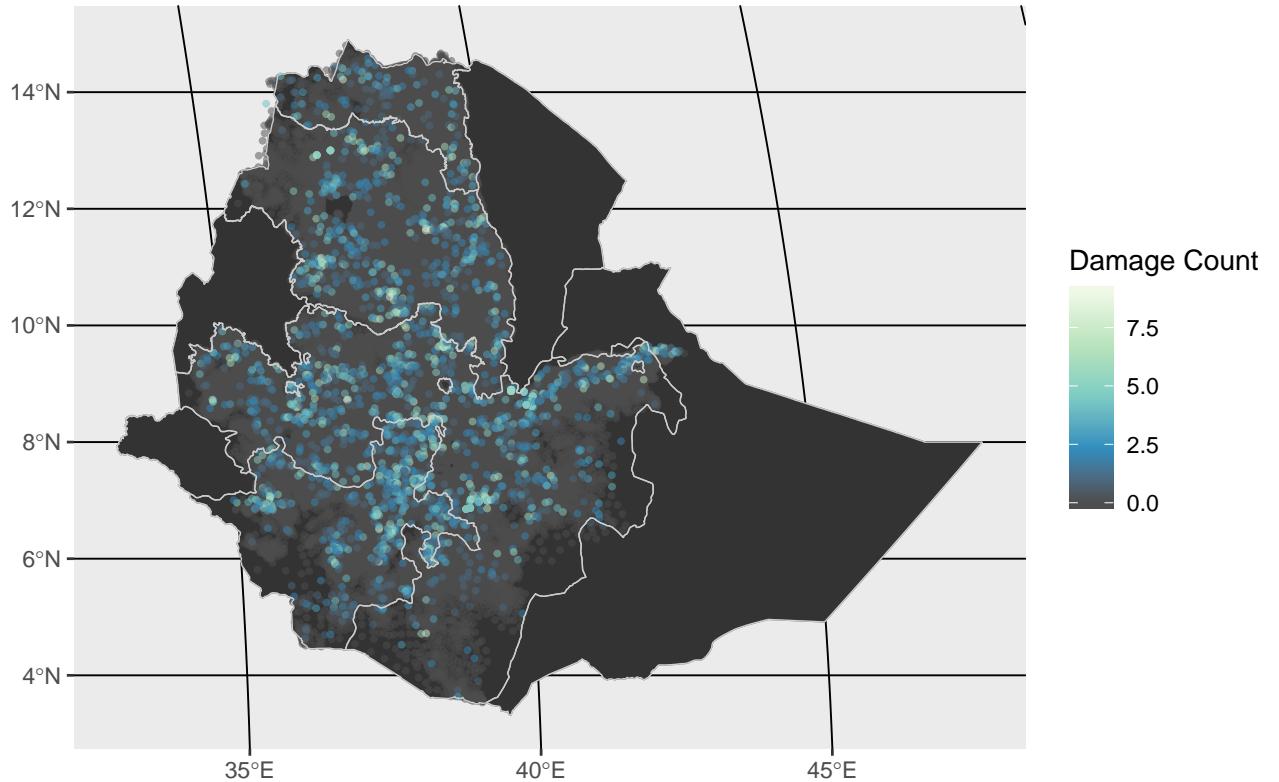
Overall, we find a high level of accuracy in our completely independent *testing* data with overall accuracy of 98 percent. For the case of teff, we see relatively good identification of reported losses with a recall rate of only 0.5, indicating that 50% of all substantial loss cases were predicted in our independent *testing* data.

In all, 26 out of 52 reported cases were predicted, and 2 cases were falsely predicted as crop failure areas. Although the recall rate is relatively low, we can still say that we can identify healthy maize harvest 1094 out of 1096 times, or 100% of the time.

3.4 Predict Damages Out-Of-Sample

One possible use case of these models is to predict losses outside of AgSS sample areas. This will allow policy makers to identify potential at-risk communities rapidly at low cost. Given the importance of maximum and 90th percentile values of NDVI in model performance, a reasonable estimate could be made mid-season (or just past peak greenness) similar to the approach in ‘Now Casting’. As example of how this might work we extracted NDVI statistics for all 75k+ sub-kebeles and used the model trained on AgSS sample areas.

Figure 20: Count of Predicted Damages (>25%) at the Sub-Kebele Level 2010-2015



In this particular example, we provide an estimate of the total number of drought related losses (for all crops) for the 2010-2015 growing seasons. As you can see maps like this could be used to identify isolated clusters of at-risk communities and especially those that experience repetitive losses. Additionally, taking from the “Now-Casting” approach, these same loss predictions could be provided midway through the growing season, acting as a method for rapid assessment, to help target food aid or agricultural extension.

4 Conclusions

In this paper we find evidence of localized clusters of drought in the 2015 period. However we find little to no evidence that these localized impacts effected crop yields in the most important growing areas. We find evidence however that particular areas are more prone to repetitive losses across an array of crops. Considering the *generally* positive impacts of improved seeds and chemical fertilizers (crop specific, please see

regressions), we believe information provided in this study could help target interventions in these especially risk-prone areas.

Although R²s and predictive accuracy are reasonably high for real-world data, we believe these models could be significantly improved with the use of 30m harmonized landsat sentinel (HLS) data. This data should be released in 2018 and will have 5-day coverage. Implying one image is captured every 5-days. This would likely provide a high enough frequency to avoid clouds, and provide reasonable estimates of NDVI curves, but at a much higher resolution (compared to 250m cells used here). We believe this will be critical to improve on our results going forward, as it will allow us to better distinguish between agricultural and non-agricultural cells, while avoiding cells that are a mixture of crops and forest. The methods developed through this study, and previous ones, could be easily transferred to the new HLS data stream, thus greatly improving our model accuracy, while minimizing further development time.

