Field-scale soil moisture bridges the ~~spatial-scale~~ spatial scale gap between drought monitoring and agricultural yields

Noemi Vergopolan[1], Sitian Xiong[2], Lyndon Estes[2], Niko Wanders[3], Nathaniel W. Chaney[4], Eric F. Wood[1], Megan Konar[5], Kelly Caylor[6,7], Hylke E. Beck[1], Nicolas Gatti[8], Tom Evans[9], and Justin Sheffield[10]

[1]Civil and Environmental Engineering Department, Princeton University, USA
[2]School of Geography, Clark University, USA
[3]Department of Physical Geography, Faculty of Geosciences, Utrecht University, The Netherlands
[4]Department of Civil and Environmental Engineering, Duke University, USA
[5]Civil and Environmental Engineering Department, University of Illinois at Urbana-Champaign, USA
[6]Department of Geography, University of California, Santa Barbara, USA
[7]Bren School of Environmental Science and Management, University of California, Santa Barbara, USA
[8]Department of Agricultural and Consumer Economics, University of Illinois at Urbana-Champaign, USA
[9]School of Geography, Development and Environment, University of Arizona, USA
[10]School of Geography and Environmental Science, University of Southampton, Southampton, UK

Correspondence: Noemi Vergopolan (noemi.v.rocha@gmail.com)

Abstract. Soil moisture is highly variable in space ~~, and its~~ and time, and deficits (i.e., droughts) ~~plays~~ play an important role in modulating crop yields~~and its variability across landscapes~~. Limited hydroclimate and yield data, however, hampers drought impact monitoring and assessment at the ~~farmer~~ farm field-scale. This study demonstrates the potential of using field-scale soil moisture simulations to ~~advance~~ support high-resolution agricultural yield prediction and drought monitoring at the smallholder farm field-scale. We present a multi-scale modeling approach that combines HydroBlocks ~~,~~— a physically-based hyper-resolution Land Surface Model (LSM) ~~, and~~— with machine learning. We used HydroBlocks to simulate root zone soil moisture and soil temperature in Zambia at 3-hourly 30-m resolution. These simulations, along with remotely sensed vegetation indices, meteorological ~~conditions, and data describing the physical properties of the landscape (~~data, and descriptors of the physical landscape (related to topography, land cover, ~~soil properties~~and soils) were combined with district-level maize data to train a random forest ~~model~~ (RF) model to predict maize yields at ~~the district- and field-scale (250-m)levels~~district and field scales (250 m). Our model predicted yields with ~~a coefficient of variation (R^2) of 0.61,~~ an average testing coefficient of determination (R^2) of 0.57 and Mean Absolute Error (MAE) of ~~349~~ 310 $\mathrm{kg\,ha^{-1}}$ ~~, and mean normalized error of 22%. We captured~~ using year-based cross-validation. Our predicted maize losses due to the 2015/2016 El Niño drought ~~at similar levels to~~ agreed well with losses reported by the Food and Agriculture Organization (FAO). Our results ~~revealed~~ reveal that soil moisture is the strongest and most reliable predictor of maize yield, driving its spatial and temporal variability. ~~Consequently, soil~~ Soil moisture was also ~~the most~~ a more effective indicator of drought impacts ~~in crops when compared with~~ on crops than precipitation, soil and air temperatures, and ~~remotely-sensed~~ remotely sensed NDVI-based drought indices. ~~By combining field-scale root zone soil moisture estimates with observed maize yield data, this research~~ This study demonstrates how field-scale modeling can help bridge the spatial scale ~~discontinuity~~ gap between drought monitoring and agricultural impacts.

1 Introduction

Droughts can significantly impact crop production, with implications for food security, particularly in smallholder farming systems (Kristjanson et al., 2012; Guilpart et al., 2017). The impacts of droughts ~~to~~ on agricultural production remain difficult to quantify, especially in developing regions, where data is generally scarce and crop production can be highly variable due to ~~high~~ substantial climate variability, highly heterogeneous landscapes, and variable farming capacities (Lobell, 2013; Donaldson and Storeygard, 2016). Challenges in understanding the precise impact of drought on crop yields exist because of the lack of high-quality data and appropriate drought metrics (Sutanto et al., 2019; Beza et al., 2017; Sadri et al., 2020). These data limitations may lead to predictions of yields, drought impacts, and, consequently, agricultural management insights that do not accurately capture the impacts at the farm-level. Improving our understanding of how drought impacts agriculture across spatiotemporal scales would improve the robustness of agricultural drought risk frameworks ~~,~~ and leverage the government's ability to design and implement policies to reduce crop losses.

For example, during the 2015/2016 El Niño, one of the strongest on record (Kintisch, 2016), drought severely impacted Sub-Saharan Africa (SSA). Crop yields dropped 20% in Zambia (Alfani et al., 2019), 63% in Somalia, 50% in Ethiopia, 49% in Zimbabwe, 31% in Swaziland (FAO, 2016a), and 40% in Malawi (FAO, 2016b), leading the region to a state of emergency due to food shortages. Despite the evident severity, in Malawi, for example, satellite-based rainfall drought indices identified that only 21,000 farmers were affected by the drought, while, in reality, survey-based assessments identified that 6.5m farmers were impacted (Economist, 2016). Although rainfall has a historical significant contribution to monitoring droughts and agricultural impacts (Zargar et al., 2011; Hao and Singh, 2015; Van Loon et al., 2016), inconsistencies, as such, emerged because rainfall-based indices do not account for the extreme heat associated with drought. By not accounting for the plant-soil-water dynamics and interactions with the landscape (Peichl et al., 2018; Franz et al., 2020), rainfall-based metrics often do not directly reflect how much water is available to plants. In fact, during the 2015/2016 El Niño, the extreme heat led to insufficient soil moisture in the rooting zone for the plants to meet the higher than normal atmospheric moisture demand, which increased the drought impacts above and beyond the deficit in rainfall supply (Kintisch, 2016; Wanders et al., 2017).

For these reasons, hydrological variables such as soil moisture and evapotranspiration are a more direct proxy of the water available in the root zone to plants. In fact, soil moisture has been shown to better predict agricultural drought impacts than precipitation and air temperature measures (Xia et al., 2014; Bachmair et al., 2016). However, in-situ soil moisture measurements or information on the root zone are virtually ~~non-existent~~ nonexistent in most of the developing world (Karthikeyan et al., 2017). Satellites can provide global information on soil moisture with a 2-3 day revisit time, but they have limited spatial resolution (e.g., 9 km for the NASA SMAP-Enhanced product) and can only measure the upper 5 cm of the soil. Thus, remote sensing based estimates of soil moisture fall short in representing conditions at the scale of agricultural fields ($\sim$ 1-10 ha) and crop rooting zones (10-150 cm). Consequently, few studies have been able to quantify the relationship between field-scale soil moisture deficit (i.e., drought) and crop yield (for a review, see Karthikeyan et al., 2020).

To aid drought assessments, data on crop yields are usually estimated through self-reported field surveys. However, these are time-consuming, expensive, and often suffer from sampling and reporting errors (Paliwal and Jain, 2020; Gourlay et al.,

2019). To compensate for these errors, survey data are generally aggregated to the scale of administrative units, which masks the

55 heterogeneity of yields that exists across small-scale farms (0-5 ha; Jayne et al., 2016). Previous studies indicate that there is a large variability between and within fields, which is substantially masked by aggregation (e.g., Lobell et al., 2007; Franz et al., 2020). This variability is due to spatiotemporal variations in weather (which occurs at kilometer scales or finer), diversity in farm management strategies, and the spatial variability in the landscape (including topography and soils; that can act at the meter scale). These spatiotemporal variations propagate into small scale variations in hydrological variables and fluxes, such as

60 soil moisture and evapotranspiration (Crow et al., 2012; Chaney et al., 2018). Variations in planting date, cultivar choice, and fertilizer/pesticide applications also create inter-field yield heterogeneity for fields with similar environmental attributes. It is, therefore, difficult to interpret spatially aggregated yields because they average out important aspects of the spatial variability in the underlying data. Not knowing the field-scale yield and drought impacts variability also complicates how drought policies are designed and implemented, especially because individual fields and farmers may respond differently and require different

65 interventions during a drought. Thus, characterizing the spatiotemporal dynamics of agricultural yields and droughts at the farm scale (1–250 m resolution) is critical to better understand the field-scale circumstances and to better guide on-the-ground interventions.

There is a long and diverse legacy of attempts to develop models that can predict how agricultural yields respond to weather and climate. These include both process-based (e.g., Jones et al., 2003; Keating et al., 2003) and empirical approaches based on

70 statistical (Lobell and Burke, 2010) and machine learning methods (Chlingaryan et al., 2018). These approaches are mostly based on predictors related to precipitation, temperature, and satellite-derived vegetation indices (VIs), which can help resolve the spatiotemporal variability in yields but are only partially correlated with actual yields (e.g., Lobell et al., 2007; Enenkel et al., 2018). Ideally, vegetation greenness can capture the combined influence of hydroclimatic variability (Koster et al., 2014; Adegoke and Carleton, 2002) and agricultural management activities (e.g., irrigation and fertilization Deines et al., 2017; Estel

75 et al., 2016; Chen et al., 2018). However, VIs are derived from visible-infrared satellite sensors that are impacted by a number of factors that can undermine yield estimates, such as long revisit times (1–2 weeks), cloud contamination, and saturation at high values (e.g. NDVI, Azzari et al., 2017; Gu et al., 2013), which limits its application.

In this study, we present a multi-scale framework that combines hyper-resolution land surface modeling and machine learning to obtain field-scale maize yield estimates and gain insight into the relationships between drought indices and yield variability.

80 Specifically, we used the HydroBlocks land surface model (Chaney et al., 2016; Vergopolan et al., 2020) to simulate root zone soil moisture and surface temperature at a high spatial and temporal resolution (30-m, 3-h interval) over a long duration (1981–2018). We combine these field-scale measures with meteorological variables, remotely sensed vegetation indices, and several other socioeconomic and physical measures with a random forest model (RF; Breiman, 2001) to predict annual maize yields at both district and field scales for Zambia (750,000 km^2), a Southern African country that is exposed to substantial

85 climate variability and where much of the population still depends on small-scale agriculture (Zhao et al., 2018). We use this modeling framework to answer the following questions:

(i) What are the most influential drivers of maize yield variability, and how do hydrological- versus meteorological-based predictors contribute to yield predictions?

3

(ii) What is the field-scale variability of the predicted yields?

90 (iii) How and what drought conditions lead to yields losses at the field-scale?

This study shows the critical role of soil moisture in modulating maize yields, outperforming precipitation, temperature, and vegetation index predictors. We demonstrate how droughts can impact yields differently across the landscape, and how field-scale soil moisture percentiles can effectively capture drought-associated crop losses.