5 Appendix

5.1 Methods - Exploratory Analysis of Output Per Hectare Regression

5.1.1 Estimation Strategy

5.1.1.1 Variable Selection

One of the key challenges is how to decide on which of the 126 variables from AgSS and MODIS we want to include in the model. For this task we relied on a new suite of variable selection tools coming out of the computer science domain. We utilized VSURF that utilizes a two-stage strategy based on preliminary ranking of all explanatory variables for prediction using the random forests permutation-based score of importance, and finalized with a forward step-wise strategy for variable introduction.

Here we utilize VSURF to select unique sets of variables for each crop (barley, maize, wheat, teff). Given the complexity of these permutations for each crop (100s of millions) we utilize 15 cores of the GWU supercomputer Colonial One. Each run took between 10-24 hours.

5.1.1.2 Panel Regression

Compared to cross-sectional approaches, panel analysis can substantially increase the degree of observed variance over both space and time. For instance, in our current application, the integration of EAS (with $n = 1752$) over the 2010-2015 sample period ($t = 6$), results in a substantially larger data ($N \sim 6900$), allowing for a much greater flexibility and degrees of freedom in the modelling of the issue of interest. Here we include a few variables of interest to the reader including Zonal and annual dummies, and a lagged value of output per hectare. Here we estimate a series of random effect models of the following form 6:

(6)

$$Y_{itc} = \alpha_i + \sum_{k=1}^K \beta_k X_{it} + \mu_{it} + \epsilon_{it}$$

where Y_{itc} is the output per hectare (OPH) for crop c for sub-kebele i for year t , α is the average output per hectare for sample population (country), $\sum_{k=1}^K \beta_k X_{it}$ is the sum of all K coefficient estimates β_k for all variables X , μ_i is the sub-kebele specific random effect that measures the difference between the average OPH at sub-kebele i and the average OPH in the entire country. The term ϵ_{it} is the individual-specific effect, which can be described as the deviation of the i th OPH from the average for that year.

Here we present preliminary results from panel regressions on output per hectare for Ethiopia's major crops. These variables reflect the initial selection from the VSURF algorithm.

5.2 Results - Exploratory Analysis of Output Per Hectare Regression

Here we present results from panel regressions as outlined in the "Estimating Crop Yields" section. The dependent variable is crop yields (output per hectare) measured in quintals (1 quintal = 100 kilograms). The independent variables are initial selected using the VSURF algorithm described in the methods section. Here we present the findings of three major crops: maize, wheat, and barley. The remainder are presented in the appendix at the end of this document

Table 19: Panel RE regression on output per hectare for maize in quintals

Panel RE regression

=====
Dependent variable:

	MAIZEOPH_W			
	Full (1)	Reduced (2)	Final (3)	No AgSS (4)
MAIZEOPH_W, 1)	0.3262*** (0.0101)	0.3227*** (0.0101)	0.3265*** (0.0100)	0.3653*** (0.0103)
G_mn	0.0029** (0.0014)	0.0021 (0.0013)	0.0025* (0.0013)	0.0029** (0.0013)
G_AUC_leading	0.000004 (0.000004)			
G_AUC_Qnt	0.000003 (0.00002)			
A_mn	-0.0053*** (0.0017)	-0.0045*** (0.0016)	-0.0034*** (0.0012)	-0.0049*** (0.0013)
A_max	-0.00003 (0.0005)			
A_AUC	0.0000* (0.0000)	0.0000 (0.0000)		
A_Qnt	0.0018*** (0.0006)	0.0020*** (0.0005)	0.0027*** (0.0004)	0.0048*** (0.0004)
A_AUC_Qnt	-0.0000 (0.0000)			
A_max_Qnt	-0.0010** (0.0004)	-0.0010*** (0.0004)	-0.0007** (0.0003)	-0.0014*** (0.0003)
T_G_Qnt	0.0013* (0.0007)	0.0012* (0.0006)		
PPT_A_max	0.0010 (0.0083)	0.0001 (0.0078)		
PPT_G_AUC_Qnt	0.0020** (0.0010)	0.0023*** (0.0007)	0.0024*** (0.0007)	0.0027*** (0.0007)
PPT_G_mx_Qnt	-0.0204** (0.0094)	-0.0183** (0.0089)	-0.0189** (0.0083)	-0.0309*** (0.0086)
PPT_T_G_Qnt	0.0201 (0.0222)			
PET_A_min	0.0108*** (0.0032)	0.0125*** (0.0032)	0.0118*** (0.0031)	0.0169*** (0.0032)
PET_G_AUC_diff_mn	0.0006*** (0.0001)	0.0005*** (0.0001)	0.0006*** (0.0001)	0.0003** (0.0001)
MAIZEAREA	0.1171*** (0.0440)	0.0545 (0.0432)		
MAIZESEED1AREA	-0.0195 (0.1486)			
Fert_Amt_Per_Area		0.0015*** (0.0004)	0.0015*** (0.0004)	
MAIZEIMSEED	0.0216*** (0.0047)	0.0186*** (0.0031)	0.0191*** (0.0031)	
MAIZEDAMAGEAREA_P	-11.1709*** (0.5803)	-11.1543*** (0.5715)	-11.0321*** (0.5673)	
MAIZEFERT_CHEMICAL_AREA	-0.2302** (0.0970)			
MAIZEFERT_CHEMICAL_AMT	0.0002 (0.0001)			
MAIZEEXTAREA	0.2161** (0.0936)	0.1061 (0.0766)	0.1513** (0.0668)	