2 Data and Methods

95 ### 2.1 Study area

Our study focuses on Zambia, which is broadly representative of the smallholder-dominated farming systems in many of Africa's savanna regions, which are also beginning to undergo a period of rapid development in which agriculture will play a key part (Searchinger et al., 2015). Savannas drylands are characterized by strong rainfall seasonality and often high inter- and intra-seasonal rainfall variability, which has important consequences for food security (Lehmann and Parr, 2016; Scanlon et al.,

100 2005; D'Odorico and Bhattachan, 2012). Zambia's annual precipitation ranges from 1400 mm in the north to 700 mm in the south, with an annual mean air temperature of 20 °C that rises to 25 to 30 °C during the growing season. The wet season is generally from October to April and the dry season is from June to September, with the date of rainy season onset earlier in the northern part of the country than in the south. The sowing period extends approximately from October to December, the growing period extends from November to May, and the harvesting season extends from April to June (Waldman et al., 2019). In 2013,

105 Zambia consisted of 72 districts (118 districts in 2020 after subdivision of some districts) with an average area of 10,450 km^2 and an average agricultural area of 3,310 km^2 per district. Figure S1 in the Supplemental Information (SI) shows the districts and land cover. While 35.8% of the Zambian agricultural area is of small-size farms (0–5 ha) and 53.0% is of medium-size farms (5–100 ha), 78.8% of the land is owned by smallholder farmers with farms of size 0–5 ha (Jayne et al., 2016). Farming is the primary livelihood activity for 85% of the population, as is the case with many other SSA countries (GYGA, 2020). Irrigation

110 systems are mostly absent in the small-scale farming sector, with agriculture heavily relying on rainfall (Mason and Myers, 2013). Maize is Zambia's key commodity~~with a~~, according to the PHS database, accounting for 60% of the agricultural area with a of 1,661,389 ha. Zambia has a potential yield of 12000 kg ha^{-1}, the same as in the United States, but with actual yields of, in average, 1600 kg ha^{-1} (GYGA, 2020).

2.2 Hyper-resolution land surface modeling

115 HydroBlocks is a field-scale Land Surface Model (LSM) that considers high-resolution ancillary datasets (soil properties, topography, and land cover at 30–250 m resolution) as drivers of landscape spatial heterogeneity (Chaney et al., 2016). HydroBlocks leverages the repeating patterns that exist over the landscape (i.e., the spatial organization) by clustering areas of assumed similar hydrologic behavior into Hydrological Response Units (HRUs). The identification of these HRUs and their spatial interactions allows the modeling of hydrological, geophysical, and biophysical processes at the field-scale over regional

to continental extents. The core of HydroBlocks is the Noah-MP LSM (Niu et al., 2011) single-column land surface scheme. HydroBlocks applies Noah-MP in an HRU framework to explicitly represent the spatial heterogeneity of surface processes down to field scale. At each time step, the land surface scheme updates the hydrological states at each HRU; and the HRUs dynamically interact laterally via subsurface flow. HydroBlocks implements a multi-scale hierarchical scheme that operates at several spatial scales identified for the underlying hydrological, geophysical, and biophysical processes (Chaney et al., 2018):

(a) Catchments: defined by topography and serve as the boundary for surface flows;

(b) Characteristic hillslopes: defined by topography and environmental similarity;

(c) Height bands: defined by the height above nearest drainage and define the primary flow directions and surface temperature gradient;

(d) HRUs: defined by multiple soil/vegetation/land cover characteristics and serve as the smallest modeling units.

With this hierarchical setup, HydroBlocks handles mass/energy exchanges within a modeling unit (at a certain scale) separately from the exchanges between units at that scale. This enables full and realistic horizontal coupling while ensuring computational efficiency.

We deployed the HydroBlocks to simulate root zone soil moisture and soil temperature from surface to 1.5 m depth at 3-h 30-m resolution between 1981-2018. As data inputs, we used hourly 9-km meteorological inputs from ERA5-Land (rainfall, 2-m air temperature, longwave radiation, shortwave radiation, wind, surface pressure, and specific humidity derived from dew point). ERA5-Land is a state-of-the-art global reanalysis product that was chosen because of its high spatial resolution and overall good performance in representing rainfall and soil moisture dynamics (Beck et al., 2020). To parameterize the HydroBlocks, we also used 30-m topography (SRTM; Farr et al., 2007); 20-m land cover type (ESA-CCI); and 250-m soil properties (SoilGrids; Hengl et al., 2017) to derive the soil hydraulic parameters via pedotransfer functions. Although not calibrated, previous validation work compared HydroBlocks soil moisture simulations with in-situ observations over similar climates, and results showed that the model represents the spatial and temporal variability of in-situ soil moisture measurement as field-scale and regional scale, with comparable performance to satellite estimates (Vergopolan et al., 2020; Cai et al., 2017), providing reasonable confidence in the variability of the yield estimates derived in this study.

2.3 Modeling maize yields at district-scale and mapping yields at field-scale

To model and predict maize yield dynamics at the field-scale, we trained a Random Forests (RF) model on district-level survey-based maize yield data, high-resolution HydroBlocks simulations, remotely sensed vegetation index, meteorological reanalysis data, and other static information on the landscape. Section 2.3.1 presents the datasets and predictors descriptions. Section 2.3.2 presents the RF model setup and evaluation. Section 2.3.3 presents the recursive feature elimination approach applied to select and rank predictors. Lastly, Section 2.3.4 presents how the RF model was deployed to predict annual maize yields in Zambia at the field-level, and the analysis performed to assess the yields spatial variability.

Table 1. List of the datasets and the predictors derived to train the random forest model

Dataset	References	Predictors description	Type
Land Cover	20-m ESA-CCI (2016)	The land cover classification was used to identify cropland areas (assuming that all cropland produce maize) and the percentage coverage of non-agricultural cover types in 250-m pixels.	Static
Topography	30-m SRTM (Farr et al., 2007)	From elevation we derived slope, aspect, topographic index, and the height above the nearest drainage.	Static
Soil properties	250-m SoilGrids (Hengl et al., 2017)	Clay, sand, silt, and organic matter content, and the following variables derived using pedo-transfer functions (Saxton and Rawls, 2006): soil saturation, field capacity, wilting point, soil water storage, and hydraulic conductivity calculated using pedotransfer functions.	Static
Climatology: – Precipitation – Air temperature	1-km WorldClim2 (Fick and Hijmans, 2017)	Climatology of air temperature and precipitation during the growing season.	Static
Socioeconomic: – GDP – Population	1-km 2010 GDP (Bank, 2012) & 2010 population (CIESIN, 2017)	GDP and population density as proxies for access to finance, technology, and infrastructure.	Static
Vegetation index: – NDVI	250-m MODIS MOD13Q1 (2000–2018)	Maximum growing season NDVI, date of maximum NDVI, and seasonal NDVI integrals calculated by using a smoothing function to fill the missing NDVI values and to remove outliers followed by a Savitzky-Golay (SG) filter (Chen et al., 2004).	Dynamic
Meteorological: – Precipitation – Air temperature	9-km monthly ERA5-Land (C3S, 2019)	Growing season and the monthly average, minimum, and maximum estimates of air temperature and precipitation.	Dynamic
Hydrological: – Soil moisture – Soil temperature	30-m monthly HydroBlocks (this study)	Seasonal and monthly average, minimum, and maximum root zone soil moisture, relative root zone soil moisture, and soil temperature from the HydroBlocks LSM.	Dynamic

2.3.1 Training data and feature engineering

To train the RF model, we used the Post-Harvest Surveys (PHS) database of maize yields (km^2). This database comprises household survey data of ∼13,000 farmers' self-reported harvested maize (kg or total bags of crop) and respective cultivated area (ha). The data was collected at the end of each harvesting season by Zambia's Central Statistical Office (a division of the Ministry of Agriculture and Livestock) from 1991 to 2005, 2007, 2008, 2011 to 2014. Due to privacy and data uncertainties, the observations were only available aggregated to the district-level (∼10,000 km^2). In this work, to match the period of availability for remotely sensed predictors, we used the PHS data from 2000–2018. This resulted in a total of 527 observations from 70 districts and 8 years. ~~The data were randomly split, with 80% used for training and 20% used for model evaluation.~~ To train and evaluate the model we used a year-based cross-validation approach. This approach relies on selecting a given year to evaluate the model and train it on all the other years. This is performed for each of the 8 years independently, with the average statistics representing the overall cross-validation performance. To train the model, we used a range of static and dynamic variables with time, as described in Table 1.

From these static and dynamic predictors, we identified 103 initial predictors. One of the challenges of working with multi-source and multi-scale datasets is the spatial scale mismatch of input data (20-m to 9-km resolution). To ~~calculate the value of~~ reconcile these different scales and calculate a district-level value for each predictor, we masked out the non-cropland pixels and calculated ~~the~~ its area-weighted average based on the cropland location/area in each district. Despite the large district

areas, the use of these area-weighted averages at only the agricultural areas helped to remove the influence from the surrounding non-cropland areas (i.e., grasslands, water bodies, urban areas), while accounting for the spatial variability of each predictor in the district. Lastly, each predictor was normalized based on the maximum and minimum values. We used NDVI retrievals from MODIS instead of Landsat because of the lower cloud coverage and a shorter revisit time. In Zambia, cloud coverage between December to February was on average 50% for Landsat and 30% for MODIS.

2.3.2 Random forest regressor for yield modeling

Machine learning models have been widely applied for crop yield prediction (Chlingaryan et al., 2018). Random Forest (RF) regressors are used with geospatial hydroclimate and satellite data to predict maize yields at fine-scales (Aghighi et al., 2018; Khanal et al., 2018; Jeong et al., 2016; Folberth et al., 2019). Advantages of RF regression models are their high predictive accuracy even when trained on small, nonlinear, and collinear datasets, their robustness to outliers, and their ability to avoid overfitting (Breiman, 2001; Archer and Kimes, 2008; Wylie et al., 2019). Dealing with data collinearity is particularly important for yield prediction because often the meteorological-, hydrological-, and vegetation-based predictors are interconnected (Archer and Kimes, 2008; Wylie et al., 2019).

To identify the optimal RF architecture, we performed a grid-search on the possible combinations of relevant hyper-parameters: number of trees, maximum depth of the trees, the minimum number of samples required to split an internal node, the minimum number of samples required to be at a leaf node, and the number of bootstraps of predictors. The best hyper-parameter combination was selected based on the average Mean Square Error (MSE) of a 3 fold cross-validation on the training sample. We further evaluate the overall model's performance by calculating the Mean Absolute Error (MAE) and the Coefficient of Determination (R^2) performance ~~with the testing sample (20% of the data)~~using a year-based cross-validation approach.

2.3.3 Predictor importance and selection

Using the district-level PHS yield data and the RF model, we performed a Recursive Feature Elimination (RFE) analysis to (i) identify and rank the most important predictors of maize yields, and (ii) to determine which of the predictors could be removed. Removing non-predictive variables is particularly helpful as it can improve model accuracy, it mitigates the model's tendency to overfit, while the smaller data volume reduces the computational cost.