Y_COORD	-0.000001 (0.000003)	0.000001 (0.000003)		
X_COORD	0.00001*** (0.000003)	0.00001*** (0.000003)	0.00001*** (0.000003)	0.000004 (0.000003)
elevation	-0.0019*** (0.0003)			
splines::bs(elevation, 3)1		15.4945*** (3.8604)	15.5929*** (3.7878)	23.8067*** (3.9425)
splines::bs(elevation, 3)2		4.6195*** (1.5420)	4.5506*** (1.5263)	7.7403*** (1.5913)
splines::bs(elevation, 3)3		4.3731 (2.6649)	4.3404* (2.6305)	8.6528*** (2.7460)
c(Year)	0.7367*** (0.1042)	0.7053*** (0.1031)	0.6262*** (0.0813)	0.5059*** (0.0838)
Constant	2.0693 (5.0900)	-9.9053* (5.5259)	-7.2621** (2.9705)	-17.1149*** (3.0796)

Zone Dummy	Yes	Yes	Yes	Yes

Observations	8,241	8,241	8,241	8,241
R2	0.4630	0.4675	0.4696	0.4211
Adjusted R2	0.4576	0.4624	0.4648	0.4163
F Statistic	84.7407***	91.8726***	99.0421***	86.1521***
=====	=====	=====	=====	=====

Note:

*p<0.1; **p<0.05; ***p<0.01

Looking at the results from Table 19 with regression results from Maize OPH (quintals per hectare) we can see the results from four model outputs. The first, “Full”, presents the results based on the model selection from the VSURF algorithm and includes zonal fixed effects (FE). The second, “Reduced”, narrows the variable selection while including critical non-linearity for variables such as elevation, and adds a variable *Fert_Amt_Per_Area* which is the amount of chemical fertilizer applied per unit area. Non-linearity is controlled for using basis-splines which outperform polynomial representations in terms of stability and forecasting. The third model, “Final”, represents the final selection of significant variables for estimation. The last model, “No AgSS”, omits all AgSS survey variables to demonstrate the ability of the model to predict yields on the basis of remotely sensed data and last years yields.

Looking at the final model we can see an impressive adjusted R² of 0.46. Residuals are also relatively small with 1st and 3rd quintiles spanning -7.66 and 6.56 quintals. These findings are relatively impressive considering the highly-localized nature of these estimates (at the sub-kebele level), and the relatively coarse spatial resolution of the remotely sensed data (250m). Looking more closely we can observe some important features.

First, remotely sensed variables, while complex, have a few tangible interpretations. For instance, *G_mn* is positive and significant (p<0.0503) indicating that yields are higher if the mean greenness (NDVI) level during the Meher growing season is higher. Similarly, increases in the 90th percentile of greenness for the entire year (*A_Qnt*), corresponds to statistically significant higher yields (p<1.66e-12). Interestingly, the variable *A_mn*, which is the mean greenness level (NDVI) for the entire year, is negative and significant (p<0.00498). Already controlling for conditions during the Meher season, this might be an indication of two issues, a) a higher *A_mn* might indicate the failure to properly till or remove weeds between seasons, or b) complications of management in dual-crop areas, where planting occurs for both the minor (Belg) and major (Meher) rainy season.

Looking at precipitation variables we can see that *PPT_G_AUC_Qnt* yields are higher (p<0.000716) in sub-kebeles where the 90th percentile of total rainfall during the Meher season is higher. The opposite is true (p<0.0223) for areas with higher 90 percentile maximum rainfalls (*PPT_G_mx_Qnt*) likely reflecting the impact of flooding or excess rain events. Additionally, areas with higher minimum potential evaporation

rates (*PET_A_min*: higher min temps, more sun) receive better yields ($p < 0.000137$) while controlling for precip and greenness, and areas with above average total PET (*PET_G_AUC_diff_mn*) also receive yield benefits ($p < 7.83e-06$).

The AgSS data also provide some insights. We see scant evidence for yield improvements based on planted area alone (*MAIZEAREA* $p < 0.207$), but we do see the positive effects of fertilization rates (*Fert_Amt_Per_Area* $p < 0.000357$), the use of improved seeds (*MAIZEIMSEED* $p < 6.31e-10$), and the area of land in the sub-kebele that has utilized extension agents (*MAIZEEXTAREA* $p < 0.0234$). Conversely, we see the substantial impact of the percent of reported crop damage (*MAIZEDAMAGEAREA_P*), with every additional one percent of land damaged corresponding to a -11 unit loss in OPH ($p < 2.21e-82$).