The RFE is an iterative process. At each iteration, the model is trained, the importance of each predictor calculated and ranked, and the least important predictor is removed (Gregorutti et al., 2016). This process continues until a convergence criterion is met. In our implementation, as an evaluation metric, we used the average maize yield R^2 ~~between the RF-predicted yields and the testing set as the evaluation metric~~of year-based cross-validation. The importance of each predictor was calculated based on the difference between the R^2 of the model with the predictor and of the model without the predictor. This iterative process of retraining and assessing the relative importance of each predictor ensures that the least important predictor is consistently removed and that discarded variables either do not contribute to or degrade model performance. Variable removal ceased once this R^2 difference fell below 0.001 (see Table S1 for a full description of the RFE process). The RFE process applied to the 103 predictors ~~in the training set~~ (section 2.3.1) resulted in the retention of ~~36~~27 important predictors.

7

In addition to variable selection, we performed a sensitivity analysis to quantify the relative value of hydrology-based predictors (soil moisture and soil temperature) and meteorology-based predictors (precipitation and air temperature) in predicting district-level maize yields. For this experiment, we trained the RF model, applied RFE, and identified the most important predictors of three different predictor sets, as shown in Table 2. For each case, we quantified the final R^2 ~~performance of the testing sample~~ year-based cross-validation performance and the delta R^2 importance of each predictor. We then compared the change in performance with and without meteorology- and hydrology-based predictors with respect to the benchmark. We also compared the relative importance of different predictors in each of the cases.

Table 2. Predictor sets applied to the maize yield modeling sensitivity experiments

Case	Name	Description	Total predictors
1	All predictors	All the predictors as a benchmark	103
2	No meteorology	All the predictors except the precipitation and air temperature predictors	70
3	No hydrology	All the predictors except the soil moisture and soil temperature predictors	61

2.3.4 Predicting maize yield at the field-scale

Previous work showed that RF models are able to successfully model fine-scale crop yields from coarse-scale physically-based crop models estimates, assuming that the fine-scale predictors are representative of the fine-scale yield variability (Folberth et al., 2019). Similarly, we deployed the trained RF model to predict maize yields at 250-m resolution, considering that predictors and model parameters that represent the yield spatiotemporal dynamics at the district-scale can represent yield dynamics at the field-scale.

To this aim, we first identified all 250-m grid cells that have at least 50% cropland coverage according to our fractional land cover map derived from the ESA-CCI data (Table 1, Figure S1), which resulted in ~1M grid cells. Since maize is Zambia's dominant crop (Ng'ombe, 2017), and it accouts for 60% of the planted area, we assume that all the identified cropland areas are maize fields. Similarly to ~~the training and testing~~ predictor set, we used the location/area of the cropland fields to calculate the area-weighted average of each predictor. We used our best RF model to predict the annual maize yield at each 250-m grid cell over Zambia between 2000 to 2018.

To characterize the spatiotemporal distribution of field-scale maize yields estimates derived from the RF model we generated maps of mean annual yields and their temporal coefficient of variation. We also calculated maps of maize yield trends between 2000–2018 to identify the locations where increases in yields were larger. As an example, to quantify the spatial distribution of losses in yield due to drought, we calculated the relative change in yields for the 2015/2016 El Niño season and compared it to FAO survey estimates (Alfani et al., 2019). Despite the sampling difference between the surveyed and predicted field-scale yields, we evaluate the estimated field-scale yields by comparing the field-scale yields aggregated to the district level with district-level PHS survey yield data. We computed the temporal R^2 and MAE, and the mean spatial Pearson correlation.

Table 3. List of the six drought impact indicators used in the crop losses analysis. Three are the most predictive variables (identified by the predictor importance analysis), and three are respectively commonly used indicators.

Indicator conditions	Variables	
	Most predictive	Commonly used
Dry/wet	Root zone soil moisture (Apr.)	Precipitation (Apr.)
Cool/hot	Max. soil temperature (Oct.)	Max. air temperature (Oct.)
Plant health	Date of NDVI peak (season)	NDVI Integral (season)

Spatial variability of field-scale maize yields and main predictors

The spatial patterns and spatial variability in maize yields can be driven by hydroclimatic conditions, soil properties, topography, and also farmer management (such planting date, seed choice, use of fertilizers, irrigation, etc.). Consequently, droughts can impact the landscape differently. To quantify the strength of each predictor in driving the spatial variability in maize yields at the local scale, we calculated the spatial Pearson correlation between the field-scale yields and the most important predictors. The time series of spatial correlation was calculated for each year over the entire country, as well as for three smaller domains (of 50-km by 50-km) in the north, central, and south of the country.

2.4 Characterizing the relationship between field-scale yields and drought indicators

The predictor importance analysis (section 2.3.3) identified the most influential variables impacting maize yields at the district-scale. To gain further insight into the potential effectiveness of these variables as drought impact indicators, we compared how they varied with the spatially co-located de-trended maize yields, hereafter called anomalies. This allows us to characterize the relationship between these indicators (e.g., dry and hot) versus local losses or gains in maize yields.

The drought indicator variables, shown in Table 3, were identified based on the three most important predictors (result from section 3.2) and three (respective) commonly used indicators. Temperatures and date of NDVI peak were used in their original units. For NDVI integrals, for each grid, we used the temporal anomaly relative to the 2000-2018 mean. For soil moisture and precipitation, drought indices were constructed from monthly values and converted to the empirical percentile. The empirical percentile was calculated for each time step and grid cell and based on the monthly historical records (1981–2018) for that location/month (Sheffield, 2004). By using percentiles, rather than absolute values, drought events can be compared in space and time, despite their local characteristics.

By comparing how each drought indicator varied with the yield anomalies, this approach allowed us to characterize the relationship between these indicators versus local losses (or gains) in maize yields. We delineated potential drought thresholds for soil moisture and precipitation percentiles and quantified the mean expected yield losses associated with each threshold. Second, we compared how each of the drought indicators co-varied with each other and with maize yield anomalies. This allowed us to quantify what are the expected yield losses (or gains) under dry and hot versus wet and hot conditions, and to identify in which conditions yield losses are driven by water deficits and/or temperature stress.

3 Results

3.1 Hydrological simulations at field-scale

HydroBlocks-simulated root zone soil moisture and soil temperature at 30-m resolution reveal substantial spatial variability at both the national and local scales. Figure 1 shows the mean April root zone soil moisture and October soil temperature
255 since these variables are the most important predictors of yield (section 3.2). At the national scale, wet and cooler conditions are observed in the northern parts of the country and near river valleys, while the south and southeast show distinctly dry and hot conditions. National and local scale variations (Figure 1 inset) reflect the interactions of meteorological conditions with the landscape, highlighting the influence of topography, soil properties, and vegetation cover on the spatial variability of hydrological processes.

3.2 District-level maize yield modeling and importance of predictors

The best-performing RF model was able to predict the values of the district-level yield data ~~in the independent testing data set with a~~ with an average Mean Absolute Error (MAE) of ~~349~~ 310 $\mathrm{kg\,ha^{-1}}$ and an R^2 of ~~0.61, with an MAE of 253 and R^2 of 0.79 for the training set (Figure 2~~ 0.57 (year-based cross-validation). Figure 2 shows the comparison between observed and predicted yields for each year (with the model trained on data of all the other years). The model captures the overall patterns of
265 district-level maize yields, but with a tendency to underestimate maize yields above ~~2000~~ 2500 $\mathrm{kg\,ha^{-1}}$ ~~.~~ in the recent years (e.g. 2011—2014), and overestimate low yields in the earlier years (e.g. 2001—2004).

The recursive feature elimination process combined with the sensitivity analysis (Table 2) revealed that soil moisture and soil temperature variables were consistently the strongest predictors of yield (Figure 3). In the first two sensitivity tests, in which Hydroblocks variables were included, soil moisture and temperature variables provided 5 or 6 of the top 7 predictors,
270 while removing soil moisture and temperature variables as predictors (case 3) resulted in the largest drop in overall model R^2 (~~0.09 and 0.11~~ 0.06 and 0.08). On the other hand, removing meteorological variables resulted in relatively little loss of model explanatory power, with an R^2 drop of only 0.02 between cases 1 and 2.

In terms of specific variables, cases 1 and 2 both showed that the mean relative soil moisture in April was the strongest predictor of yield, followed by precipitation climatology, October and February soil temperatures, and then March soil moisture
275 (Figure 3). Two static variables appear to have some strong influence as well, ~~particularly~~ shrubland percentage, which is ranked third in Case ~~1 and second in Case 2. Precipitation~~ 2, and precipitation climatology, which represents the spatial distribution of rainfall, ~~is the other static variable with substantial influence, appearing fourth~~ appearing second in Case 1 and accounting for an R^2 drop of 0.02 if removed from the model. Besides this variable, no other meteorologically-derived variable, including all the dynamic measures, ranks highly in the presence of HydroBlocks variables. The highest of these is December ~~minimum~~
280 mean air temperature, which ranks ~~19~~ 10th in Case 1. Removing HydroBlocks variables (Case 3) increases the influence of these variables, but they still remain less predictive than the static precipitation climat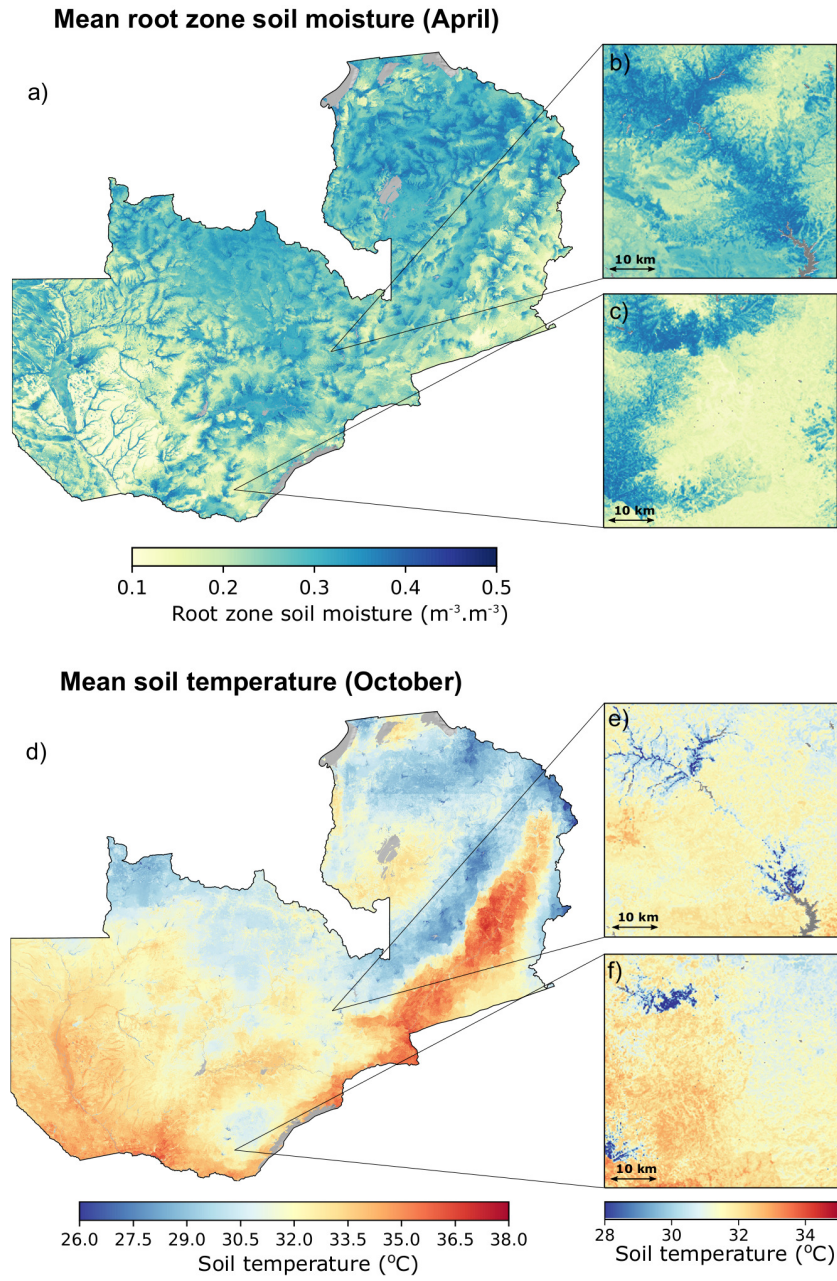ology ~~and shrubland percentage (ranked second and third)~~, while the dynamic vegetation measures, date of maximum NDVI and the corresponding value are respectively the

10

Figure 1. April root zone soil moisture and October soil temperature were the most predictive variables in the RF model. Figures a) and b) shows respective mean values (2000–2018), at the 30-m spatial resolution, as simulated by HydroBlocks land surface model. Large lakes are excluded from the simulations (grey areas).
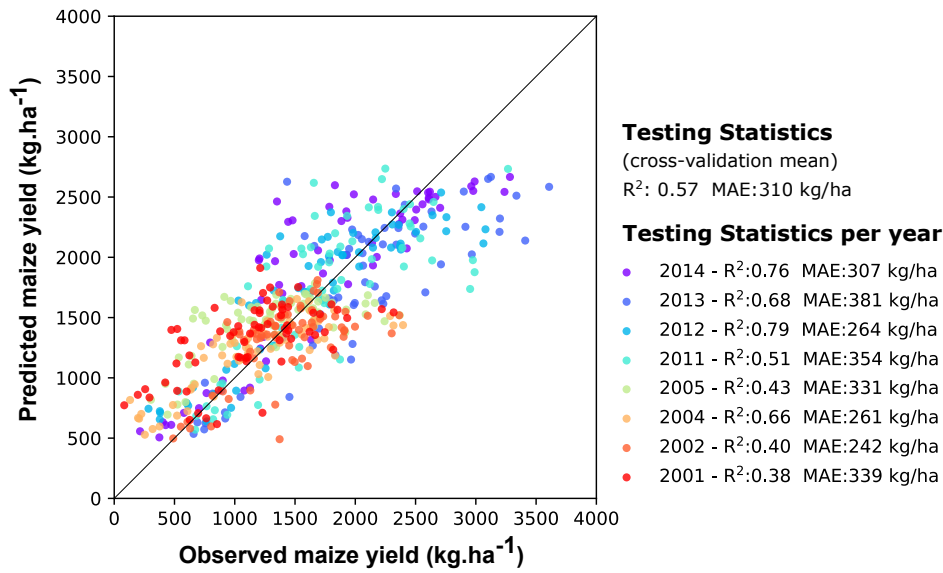
Figure 2. Comparison between observed and predicted maize yields at the district-scale. The predictions were obtained by training a random forest model with survey-based maize yield data and fine-scale geospatial environmental datasets. Each color shows the maize yield evaluation for a given year (trained on data of the other years).

first and ~~fourth~~ sixth most influential. NDVI-derived variables otherwise rank ~~10~~12th ~~and 8th~~ in Cases 1 and were not retained in Case 2.

3.3 Field-scale maize yields for Zambia

At the field-scale, RF-predicted maize yields averaged $1557 \, \mathrm{kg \, ha^{-1}}$ ($\pm 219 \, \mathrm{kg \, ha^{-1}}$) for the 2000–2018 period. As expected, we observe higher average maize yields in the northern parts of the country and average lower yields in the southern regions (Figure 4a), which reflects the spatial distribution of mean rainfall. Figure S6 shows time-series of the predicted yields for different locations across the country. In terms of the temporal coefficient of variation (Figure 4d), yield variability is highest in the central, southern, and eastern parts of the country, and lowest in northern and northwestern Zambia. In general, there is an inverse relationship between mean annual yields and their coefficient of variation (Figure S2), however, there are several notable departures. For example, a landscape-level view in central Zambia reveals areas of moderate to high yields (Figure 4b) with corresponding moderate to high levels of variability (Figure 4f) due to proximity to rivers and extensive floodplains, where inundation is frequent. The opposite pattern is seen in Southern Zambia, where portions of the landscape show low yields and relatively low variability (Figures 4c and 4e) due to more consistent dry conditions.

Analyzing the change in maize yields trends for the period 2000–2018 (Figure 4g, 4h, 4i), we found that on the whole maize productivity increased by $3.5 \, \mathrm{kg \, ha^{-1} \, y^{-1}}$ (± 4.6 ~~kg~~ $\mathrm{kg \, ha^{-1} \, y^{-1}}$). The gain was more prominent in the northern and central
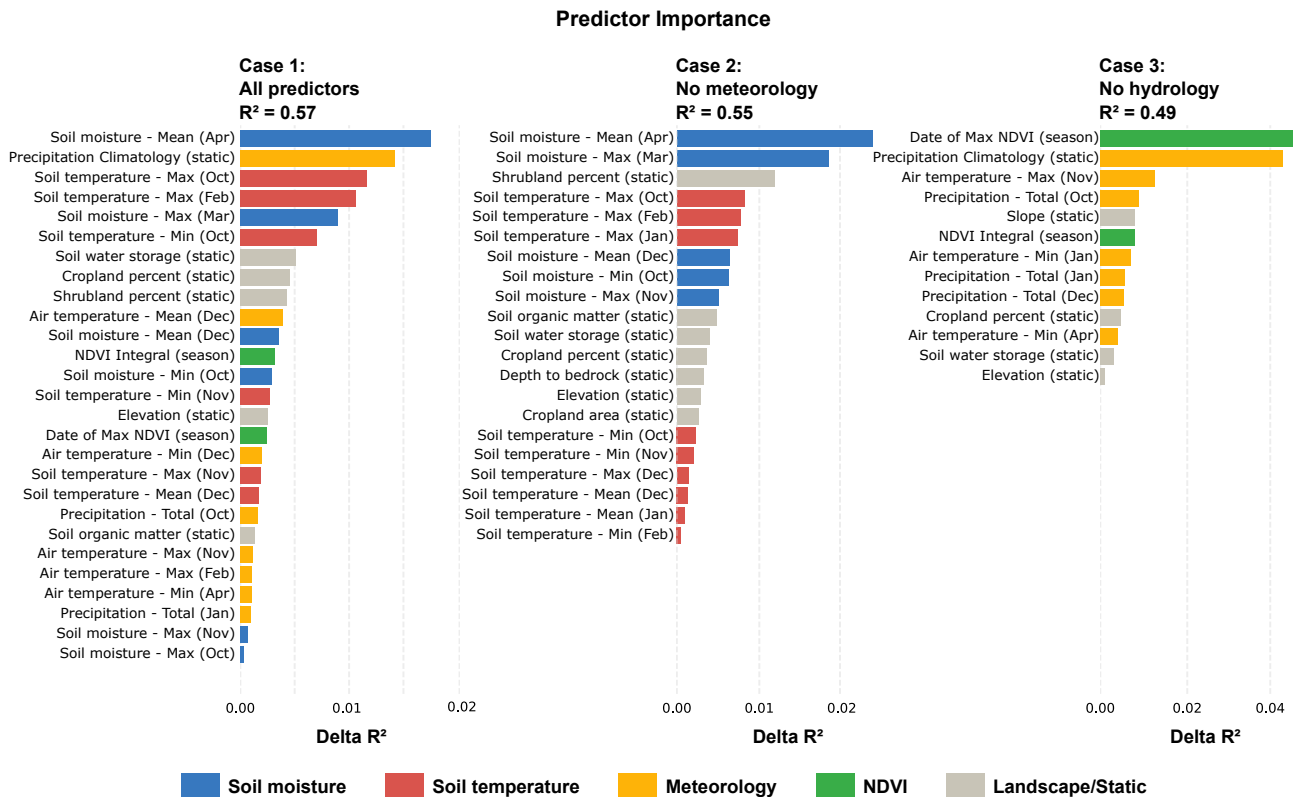
Predictor Importance

Figure 3. The most important predictors for maize yield at the district-scale. The predictors were selected and ranked via Recursive Feature Elimination, with the importance rank shown in terms of delta R^2. Results are shown for case 1 (considering all the variables), case 2 (without precipitation and air temperature predictors), and case 3 (without soil moisture and soil temperature predictors). Each color represents different categories of predictors data.

parts of the country, particularly in floodplains (Figure 4i), rising to 15 ~~kg~~ kg ha^{-1} y^{-1}, while there was little change in southern Zambia. The productivity also tends to be higher at the locations near floodplains (Figure 4i). Nonetheless, the overall yield
300 gains rates observed were far behind than the rates required to match the 12,000 ~~kg~~ kg ha^{-1} Zambia yield potential (Mueller et al., 2012).