Table 20: Panel RE regression on output per hectare for wheat in quintals

Panel RE regression

	Dependent variable: WHEATOPH_W			
	Full (1)	Reduced (2)	Final (3)	No AgSS (4)
WHEATOPH_W, 1)	0.2819*** (0.0123)	0.2786*** (0.0123)	0.2800*** (0.0123)	0.2995*** (0.0124)
G_mx	0.0011*** (0.0004)	0.0012*** (0.0004)	0.0012*** (0.0004)	0.0027*** (0.0004)
A_max_Qnt	-0.0010** (0.0004)	-0.0011*** (0.0004)	-0.0011*** (0.0004)	-0.0014*** (0.0004)
T_G_Qnt	0.0014*** (0.0005)	0.0014*** (0.0005)	0.0015*** (0.0005)	0.0007 (0.0005)
PPT_G_Qnt	-0.0486* (0.0278)	-0.0504* (0.0277)	-0.0514* (0.0277)	-0.0385 (0.0284)
PPT_G_sd	0.2000*** (0.0761)	0.2064*** (0.0760)	0.2084*** (0.0759)	0.1924** (0.0780)
PPT_T_G_Qnt	-0.0851*** (0.0157)	-0.0881*** (0.0155)	-0.0874*** (0.0154)	-0.0935*** (0.0158)
PET_G_mn	0.0073** (0.0037)	0.0071* (0.0037)	0.0074** (0.0037)	0.0040 (0.0038)
PET_G_sd	-0.0378*** (0.0105)	-0.0387*** (0.0105)	-0.0388*** (0.0104)	-0.0430*** (0.0107)
WHEATAREA	0.0524 (0.0731)	-0.0185 (0.0630)	0.0830*** (0.0292)	
WHEATNIMSEED	0.0005* (0.0003)	0.0004 (0.0003)		
WHEATNIMSEED_P	0.0008 (0.0008)			
WHEATDAMAGEAREA_P	-9.9877*** (0.6623)	-9.8762*** (0.6606)	-9.8837*** (0.6607)	
Fert_Amt_Per_Area		-0.0013 (0.0008)		
WHEATFERT_CHEMICAL_AMT	0.0007*** (0.0002)	0.0009*** (0.0003)	0.0006*** (0.0002)	
WHEATFERT_CHEMICAL_AREA	-0.0701 (0.0726)			
WHEATFERT_CHEMICAL_AREA_P	1.8323*** (0.3831)	1.3742*** (0.3406)	1.4219*** (0.3365)	
X_COORD	0.00001** (0.000003)	0.00001** (0.000003)	0.00001** (0.000003)	0.000001 (0.000004)
Y_COORD	0.000003 (0.000003)	0.000001 (0.000003)		
dist_rcap	0.000003 (0.000004)			
elevation	0.0017*** (0.0003)			
bs(elevation, 3)1		9.2350***	9.5196***	13.2382***

	(2.4118)	(2.3844)	(2.4551)
bs(elevation, 3)2	5.6638*** (1.2403)	5.4246*** (1.2319)	9.1046*** (1.2223)
bs(elevation, 3)3	7.2098*** (1.9324)	7.4203*** (1.9158)	8.9010*** (1.9799)
c(Year)	1.0661*** (0.0934)	1.0796*** (0.0931)	1.0799*** (0.0930)
Constant	-9.0136 (6.7203)	-7.0887 (6.7159)	-6.4993 (4.6303)
			-10.3012** (4.7753)
Zone Dummy	Yes	Yes	Yes
Observations	5,660	5,660	5,660
R2	0.4039	0.4059	0.4064
Adjusted R2	0.3962	0.3983	0.3990
F Statistic	52.5652***	53.0250***	55.4557***
			49.9115***

Note:

*p<0.1; **p<0.05; ***p<0.01

Looking at the results from Table 20 with regression results from Wheat OPH (quintals per hectare) we can see the results from four model outputs. The first, “Full”, presents the results based on the model selection from the VSURF algorithm and includes zonal fixed effects (FE). The second, “Reduced”, narrows the variable selection while including critical non-linearity for variables such as elevation, and adds key variables. The third model, “Final”, represents the final selection of significant variables for estimation. The last model, “No AgSS”, omits all AgSS survey variables to demonstrate the ability of the model to predict yields on the basis of remotely sensed data and last years yields.

In Table 20 we can see an impressive adjusted R² of 0.4. Residuals are also relatively small with 1st and 3rd quintiles spanning -5.64 and 4.62 quintals per hectare. Although interpretation is complex for this prediction-oriented model, most variables have the expected sign and significance. Increased maximum greeness during the growing season *G_mx* corresponds to significantly higher yields (p<0.00269). Interestingly, most measures of precipitation like *PPT_G_Qnt* (while controling for greeness, a proxy of plant health) correpsonds to lower yields, perhaps indicating that most wheat growing areas have suffered from excess rain during the sample period (p<0.0634). While higher mean evaporative demand *PET_G_mn* (related to higher temperatures and lower humidity) correspond to higher yields (p<0.0431). Like maize *A_max_Qnt* has the opposite sign perhaps for similar reason (poor management).

Looking at AgSS survey data we can see that improved seeds *WHEATNIMSEED* have a positive and nearly statistically significant effect (p<0.0537). Both the amount *WHEATFERT_CHEMICAL_AMT* and percentage covered *WHEATFERT_CHEMICAL_AREA_P* by chemical fertilizers had large statistically significant effects (p<0.000301, p<2.42e-05 respectively). Although unstable, we also see the possibility of gains from scale with area planted *WHEATAREA* having a small significant effect in the final model (p<2.42e-05).