During the 2015/2016 El Niño drought, Zambia agricultural production was severely impacted with overall losses across the country. Our field-scale RF-predicted yields were able to capture these losses (Figure 5). The country-wide predicted mean yield for 2015/2016 was 1514 kg ha^{-1} ($\pm$ 233 kg ha^{-1}), which represents an overall loss of 84 kg ha^{-1} ($\pm$ 60 kg ha^{-1}), or 5.3%,
305 relative to 2010–2014 mean. Predicted yield reductions were more severe in southern and southeastern Zambia with losses of 200 kg ha^{-1} (15%) across much of this area, and reaching as high as 522 kg ha^{-1} (28%) (Figure 5). These estimates align with those of the 20% losses estimated by the Food and Agriculture Organization (FAO) for this same area during the 2015/2016 El Niño drought (Alfani et al., 2019). However, when evaluating drought impact at the local scales (Figure 5b, 5c), we observe that
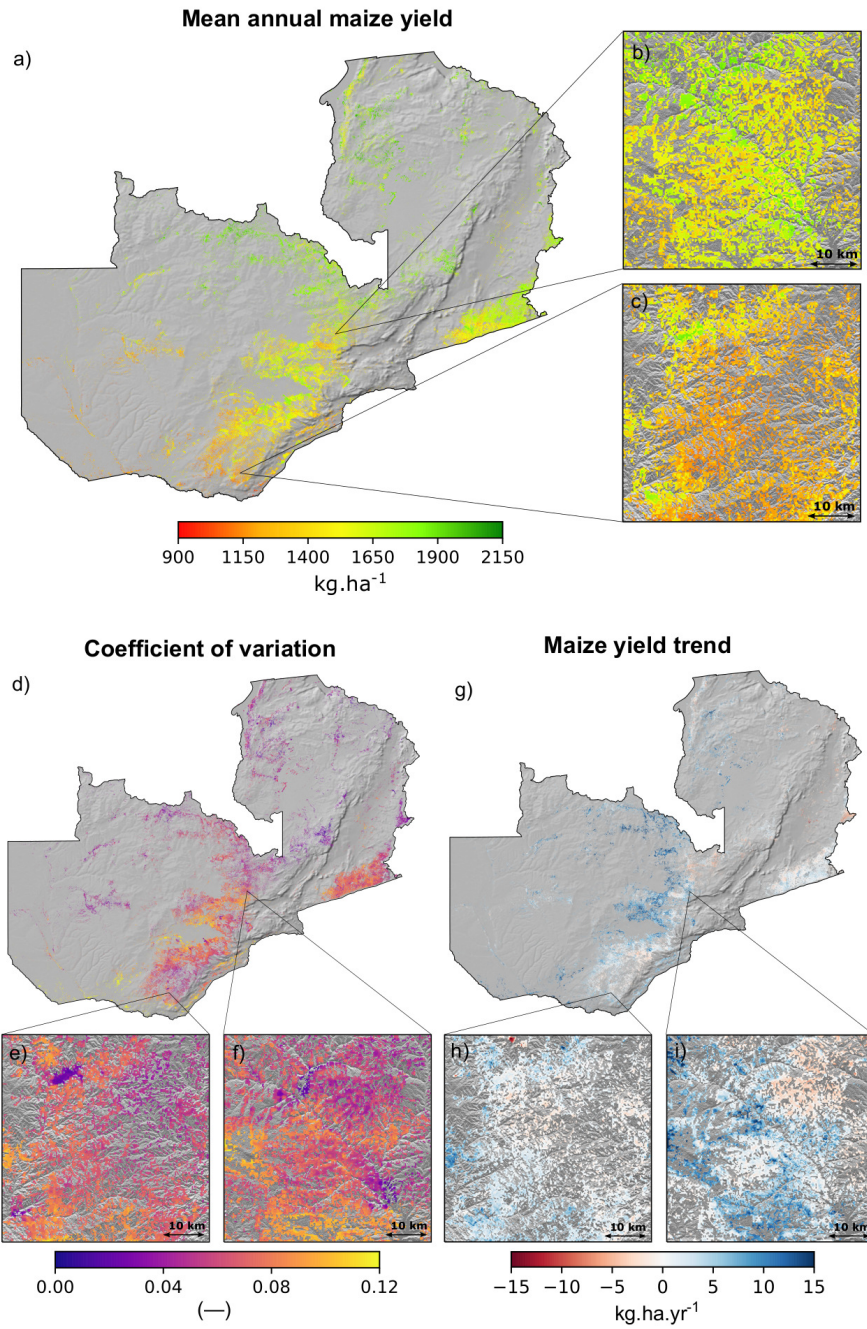
Figure 4. Annual maize yields (a), coefficient of variation (d), and maize yield trends (g) for the period between 2000–2018 estimated using a random forest model. Each zoom panel (b, c, e, f, h, i) shows the respective estimates at a 250-m resolution for a 50-km x 50-km area.
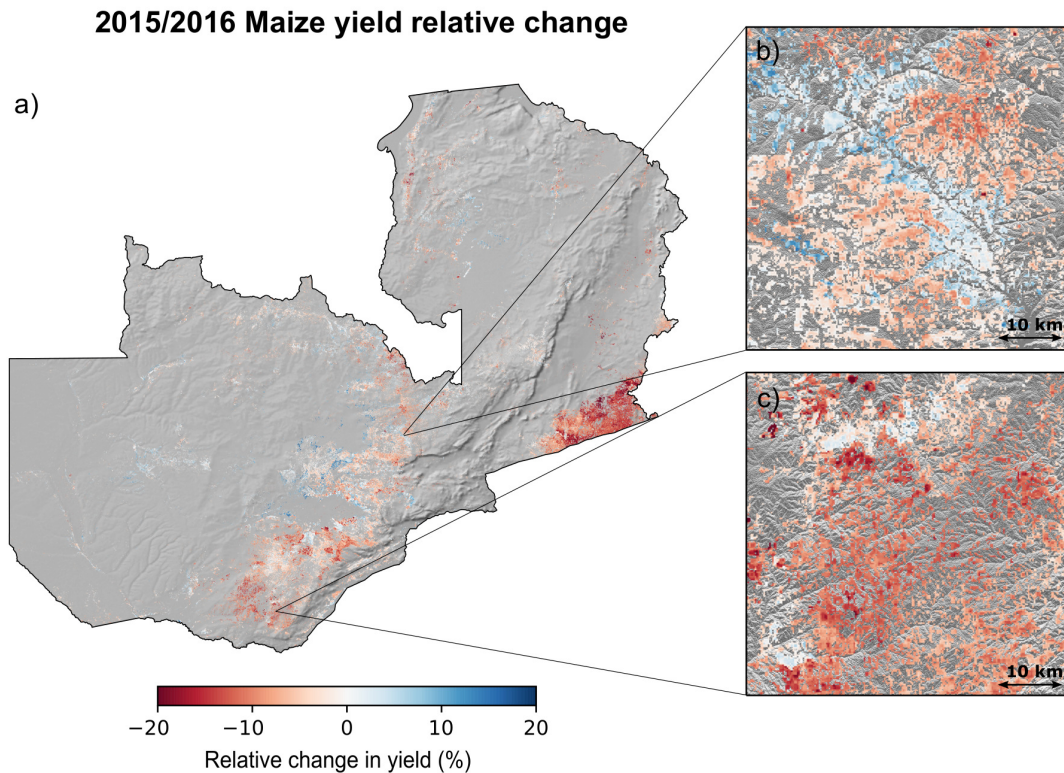
Figure 5. (a) Relative change in maize yields (%) for the 2015/2016 season with respect to 2010–2014 mean. Each zoom panel (b, c) shows the respective data at a 250-m resolution for a 50-km by 50-km area.

drought impacts yield differently across the landscape. Areas nearby to rivers and floodplains were less prone to crop losses (Figure 5b), given its wetter and cooler soil moisture conditions.

We compared the field-scale yield data aggregated to the district level with the PHS survey district-level yield data. We obtained a Pearson correlation of 0.67 (R^2=0.46), and an MAE of 450 $\mathrm{kg\,ha^{-1}}$. In terms of spatial patterns, aggregated field-scale yield and PHS have a spatial correlation of 0.84. This strong spatial relationship is also evident in the PHS and estimated z-scores shown in Figure S3 in the SI. The weak strength and limitations of this aggregated data comparison are discussed in section 4.4

Field-scale spatial variability of yields and dominant predictors

As observed in Figure 5b, droughts impact yields differently across the landscape. To quantify to what extent the spatial variability in the hydroclimate and landscape conditions are driving the spatial variability in the yields, we calculated the spatial correlation between the field-scale yields and the dominant predictors (Figure 6). The spatial correlation was calculated each year, to assess whether these associations changed inter-annually, and over different locations (the entire country, and for 50 x 50-km boxes in southern, central, and northern Zambia) to assess how the relationships change regionally.
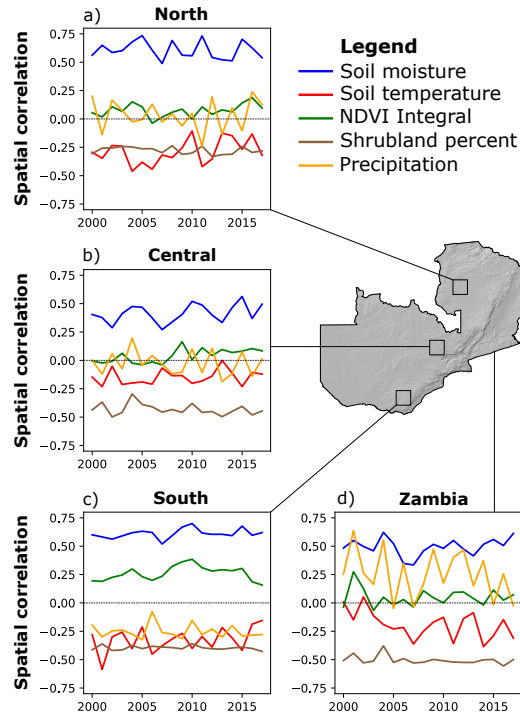
15

Figure 6. Time series of Pearson spatial correlation of maize yields and the most important predictors. Top panels show the spatial correlation within 50 km by 50 km boxes in a) south, b) central, and c) north of Zambia. Panel d) shows the spatial correlation for the entire country.

Soil moisture exhibited the largest impact on the spatial variability in yields, with the highest spatial correlation across all three sub-regions, and the entire country (Figure 6a–d). Soil temperature and shrubland percent are negatively correlated with yield. The relative ranking and temporal dynamics of the spatial correlations are generally consistent across regions, although the strength of correlation varies between regions, with close to zero correlation for most predictors in the central and northern regions. Given its coarser ~~spatial-scale~~ spatial scale and smoother spatial variability, precipitation climatology showed no spatial correlation with the field-scale yield for each region, but strong correlations at the national scale due to the substantial spatial gradient. However, the high inter-annual variability indicates the influence of other factors. With the exception of southern Zambia, NDVI showed a low spatial correlation with yield over time.

3.4 The impact of drought on field-scale maize yields

To examine the effectiveness of the 6 potential drought indicators (Table 3), we evaluated the degree of influence that each indicator had on the anomalies of the predicted field scale yields (Figure 7). Overall the relationship between these indicators and yield anomalies was highly non-linear. Soil moisture and precipitation percentiles capture the largest yield losses of all six indicators (Figure 7a, 7d), with both showing substantial negative anomalies below their 25th percentile values. Yield losses
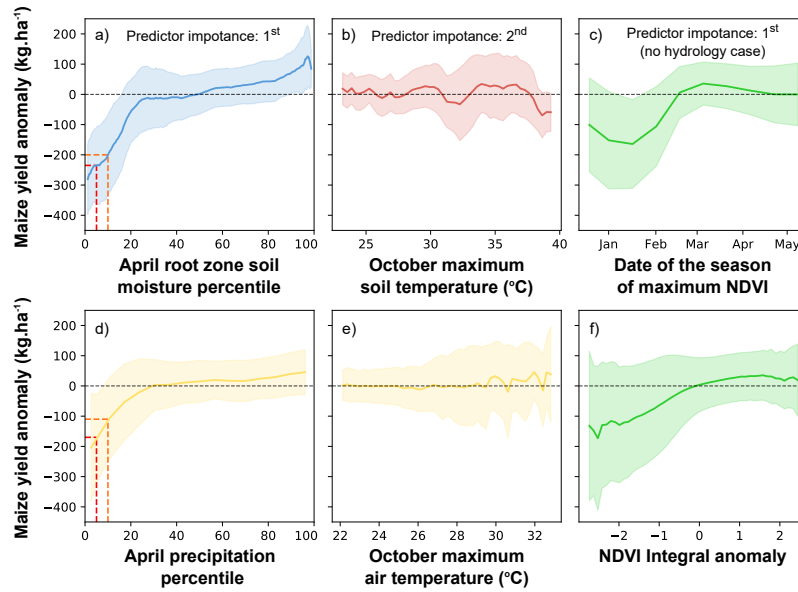
Figure 7. The relationship between the field-scale maize yield anomalies with respect to the six drought indicators identified in Table 3. The top row shows the most predictive variables, the bottom row shows the (respective) commonly used variables. Dark lines show the mean values, shades show the standard deviation. Red and orange dashed lines illustrate potential drought thresholds.

associated with low soil moisture conditions are larger and more certain than those related to low precipitation. For instance,

335   yield losses associated with the 10th soil moisture percentile ($202 \pm 134 \,\mathrm{kg\,ha^{-1}}$) were 89% greater than those related to the 10th precipitation percentile ($107 \pm 135 \,\mathrm{kg\,ha^{-1}}$); orange dashed lines in Figures 7a, 7d. At the 5th percentile (red dashed line), average soil moisture-related yield loss ($235 \pm 128 \,\mathrm{kg\,ha^{-1}}$) was 26% greater than the yield loss associated with the precipitation ($187 \pm 161 \,\mathrm{kg\,ha^{-1}}$), and furthermore has a 20% narrower confidence interval.

Negative NDVI integral anomaly and early NDVI peak date were also associated with yield reductions (Figure 7c, 7f). The

340   strongest and most certain of these was the date of peak NDVI, which resulted in yield anomalies when maximum NDVI occurred before March, with the largest reductions ($164 \pm 146 \,\mathrm{kg\,ha^{-1}}$) occurring for peak dates between January and February. Negative NDVI integral anomalies also showed substantial but more uncertain yield losses. For example, an NDVI integral anomaly of -2 identified a yield loss of $140 \pm 210 \,\mathrm{kg\,ha^{-1}}$.

In contrast to the previous two indicators, there was little variation in yield anomalies associated with soil and air temperature,

345   although the uncertainty in anomalies increased with both measures. However, we observe nearly consistent yield losses with October maximum soil temperatures above 37.5 °C, which is near known critical temperature thresholds for maize (Luo, 2011; Schauberger et al., 2017).

To gain further insight into the conditions (e.g., hot and dry vs. hot and wet) associated with yield losses, ~~we examined in Figure 8 we compared~~ how the yield anomalies co-varied ~~within pairwise comparisons of drought indicators(Figure 8)with~~

17

350 pairwise drought indicators. For example, Figure 8a shows what is the mean yield anomaly associated with its respective October maximum temperature and April root zone soil moisture. As expected from previous findings (Figure 7), soil moisture (Figure 8a, 8c, 8e) and precipitation (Figure 8b, 8d, 8f) percentiles are the dominant influences on maize yield responses. Both indicators show similar patterns, but the yield responses associated with precipitation are noisier. Extreme soil temperature (Figure 8a and 8b) and air temperature (Figure 8c and 8d) only lead to yield losses (red) when the soil moisture and precipitation

355 percentiles are low ($< 25^{th}$). High temperatures under wet conditions (top right corners, Figures 8a-d) show increased productivity (blue). Yield losses occur for the full range of NDVI peak dates when soil moisture and precipitation percentiles are below $10–15^{th}$, but when peak date is earlier than March yield losses occur when soil moisture and precipitation fall below their median values (Figure 8g, 8h). NDVI integral anomalies below zero and below-median soil moisture values show a similar relationship with yield (Figure 8e), but this tendency was much weaker when assessed with precipitation (Figure 8f).

360 ## 4   Discussion

4.1   Key findings

We presented a modeling framework that combines physically-based hyper-resolution land surface modeling and machine learning for maize yield prediction at the field-scale. Our work advances on previous approaches by integrating field-scale hydrological variables into yield prediction, and by more effectively quantifying yield sensitivity to drought. Our key findings

365 are:

- **Model skill**: The modeling approach that we developed was able to estimate maize yields at district scale with comparable or higher skill (R^2=~~0.61~~0.57, MAE=~~349~~310 $kg\,ha^{-1}$) compared to state-of-the-art approaches based on mechanistic yield models (e.g., Jin et al., 2017; Azzari et al., 2017) and higher skill than standard empirical approaches based on weather variables or vegetation indices (Estes et al., 2013; Zhao et al., 2018).

370 - **Yield estimates**: The field-scale results showed a mean maize yield of 1557 $kg\,ha^{-1}$ ($\pm$ 219 $kg\,ha^{-1}$) across Zambia, with an overall increasing production trend of 3.5 $kg\,ha^{-1}\,y^{-1}$ ($\pm$ 4.6 $kg\,ha^{-1}\,y^{-1}$) between 2000–2018. The field-scale yields captured maize losses during the 2015/2016 El Niño drought at similar levels to losses reported by the FAO based on actual yield data (Alfani et al., 2019).

- **Drivers of yield**: We identified soil moisture as the main driver of maize yield variability at both the district-scale and

375 field-scale. At the district-scale, soil moisture was followed in importance by soil temperature, shrubland percent coverage, and precipitation climatology. Time-varying meteorological predictors (precipitation and air temperature) played a minor role. NDVI-based predictors only showed meaningful contribution when soil moisture and soil temperature predictors were absent.

- **Drought impacts**: There is a highly non-linear behavior between drought indices and yield losses. However, consistent

380 maize losses are observed when soil moisture or precipitation drop below the 25th percentile. At extreme dry conditions
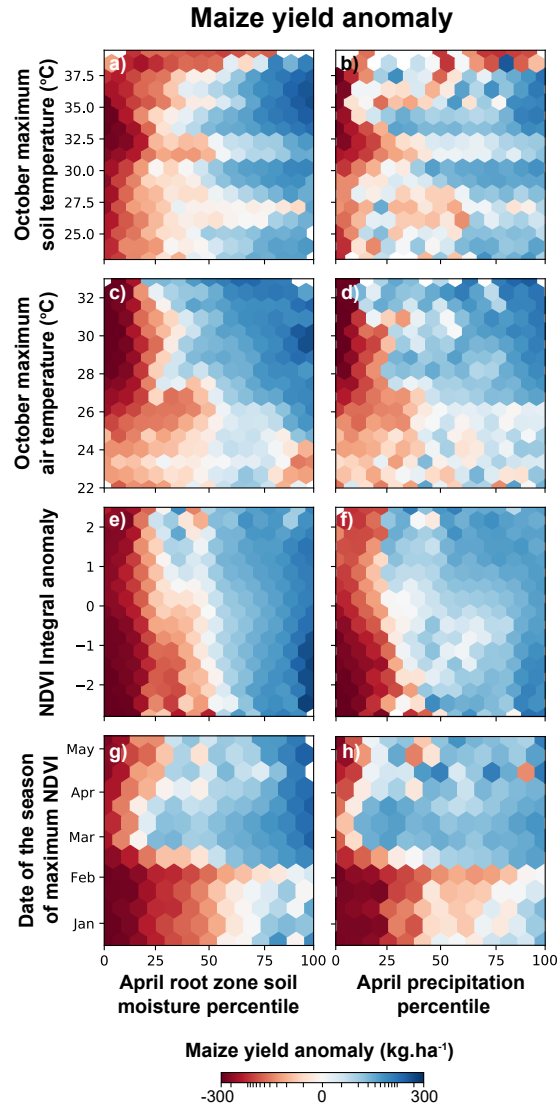
18

Figure 8. Field-scale estimated maize anomalies, as a function of soil moisture and precipitation versus soil temperature, air temperature, date of maximum NDVI, and NDVI integral anomaly. Each hexagon shows the mean yield anomaly, where red (blue) shows overall yield losses (gains). Figure S4 in the Appendix shows the data density of each hexagon.

(5th percentile), soil moisture identifies 26% more losses with 21% less uncertainty than precipitation, providing an effective measure of drought impact. Significant yield losses are also predicted when soil temperature exceeded 37.5 °C in the early growing season. Drought impacted yields differently across the landscape (Figure 5), with most of the spatial variability coming from soil moisture (Figure 6).

4.2 Drivers of maize yields predictability

Soil moisture and precipitation

Soil moisture and soil temperature showed a strong contribution to yield prediction because they represent the water and temperature balances at the root zone, accounting for the soil-water-plant interactions via infiltration, surface and subsurface flow, vertical drainage, water uptake by plants, and evaporation. Consequently, in our sensitivity experiments, in case 1 and case 2 the meteorological predictors added only minor improvements to yield prediction beyond the contribution of root zone predictors. Mainly because precipitation does not necessarily translate into water availability for plants. For example, rainfall from intense storms may contribute mostly to runoff, rather than infiltration, leading to a spatial mismatch between where the water is supplied (precipitation) and where water is finally available to plants (soil moisture). Furthermore, whilst temperature drives the atmospheric demand for evapotranspiration, yet it does not reflect the complexity of controls on transpiration, which are more water-limited (rather than energy limited) in drier regions such as much of Zambia (Berg et al., 2016; Williams et al., 2012; McNaughton and Jarvis, 1991).

Distinctions between atmospheric and ~~root-zone~~ root zone processes are particularly important at the field-scale, due to different natural scales of variability that influence crops in different ways. Precipitation and air temperature operate mostly at larger scales with most of the spatiotemporal variability coming from large-scale atmospheric circulation and large-scale topographic features (Grayson and Blöschl, 2001). Soil moisture, on the other hand, operates at multiple scales, influenced by the meteorological conditions, but with most of the spatial variability coming from the heterogeneity of the landscape (topography, soil properties, land cover) (Crow et al., 2012; Famiglietti et al., 2008; Grayson and Blöschl, 2001). By providing soil moisture and soil temperature estimates at a spatial scale closer to its natural scale of variability, HydroBlocks allows for improved yield predictions. This is indicated by comparison of the joint distribution of soil moisture and precipitation percentiles (Figure 9), which shows that most of the time their condition will be similar (0.78 correlation), yet when they are different (due to the above reasons), soil moisture better captures losses in yield.