Table 21: Panel RE regression on output per hectare for barley in quintals

Panel RE regression

	Dependent variable:			
	BARLEYOPH_W			
	Full (1)	Reduced (2)	Final (3)	No AgSS (4)
BARLEYOPH_W, 1)	0.2247*** (0.0122)	0.2273*** (0.0121)	0.2272*** (0.0121)	0.2420*** (0.0124)
G_mx	0.0007* (0.0004)	0.0007* (0.0004)	0.0007* (0.0004)	0.0010*** (0.0004)
G_Qnt	0.0020*** (0.0004)	0.0019*** (0.0004)	0.0019*** (0.0004)	0.0026*** (0.0004)
A_AUC_Qnt	-0.0000 (0.0000)			
A_max_Qnt	-0.0015*** (0.0003)	-0.0015*** (0.0003)	-0.0015*** (0.0003)	-0.0020*** (0.0003)
PPT_A_max	-0.0008 (0.0140)			
PPT_G_AUC	-0.0038 (0.0024)			
PPT_A_mn	0.1328 (0.0930)			
PPT_G_sd	-0.2603* (0.1413)	-0.0838 (0.1006)		
PPT_A_sd	0.3365** (0.1570)	0.1225 (0.1022)	0.0393** (0.0200)	0.0502** (0.0204)
PPT_G_AUC_leading	-0.0005 (0.0004)			
PPT_G_AUC_Qnt	0.0028*** (0.0010)	0.0022** (0.0008)	0.0022*** (0.0008)	0.0015* (0.0008)
PPT_G_mx_Qnt	-0.0228*** (0.0078)	-0.0168** (0.0075)	-0.0166** (0.0075)	-0.0188** (0.0075)
PPT_T_G_Qnt	-0.0717*** (0.0199)	-0.0704*** (0.0193)	-0.0723*** (0.0191)	-0.0651*** (0.0193)
PET_G_mx	0.0009 (0.0023)	0.0007 (0.0022)		
BARLEYFERT_CHEMICAL_AMT	0.0003 (0.0004)			
Fert_Amt_Per_Area		0.00003 (0.0006)		
BARLEYDAMAGEAREA_P	-7.4954*** (0.4995)	-7.4406*** (0.4959)	-7.4451*** (0.4953)	
BARLEYFERT_CHEMICAL_AREA	0.2160*** (0.0638)	0.2336*** (0.0505)	0.2330*** (0.0503)	
BARLEYFERT_NATURAL_AREA	-0.0454 (0.1131)			
Y_COORD	-0.000004 (0.000003)	-0.000004 (0.000003)	-0.000004 (0.000003)	-0.00001* (0.000003)
X_COORD	0.00001***	0.00001***	0.00001***	0.00001***

	(0.000003)	(0.000003)	(0.000003)	(0.000003)
dist_rcap	0.000004			
	(0.000003)			
elevation	0.0030***			
	(0.0003)			
bs(elevation, 3)1	1.7332	1.6812	2.2073	
	(2.0808)	(2.0792)	(2.0972)	
bs(elevation, 3)2	6.1018***	6.1017***	7.7045***	
	(0.9744)	(0.9717)	(0.9733)	
bs(elevation, 3)3	6.2416***	6.1964***	7.8908***	
	(1.5140)	(1.5080)	(1.5236)	
c(Year)	0.5863***	0.6669***	0.6811***	0.6120***
	(0.0830)	(0.0660)	(0.0638)	(0.0652)
Constant	2.5650	6.2490	6.2445	4.4223
	(4.7692)	(4.7926)	(4.7879)	(4.8361)
<hr/>				
Zone Dummy	Yes	Yes	Yes	Yes
<hr/>				
Observations	5,864	5,864	5,864	5,864
R2	0.3992	0.4003	0.4003	0.3793
Adjusted R2	0.3915	0.3931	0.3933	0.3723
F Statistic	51.2850***	55.2420***	57.7385***	54.4996***
<hr/>				

Note:

*p<0.1; **p<0.05; ***p<0.01

Looking at the results from Table 21 with regression results from Barley OPH (quintals per hectare) we can see the results from four model outputs. The first, “Full”, presents the results based on the model selection from the VSURF algorithm and includes zonal fixed effects (FE). The second, “Reduced”, narrows the variable selection while including critical non-linearity for variables such as elevation, and adds key variables. The third model, “Final”, represents the final selection of significant variables for estimation. The last model, “No AgSS”, omits all AgSS survey variables to demonstrate the ability of the model to predict yields on the basis of remotely sensed data and last years yields.

In Table 21 we can see an impressive adjusted R^2 of 0.39. Residuals are also relatively small with 1st and 3rd quintiles spanning -4.78 and 3.85 quintals per hectare. Although interpretation is complex for this prediction-oriented model, most variables have the expected sign and significance. We can see that increased maximum and 90th percentile greeness during the growing season (G_{max}, G_{Qnt}) corresponds to significantly higher yields ($p<0.0512$ and $8.65e-08$ respectively). Interestingly, controlling for precip and growing season greeness the annual maximum greenes A_{max_Qnt} is negatively related to yields ($p<3.02e-07$). This could be related to differences between dual and single crop areas. This however should be investigated further. Variability in precipitation during the Meher season PPT_G_{sd} has a negative but insignificant effect on yields ($p<0.405$), while higher precipitation variability during the entire year PPT_A_{sd} increases yields ($p<0.0495$). These findings likely point to the desirability of a consistent Meher rainy season, but more importantly to critical role the Belg rains play in soil moisture availability for both growing seasons. We can also see that the sum of precipitation across the Meher season $PPT_G_{AUC_Qnt}$ significantly increases yields ($p<0.00723$), while having a higher maximum observed rainfall during that season $PPT_G_{mx_Qnt}$ can reduce yields ($p<0.0258$) through greater damage to crops.

Looking at AgSS variables we can see the negative impacts of reported damage $BARLEYDAMAGEAREA_P$ ($p<4.03e-50$) and the positive impacts of increasing the coverage area of chemical fertilizers $BARLEYFERT_CHEMICAL_AREA$ ($p<3.67e-06$)

Table 22: Panel RE regression on output per hectare for teff in quintals

Panel RE regression

Dependent variable:				
	TEFFOPH_W			
	Full (1)	Reduced (2)	Final (3)	No AgSS (4)
TEFFOPH_W, 1)	0.2873*** (0.0109)	0.2873*** (0.0109)	0.2882*** (0.0108)	0.3113*** (0.0110)
G_mx_Qnt	-0.0001 (0.0036)			
G_AUC_diff_mn	0.000004 (0.00001)	0.000003 (0.00001)		
G_AUC_Qnt	0.00001* (0.00001)	0.00001*** (0.000003)	0.00001*** (0.000003)	0.00001** (0.000003)
A_sd	0.0020*** (0.0004)	0.0022*** (0.0003)	0.0023*** (0.0003)	0.0029*** (0.0003)
A_max_Qnt	0.0002 (0.0039)			
T_G_Qnt	0.0001 (0.0004)			
PPT_G_AUC	-0.0004 (0.0006)			
PPT_G_AUC_trailing	-0.0002 (0.0003)			
PPT_G_sd	0.0273 (0.0207)	0.0190 (0.0168)		
PPT_G_mx_Qnt	-0.0118** (0.0060)	-0.0079 (0.0058)		
PPT_G_AUC_Qnt	0.0019** (0.0008)	0.0015** (0.0006)	0.0013** (0.0006)	0.0018*** (0.0006)
PET_A_sd	-0.0151** (0.0063)	-0.0155** (0.0062)	-0.0167*** (0.0061)	-0.0223*** (0.0062)
PPT_T_G_Qnt	-0.0336** (0.0149)	-0.0331** (0.0147)	-0.0328*** (0.0126)	-0.0299** (0.0128)
TEFFDAMAGEAREA_P	-7.2108*** (0.4368)	-7.0764*** (0.4351)	-7.0871*** (0.4349)	
TEFFFERT_CHEMICAL_AMT	0.0001 (0.0001)	0.0003** (0.0001)	0.0003** (0.0001)	
Fert_Amt_Per_Area		-0.0010** (0.0005)	-0.0010** (0.0005)	
TEFFAREA	0.1034*** (0.0148)	0.0807*** (0.0163)	0.0805*** (0.0163)	
Y_COORD	-0.000004* (0.000002)	-0.000003* (0.000002)	-0.000003* (0.000002)	-0.000005** (0.000002)
X_COORD	0.00001*** (0.000002)	0.00001*** (0.000002)	0.00001*** (0.000002)	0.00001*** (0.000002)
dist_pp50k	0.000001 (0.000003)			
dist_rcap	0.00001** (0.000002)	0.00001*** (0.000002)	0.00001*** (0.000002)	0.00001** (0.000002)
elevation	0.0019*** (0.0002)			
bs(elevation, 3)1		-6.2332***	-6.3313***	-5.1210**

	(2.0305)	(2.0264)	(2.0549)
bs(elevation, 3)2	6.8509*** (0.8863)	6.8793*** (0.8823)	9.3317*** (0.8802)
bs(elevation, 3)3	-2.7897* (1.6085)	-2.9933* (1.6004)	-3.4363** (1.6195)
c(Year)	0.6047*** (0.0597)	0.5814*** (0.0587)	0.5414*** (0.0478)
Constant	0.1557 (3.3280)	4.5775 (3.2890)	5.0360 (3.2636)
5.6933* (3.3054)			
Zone Dummy	Yes	Yes	Yes
Observations	7,373	7,373	7,373
R2	0.3039	0.3094	0.3100
Adjusted R2	0.2965	0.3024	0.3033
F Statistic	41.3527***	44.1837***	46.1929***
			42.8010***

Note: *p<0.1; **p<0.05; ***p<0.01

In Table 22 we can see an impressive adjusted R² of 0.3. Residuals are also relatively small with 1st and 3rd quintiles spanning -4.42 and 3.1 quintals per hectare. Although interpretation is complex for this prediction-oriented model, most variables have the expected sign and significance. We can see that increased sum of greeness during the growing season (*G_AUC_Qnt*) corresponds to significantly higher yields (p<0.000385). Similar to barley we see an increase in yields corresponding to higher annual variability but this time in greeness *A_sd* (p<0.000385), here we assume again this has to do with better performance in areas with a strong Belg rain. We also see an increase in yields corresponding to higher total rainfall across the growing season *PPT_G_AUC_Qnt* (p<0.0425), and lower yields in areas with higher annual variability in rainfall *PET_A_sd* (p<0.00625). Looking at the basis splines for elevation *bs(elevation, i)*, teff is one of the few crops that seems to have a highly non-linear response where yeilds decline in low-land and upper highland areas.

Looking at AgSS variables we can see the positive impact of chemical fertilizers with *TEFF-FERT_CHEMICAL_AMT* (p<0.0356), but interestingly a negative influence from higher fertilizer amounts per unit area *Fert_Amt_Per_Area* (p<0.0413). Although this should be evaluated much further this might point to overly intensive use of chemical fertilizers in teff growing areas. This may be confirmed by the finding that yields improve with greater distance to regional capitals *dist_rcap* (p<0.00155). This might confirm the finding with fertilizers since more rural areas would have limited access to input markets, thereby reducing their utilization of fertilizers etc.