April soil moisture ranked as the strongest predictor of yield. April covers the grain filling stage of the maize calendar in Zambia (~35 days prior to maturity; Yonts et al., 2008), when soil moisture is critical for plants because of the large water uptake demands (Yonts et al., 2008; Borras et al., 2003). Despite the high greenness of the season (measured by vegetation indices, such as NDVI, for example), if plants do receive enough water in the reproductive period the cob will not develop well, compromising productivity. Although several studies identified NDVI as the strongest predictors for maize (Table 1 in Funk and Budde, 2009; Johnson, 2016; Petersen, 2018; Karthikeyan et al., 2020), our sensitive analysis results (Figure 3, case 3) showed that only in the absence of soil moisture and soil temperature variables do NDVI-based variables emerge as strong predictors of yield. This is evident in Figure 8e, 8g which shows that NDVI-based anomalies have little sensitivity to change in yields when compared to soil moisture. ~~NDVI limitations are discussed~~ This could also be a limitation of visible satellite sensors in capturing under canopy plant-soil-water dynamics. A further discussion on NDVI limitations is presented in section 4.4.
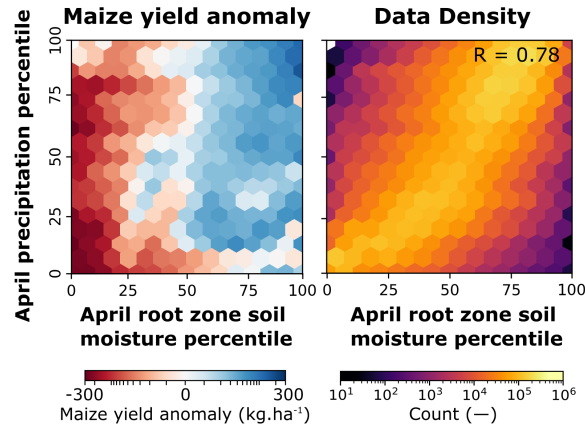
Figure 9. Mean maize anomalies and data density as a function of root zone soil moisture and precipitation percentiles.

Soil temperature and extreme heat

Soil temperatures in October (sowing period) are also critical in controlling yields. When the rainy season is delayed, extreme temperatures in the early stages can potentially damage seeds prior to their germination (Mulenga et al., 2016), or cause an earlier start of the maize reproductive period, increasing the susceptibility to heat and water stress (Harrison et al., 2011; Hatfield and Prueger, 2015). However, elevated temperatures in the early season only lead to yield losses when the soil moisture content at the end of the season was low, otherwise, an overall yield gain is observed (Figure 8a). Thus, wet soil moisture conditions (e.g., irrigation) could play an important role in mitigating extreme heat effects (Troy et al., 2015; Thomas et al., 2020) and potentially even increase productivity (Steward et al., 2018), such as observed in Figure 5b. Nonetheless, because of maize's overall sensitivity to elevated temperatures (Lobell and Field, 2007; Lobell et al., 2013), soil temperature above 37.5 °C would lead to yield losses, despite the wet conditions in the late season (Figure 8a and 8b).

Static landscape predictors

In terms of the static predictors, shrubland ranked ~~second~~ third in Case 2 as it may represent the heterogeneity of agricultural landscapes. While high cropland percent could be associated with more homogenous agricultural fields (often associated with commercial farming), high shrubland percent may indicate a higher presence of fragmented agricultural areas, reflecting poorer agricultural landscapes (from a physical and management perspective), and consequently lower yields (Maggio et al., 2018). Figure S5 in the SI illustrates this relationship. On the other hand, actual shrublands/cropland mosaics may be more likely to be misclassified as a cropland-pixel (commission error), which could explain the negative influence that shrubland percent has on yield. Landscape characteristics such as slope and properties of the soil (residual soil moisture, soil water storage capacity, and wilting point), were also identified as important predictors. These predictors control soil moisture dynamics and the water holding capacity of the soil, as more water available from plants for longer increases the likelihood of higher yields.

Climatological-based predictors

Climatological precipitation over the growing season also ranked as a strong predictor, as it represents the mean spatial distribution of rainfall. Historically, yields in Zambia are higher in locations with more precipitation during the growing season (Zhao et al., 2018). On the other hand, ~~RFE monthly or seasonal~~ in the RFE analysis monthly precipitation (i.e., dynamic) ~~precipitation did not remain as a model predictor~~only ranked 10th. This could be because soil moisture dominated the predictive power at the sub-annual time scale, but also because of the spatial scale mismatch between the input data (30-m soil moisture, 1-km climatological precipitation, and 9-km dynamic precipitation predictors). Conversely, climatological air temperature over the growing season did not remain after RFE, but monthly air temperatures in the early season showed predictive importance, highlighting its contribution at sub-annual time scales.

4.3 Drivers of maize yield spatial variability

Our results show that there is large spatial variability in yields at national and local scales (Figure 4) that is consistently and mainly driven by the spatial variability in soil moisture (Figure 6). The spatial variability in soil moisture originates from the interactions between the meteorological conditions and the landscape. For instance, topography controls lateral flows at the root zone, driving the wet (Figure 1b) and cooler (Figure 1e) soil conditions at the river valleys. Soil properties also influence the spatial variability in soil moisture by driving soil moisture dynamics (e.g., infiltration rates), but also thermal properties of the soil (e.g., thermal conductivity). Land cover and vegetation types (Figure S1) control land-atmosphere interactions (i.e., evapotranspiration), water retention in the root zone, and surface runoff processes. This spatial variability in hydroclimate processes, along with the different farmer management practices, leads to different drought impacts on yields across the landscape. This is evident in the 2015/2016 El Niño drought. Figure 5a, 5b shows that, despite the severity of the event, areas frequently wet and cooler at river valleys are less prone to agriculture losses. ~~This spatial~~ If this modeling framework was implemented using coarser hydrological data we do not expect significant changes in the RF performance (as the data is aggregated to the district-level for training), but we do expect smoother "field-scale" predicted yield. This is expected because hydrology plays an important role in the spatial patterns of the predicted yields, as soil moisture and soil temperature were the strongest predictors (Figure 3). This spatial variability in the impact of drought on yields has important implications for decision making (at national and local scales) as traditional coarse-scale drought indices (and aggregated yield surveys) would have only captured averaged impacts, likely missing the extremes. However, future research should further evaluate the role of spatial scaling in field-scale yield predictions.

4.4 Challenges and limitations for field-scale maize yield prediction

Extending the RF-model to predict maize yields at the field-scale is done under the assumption that the model predictors and parameters trained at the district-scale reflect the relationships at the field-scale. However, there are uncertainties in the training data, predictors, and model parameters that lead to uncertainties in the district-level and field-scale level maize yield estimates.

There are also uncertainties in how yields vary at small scales and the physical processes and variability in management that drive this.

470 *Maize yield training data*

The PHS survey-based dataset, used to train the RF model, was computed by aggregating individual farmer self-reported harvest (number of maize bags) and field areas to the district-level. While aggregating the data provides a more reliable estimate of yield, as it smooths out the noise and removes outliers, it also constrains the ability of the model to predict high and low yields. Similarly, self-reported field-scale data is also very uncertain and sensitive to errors when estimating yields (harvest/area)

475 (Paliwal and Jain, 2020; Gourlay et al., 2019). Ideally, to improve the capability of modeling yield extremes and to ensure the reliability of the field-scale yield estimates, models should be trained on yield data that was properly measured (in terms of weight and area). Currently, such data is hard to obtain for Zambia (and other countries in the region), except for field trials and focused research projects. However, by assuming that the RF model trained on district-level data can reflect the major drivers of maize yields, we provide an estimate of field-scale yield variability that is, at least, grounded in reliable district-level estimates.

480 Comparing and validating field-scale yield estimates is also challenging, mainly because of the mismatches between survey-observed and model-estimated yields. For instance, we assume that the PHS survey (district-level yield data) is representative of the overall yield at each district (i.e., there were no survey sampling issues) and that all 250-m grid cells with at least 50% cropland were maize fields. These mismatches may lead to biases in the aggregated-level estimates.

Modeling framework and predictors

485 ~~Our~~ Despite the good potential to predict yields, our yield predictions showed a tendency to underestimate high yields (~~Figure 2)~~ $> 2500 \, \mathrm{kg \, ha^{-1}}$) and slightly overestimate low yields ($< 500 \, \mathrm{kg \, ha^{-1}}$), as observed in Figure 2. This could be due to above reasons, but also due to the limited training data of ~~high yields (> 2000)~~ extreme high and low values, and a consequence of the RF model ~~averaging many different~~ being somewhat limited in predicting extreme values, as its final estimates are an average of an ensemble of decision trees (e.g., Baccini et al., 2004; Bourgoin et al., 2018). The underestimation of high yields could

490 also be associated with the lack of information on farmer management, such as the use of drought-resistant seeds, fertilizers, or irrigation systems, that are not accounted for in the model. The human-driven factors can play a major role in modulating yield outcomes, especially when hydroclimate conditions are not favorable. We expected vegetation index based predictors to improve model capabilities of identifying high yields, however, NDVI only ranked ~~10th~~ 12th on the RFE (~~0.01~~ <0.005 change in R^2, Figure 3). This could be attributed to mixed NDVI signals from other crops (as we only map cropland occurrence,

495 not maize occurrence), infrequent retrievals (biweekly) with high cloud coverage, and the NDVI tendency to saturate at high LAI values. We also expected socioeconomic based predictors, such as GDP and population, to reflect farmers' access to technology, infrastructure, and markets. However, these predictors were removed by the RFE analysis, showing that yields were not influenced by these socioeconomic variables (similar to Jain, 2007), but mostly influenced by early-season temperature and mid-late season rainfall.

500 In addition, the uncertainties associated with the strongest predictors could be also reflected in uncertainties in the maize yield prediction. As most of the temporal and spatial variability in yields comes from the variability in soil moisture and soil temperature (Figure 6), considering the uncertainties from these estimates is fundamental. Uncertainties arise mostly from LSM input data such as soil properties, meteorological conditions, and land cover, as well as model parameterizations. While validation allows us to quantify the reliability of the simulation, in-situ observations are nonexistent in most of the developing
505 world. Given that agricultural yield dynamics are also heavily influenced by human intervention, pathways forward should consider predictors that better account for human management of crop yields. This can be achieved by accounting for water resources management (e.g., irrigation) when modeling soil moisture dynamics, but also by including additional predictors that reflect farmer decision making (e.g., information on planting dates, seed, and varietal choice, use of pesticides, fertilizers, and machinery, etc.).

510 **5 Conclusions**

Current drought monitoring often relies on hydroclimate data at coarse spatial resolutions or (infrequent) vegetation index retrievals that do not always indicate the conditions farmers face in the field. As a result, few studies were able to link drought indices to agricultural losses in the field (Karthikeyan et al., 2020; Sutanto et al., 2019). Consequently, decision making (by governments and policy-makers, insurance payouts, water resources management, etc.) based on these indices can often be
515 disconnected from the farmer reality.

With the advancement of hyper-resolution modeling (Wood et al., 2011; Bierkens et al., 2014; Chaney et al., 2016; Vergopolan et al., 2020), a new paradigm has been established that allows field-scale agricultural prediction and drought monitoring. In this work, this is achieved by accounting for the water balance in the root zone, as a critical variable for crop productivity, and by representing soil moisture and soil temperature dynamics at a scale that is representative of farm-scale spatial variability. In
520 specific, we used HydroBlocks model to estimate 30-m soil moisture and soil temperature (1981–2018), and combine these and other predictors with a Random Forests to model and map 250-m maize yields (2000–2018) across Zambia. To our knowledge, this study is the first to estimate historical field-scale soil moisture and temperature in this region. Our work advances on previous approaches by integrating these variables into yield prediction, and by more effectively quantifying yield sensitivity to drought at the smallholder farm-scale.