Table 23: Panel RE regression on output per hectare for sorghum in quintals

Panel RE regression

	Dependent variable:			
	SORGHUMOPH_W			
	Full (1)	Reduced (2)	Final (3)	No AgSS (4)
SORGHUMOPH_W, 1)	0.2571*** (0.0135)	0.2585*** (0.0135)	0.2591*** (0.0134)	0.2592*** (0.0137)
A_Qnt	0.0016*** (0.0004)	0.0016*** (0.0004)	0.0017*** (0.0002)	0.0024*** (0.0002)
G_mn	0.0002 (0.0003)	0.0002 (0.0003)		
PPT_T_G_Qnt	-0.0084 (0.0147)	-0.0085 (0.0147)		
PET_A_min	-0.0010 (0.0036)			
SORGHUMDAMAGEAREA_P	-9.5630*** (0.7066)	-9.5516*** (0.7049)	-9.5357*** (0.7043)	
SORGHUMFERT_CHEMICAL_AMT	-0.00002 (0.0002)			
Fert_Amt_Per_Area		0.0003 (0.0010)		
Y_COORD	0.00001** (0.000004)	0.00001** (0.000004)	0.00001** (0.000004)	0.00001** (0.000004)
X_COORD	0.00001*** (0.000004)	0.00001*** (0.000004)	0.00001*** (0.000004)	0.00001*** (0.000004)
dist_rcap	0.00001 (0.000005)	0.00001 (0.000005)	0.00001 (0.000004)	0.000001 (0.000004)
elevation	-0.0016*** (0.0004)			
bs(elevation, 3)1		-1.1277 (4.4048)	-1.1653 (4.3909)	-0.3590 (4.4724)
bs(elevation, 3)2		-2.9179 (2.5588)	-2.9431 (2.5297)	-2.4572 (2.5762)
bs(elevation, 3)3		-3.8284 (5.0268)	-3.9930 (5.0174)	-3.1931 (5.1099)
c(Year)	0.8416*** (0.0856)	0.8323*** (0.0795)	0.8378*** (0.0786)	0.7175*** (0.0795)
Constant	-9.8415 (6.2913)	-11.1115* (6.4818)	-11.4730* (6.4577)	-15.4878** (6.5710)
Region Dummy	Yes	No	No	No
Zone Dummy	Yes	Yes	Yes	Yes
Observations	5,108	5,108	5,113	5,113
R2	0.3066	0.3083	0.3083	0.2827
Adjusted R2	0.2977	0.2992	0.2997	0.2739
F Statistic	34.3031***	34.0366***	35.7203***	32.1070***

Note:

*p<0.1; **p<0.05; ***p<0.01

In Table 23 we can see an impressive adjusted R^2 of 0.3. Residuals are also relatively small with 1st and 3rd quintiles spanning -6.5 and 5.16 quintals per hectare. Although interpretation is complex for this prediction-oriented model, most variables have the expected sign and significance. We can see that increased 90th percentile of greenness during the growing season (A_Qnt) corresponds to significantly higher yields ($p<7.78e-14$). Interestingly we see no improvements from the amount $SORGHUMFERT_CHEMICAL_AMT$ or intensity $Fert_Amt_Per_Area$ of chemical fertilizer use for sorghum ($p<0.914$ and $p<0.752$ respectively).

6 References

- Bachewe, Fantu, Feiruz Yimer, and Bart Minten. 2016. *Agricultural Prices During Drought in Ethiopia: An Updated Assessment Using National Producer Data (January 2014 to June 2016)*. *Agricultural Prices During Drought in Ethiopia: An Updated Assessment Using National Producer Data (January 2014 to June 2016)*.
- Becker-Reshef, Inbal, Chris Justice, Mark Sullivan, Eric Vermote, Compton Tucker, Assaf Anyamba, Jen Small, et al. 2010. "Monitoring Global Croplands with Coarse Resolution Earth Observations: The Global Agriculture Monitoring (Glam) Project." *Remote Sensing* 2 (6). Molecular Diversity Preservation International: 1589–1609.
- Becker-Reshef, Inbal, Eric Vermote, Mark Lindeman, and Christopher 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). Elsevier: 1312–23.
- Benedetti, Roberto, and Paolo 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). Elsevier: 311–26.
- Biradar, Chandrashekhar M, and Xiangming Xiao. 2011. "Quantifying the Area and Spatial Distribution of Double-and Triple-Cropping Croplands in India with Multi-Temporal Modis Imagery in 2005." *International Journal of Remote Sensing* 32 (2). Taylor & Francis: 367–86.
- Bolton, Douglas K, and Mark A Friedl. 2013. "Forecasting Crop Yield Using Remotely Sensed Vegetation Indices and Crop Phenology Metrics." *Agricultural and Forest Meteorology* 173. Elsevier: 74–84.
- Butler, Ethan E, and Peter Huybers. 2015. "Variations in the Sensitivity of Us Maize Yield to Extreme Temperatures by Region and Growth Phase." *Environmental Research Letters* 10 (3). IOP Publishing: 034009.
- Didan, Kamel, and Alfredo Huete. 2006. "MODIS Vegetation Index Product Series Collection 5 Change Summary." *TBRS Lab, the University of Arizona*.
- Doraiswamy, Paul C, Sophie Moulin, Paul W Cook, and Alan Stern. 2003. "Crop Yield Assessment from Remote Sensing." *Photogrammetric Engineering & Remote Sensing* 69 (6). American Society for Photogrammetry; Remote Sensing: 665–74.
- Doraiswamy, Paul C, Alan J Stern, and Bakhyt Akhmedov. 2007. "Crop Classification in the Us Corn Belt Using Modis Imagery." In *Geoscience and Remote Sensing Symposium, 2007. Igarss 2007. Ieee International*, 809–12. IEEE.
- "Ethiopia: Drought - 2015-2017." 2017. *ReliefWeb*. <https://reliefweb.int/disaster/dr-2015-000109-eth>.
- Funk, Chris, and Michael E Budde. 2009. "Phenologically-Tuned Modis Ndvi-Based Production Anomaly Estimates for Zimbabwe." *Remote Sensing of Environment* 113 (1). Elsevier: 115–25.
- Funk, Chris, Pete Peterson, Martin Landsfeld, Diego Pedreros, James Verdin, Shraddhanand Shukla, Gregory Husak, et al. 2015. "The Climate Hazards Infrared Precipitation with Stations—a New Environmental Record for Monitoring Extremes." *Scientific Data* 2. Nature Publishing Group: 150066.
- Gao, Feng, Martha C Anderson, Xiaoyang Zhang, Zhengwei Yang, Joseph G Alfieri, William P Kustas, Rick Mueller, David M Johnson, and John H Prueger. 2017. "Toward Mapping Crop Progress at Field Scales Through Fusion of Landsat and Modis Imagery." *Remote Sensing of Environment* 188. Elsevier: 9–25.
- Gray, Josh, Mark Friedl, Steve Frolking, Navin Ramankutty, Andrew Nelson, and Murali Krishna Gumma. 2014. "Mapping Asian Cropping Intensity with Modis." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (8). IEEE: 3373–9.
- Idso, Sherwood B, Ray D Jackson, and Robert J Reginato. 1977. "Remote-Sensing of Crop Yields." *Science* 196 (4285). American Association for the Advancement of Science: 19–25.
- Li, Le, Mark A Friedl, Qinchuan Xin, Josh Gray, Yaozhong Pan, and Steve Frolking. 2014. "Mapping Crop

Cycles in China Using Modis-Evi Time Series.” *Remote Sensing* 6 (3). Multidisciplinary Digital Publishing Institute: 2473–93.

Lobell, David B., David Thau, Christopher Seifert, Eric Engle, and Bertis Little. 2015. “A Scalable Satellite-Based Crop Yield Mapper.” *Remote Sensing of Environment* 164. Elsevier: 324–33.

Lobell, David B., Wolfram Schlenker, and Justin Costa-Roberts. 2011. “Climate Trends and Global Crop Production Since 1980.” *Science* 333 (6042). American Association for the Advancement of Science: 616–20. doi:10.1126/science.1204531.

Major, JACK. 1967. “Potential Evapotranspiration and Plant Distribution in Western States with Emphasis on California.” *AMER ASSOC Adv SCI PUBL* 86.

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

Mann, Michael L., Enric Batllori, Max A Moritz, Eric K Waller, Peter Berck, Alan L Flint, Lorraine E Flint, and Emmalee 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). Public Library of Science: e0153589.

Mann, ML 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.

“Mid-Year Review, Ethiopia Humanitarian Requirements Document, July 2017.” 2017. *ReliefWeb*. <https://reliefweb.int/report/ethiopia/mid-year-review-ethiopia-humanitarian-requirements-document-july-2017>.

Mkhabela, MS, Paul 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). Elsevier: 385–93.

Portmann, Felix T, Stefan Siebert, and Petra Döll. 2010. “MIRCA2000—Global Monthly Irrigated and Rainfed Crop Areas Around the Year 2000: A New High-Resolution Data Set for Agricultural and Hydrological Modeling.” *Global Biogeochemical Cycles* 24 (1). Wiley Online Library.

Rasmussen, Michael 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). Taylor & Francis: 3431–42.

Ray, Deepak K, James S Gerber, Graham K MacDonald, and Paul C West. 2015. “Climate variation explains a third of global crop yield variability.” *Nature Communications* 6 (January). The Author(s): 5989. <http://dx.doi.org/10.1038/ncomms6989> <http://10.1038/ncomms6989> <http://www.nature.com/articles/ncomms6989{\#}supplementary-information>.

Schemm, Paul. 2016. “Ethiopia Is Facing a Devastating Drought, and Food Aid May Soon Run Out.” *The Washington Post*. WP Company. http://www.washingtonpost.com/sf/world/2016/02/22/history-repeats-itself-in-ethiopia/?utm_term=.2b7bb38f3a3e.

———. 2017. “Ethiopia Is Facing a Killer Drought. but It’s Going Almost Unnoticed.” *The Washington Post*. WP Company. https://www.washingtonpost.com/news/worldviews/wp/2017/05/01/ethiopia-is-facing-a-killer-drought-but-its-going-almost-unnoticed/?utm_term=.ad44d46b0631.

Thenkabail, Prasad S, Chandrashekhar M Biradar, Praveen Noojipady, Venkateswarlu Dheeravath, Yuanjie Li, Manohar Velpuri, Muralikrishna Gumma, et al. 2009. “Global Irrigated Area Map (Giam), Derived from Remote Sensing, for the End of the Last Millennium.” *International Journal of Remote Sensing* 30 (14). Taylor & Francis: 3679–3733.

Tucker, CJ, and PJ Sellers. 1986. “Satellite Remote Sensing of Primary Production.” *International Journal of Remote Sensing* 7 (11). Taylor & Francis: 1395–1416.

2017. *Data Portals*. FEWS NET. <https://earlywarning.usgs.gov/fews/downloads/index.php>.