525 By bridging the spatial scale gap between drought monitoring and field-scale agricultural impacts, this work paves the way towards applying field-scale soil moisture monitoring to inform agricultural decision-making across scales. Although our approach is complex and involves integrating remote sensing data, hyper-resolution land surface modeling, long-term district level yield data, and machine learning, it can provide the basis for a practical approach to field-scale monitoring that is an improvement over current approaches that are less accurate and at coarse resolution. In addition, it complements and can help to
530 minimize the difficult task of collecting field-based yield data, which is one of the primary limitations for remote agricultural impact assessments. At the same time, accurate field-scale yield data, and information on biophysical parameters (e.g., soils) and farmer-level management practices are still needed to improve and further validate the approach.

References

Adegoke, J. O. and Carleton, A. M.: Relations between Soil Moisture and Satellite Vegetation Indices in the U.S. Corn Belt, Journal of Hydrometeorology, 3, 395–405, https://doi.org/10.1175/1525-7541(2002)003<0395:rbsmas>2.0.co;2, 2002.

Aghighi, H., Azadbakht, M., Ashourloo, D., Shahrabi, H. S., and Radiom, S.: Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11, 4563–4577, https://doi.org/10.1109/jstars.2018.2823361, 2018.

Alfani, F., Arslan, A., McCarthy, N., Cavatassi, R., and Sitko, N.: Climate-change vulnerability in rural Zambia: the impact of an El Niño-induced shock on income and productivity, http://www.fao.org/3/ca3255en/CA3255EN.pdf, 2019.

Archer, K. J. and Kimes, R. V.: Empirical characterization of random forest variable importance measures, Computational Statistics & Data Analysis, 52, 2249–2260, https://doi.org/10.1016/j.csda.2007.08.015, 2008.

Azzari, G., Jain, M., and Lobell, D. B.: Towards fine resolution global maps of crop yields: Testing multiple methods and satellites in three countries, Remote Sensing of Environment, 202, 129–141, https://doi.org/10.1016/j.rse.2017.04.014, 2017.

Baccini, A., Friedl, M. A., Woodcock, C. E., and Warbington, R.: Forest biomass estimation over regional scales using multisource data, Geophysical Research Letters, 31, n/a–n/a, https://doi.org/10.1029/2004gl019782, 2004.

Bachmair, S., Svensson, C., Hannaford, J., Barker, L. J., and Stahl, K.: A quantitative analysis to objectively appraise drought indicators and model drought impacts, Hydrology and Earth System Sciences, 20, 2589–2609, https://doi.org/10.5194/hess-20-2589-2016, 2016.

Bank, T. W.: Global Risk Data Platform: Gross Domestic Product. United Nations Environment Programme, http://preview.grid.unep.ch/, 2012.

Beck, H. E., Pan, M., Miralles, D. G., Reichle, R. H., Dorigo, W. A., Hahn, S., Sheffield, J., Karthikeyan, L., Balsamo, G., Parinussa, R. M., van Dijk, A. I. J. M., Du, J., Kimball, J. S., Vergopolan, N., and Wood, E. F.: Evaluation of 18 satellite- and model-based soil moisture products using in situ measurements from 826 sensors, Hydrology and Earth System Sciences Discussions, 2020, 1–35, https://doi.org/10.5194/hess-2020-184, https://www.hydrol-earth-syst-sci-discuss.net/hess-2020-184/, 2020.

Berg, A., Findell, K., Lintner, B., Giannini, A., Seneviratne, S. I., van den Hurk, B., Lorenz, R., Pitman, A., Hagemann, S., Meier, A., Cheruy, F., Ducharne, A., Malyshev, S., and Milly, P. C. D.: Land–atmosphere feedbacks amplify aridity increase over land under global warming, Nature Climate Change, 6, 869–874, https://doi.org/10.1038/nclimate3029, https://www.nature.com/articles/nclimate3029, 2016.

Beza, E., Silva, J. V., Kooistra, L., and Reidsma, P.: Review of yield gap explaining factors and opportunities for alternative data collection approaches, European Journal of Agronomy, 82, 206–222, https://doi.org/10.1016/j.eja.2016.06.016, 2017.

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., de Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H., and Wood, E. F.: Hyper-resolution global hydrological modelling: what is next?, Hydrological Processes, 29, 310–320, https://doi.org/10.1002/hyp.10391, 2014.

Borras, L., E, W. M., and E, O. M.: Control of kernel weight and kernel water relations by Post-flowering Source–sink ratio in maize, Annals of Botany, 91, 857–867, https://doi.org/10.1093/aob/mcg090, https://doi.org/10.1093/aob/mcg090, 2003.

Bourgoin, C., Blanc, L., Bailly, J.-S., Cornu, G., Berenguer, E., Oszwald, J., Tritsch, I., Laurent, F., Hasan, A., Sist, P., and Gond, V.: The Potential of Multisource Remote Sensing for Mapping the Biomass of a Degraded Amazonian Forest, Forests, 9, 303, https://doi.org/10.3390/f9060303, 2018.

Breiman, L.: Random Forests, Machine Learning, 45, 261–277, https://doi.org/10.1023/a:1017934522171, 2001.

C3S: ERA5-Land reanalysis, https://cds.climate.copernicus.eu, 2019.

Cai, X., Pan, M., Chaney, N. W., Colliander, A., Misra, S., Cosh, M. H., Crow, W. T., Jackson, T. J., and Wood, E. F.: Validation of SMAP soil moisture for the SMAPVEX15 field campaign using a hyper-resolution model, Water Resources Research, 53, 3013–3028, https://doi.org/10.1002/2016wr019967, 2017.

580   Chaney, N. W., Metcalfe, P., and Wood, E. F.: HydroBlocks: a field-scale resolving land surface model for application over continental extents, Hydrological Processes, 30, 3543–3559, https://doi.org/10.1002/hyp.10891, 2016.

Chaney, N. W., Van Huijgevoort, M. H. J., Shevliakova, E., Malyshev, S., Milly, P. C. D., Gauthier, P. P. G., and Sulman, B. N.: Harnessing big data to rethink land heterogeneity in Earth system models, Hydrology and Earth System Sciences, 22, 3311–3330,

585   https://doi.org/10.5194/hess-22-3311-2018, 2018.

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., and Eklundh, L.: A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter, Remote Sensing of Environment, 91, 332–344, https://doi.org/10.1016/j.rse.2004.03.014, 2004.

Chen, Y., Lu, D., Moran, E., Batistella, M., Dutra, L. V., Sanches, I. D., da Silva, R. F. B., Huang, J., Luiz, A. J. B., and de Oliveira, M. A. F.: Mapping croplands, cropping patterns, and crop types using MODIS time-series data, International Journal of Applied Earth

590   Observation and Geoinformation, 69, 133–147, https://doi.org/10.1016/j.jag.2018.03.005, https://www.sciencedirect.com/science/article/pii/S0303243418302472, 2018.

Chlingaryan, A., Sukkarieh, S., and Whelan, B.: Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review, Computers and Electronics in Agriculture, 151, 61–69, https://doi.org/10.1016/j.compag.2018.05.012, https://www.sciencedirect.com/science/article/pii/S0168169917314710, 2018.

595   CIESIN: Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 10, https://doi.org/10.7927/H4DZ068D, 2017.

Crow, W. T., Kumar, S. V., and Bolten, J. D.: On the utility of land surface models for agricultural drought monitoring, Hydrology and Earth System Sciences, 16, 3451–3460, https://doi.org/10.5194/hess-16-3451-2012, 2012.

Deines, J. M., Kendall, A. D., and Hyndman, D. W.: Annual Irrigation Dynamics in the U.S. Northern High Plains Derived from Landsat Satellite Data, Geophysical Research Letters, 44, 9350–9360, https://doi.org/10.1002/2017gl074071, 2017.

600   D'Odorico, P. and Bhattachan, A.: Hydrologic variability in dryland regions: impacts on ecosystem dynamics and food security, Philosophical Transactions of the Royal Society B: Biological Sciences, 367, 3145–3157, https://doi.org/10.1098/rstb.2012.0016, 2012.

Donaldson, D. and Storeygard, A.: The View from Above: Applications of Satellite Data in Economics, Journal of Economic Perspectives, 30, 171–198, https://doi.org/10.1257/jep.30.4.171, https://olc.worldbank.org/system/files/Primer_Satellite%20Data_Econ.pdf, 2016.

Economist, T.: ARC's covenant, https://www.economist.com/finance-and-economics/2016/08/25/arcs-covenant, 2016.

605   Enenkel, M., Farah, C., Hain, C., White, A., Anderson, M., You, L., Wagner, W., and Osgood, D.: What Rainfall Does Not Tell Us—Enhancing Financial Instruments with Satellite-Derived Soil Moisture and Evaporative Stress, Remote Sensing, 10, 1819, https://doi.org/10.3390/rs10111819, 2018.

Estel, S., Kuemmerle, T., Levers, C., Baumann, M., and Hostert, P.: Mapping cropland-use intensity across Europe using MODIS NDVI time series, Environmental Research Letters, 11, 024 015, https://doi.org/10.1088/1748-9326/11/2/024015, 2016.

610   Estes, L. D., Bradley, B. A., Beukes, H., Hole, D. G., Lau, M., Oppenheimer, M. G., Schulze, R., Tadross, M. A., and Turner, W. R.: Comparing mechanistic and empirical model projections of crop suitability and productivity: implications for ecological forecasting, Global Ecology and Biogeography, 22, 1007–1018, https://doi.org/10.1111/geb.12034, 2013.

Famiglietti, J. S., Ryu, D., Berg, A. A., Rodell, M., and Jackson, T. J.: Field observations of soil moisture variability across scales, Water Resources Research, 44, https://doi.org/10.1029/2006wr005804, 2008.

615　FAO: 2015–2016 El Niño Early action and response for agriculture, food security and nutrition (Update 6), http://www.fao.org/fileadmin/user_upload/emergencies/docs/FAOEl%20NinoReportMarch2016.pdf, 2016a.

　　　FAO: 2015–2016 El Niño Early action and response for agriculture, food security and nutrition (Update 10), http://www.fao.org/3/a-i6049e.pdf, 2016b.

　　　Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer,
620　　S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., and Alsdorf, D.: The Shuttle Radar Topography Mission, Reviews of Geophysics, 45, https://doi.org/10.1029/2005rg000183, 2007.

　　　Fick, S. E. and Hijmans, R. J.: WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, International Journal of Climatology, 37, 4302–4315, https://doi.org/10.1002/joc.5086, 2017.

　　　Folberth, C., Baklanov, A., Balkovič, J., Skalský, R., Khabarov, N., and Obersteiner, M.: Spatio-temporal downscaling of gridded crop model
625　　yield estimates based on machine learning, Agricultural and Forest Meteorology, 264, 1–15, https://doi.org/10.1016/j.agrformet.2018.09.021, 2019.

　　　Franz, T. E., Pokal, S., Gibson, J. P., Zhou, Y., Gholizadeh, H., Tenorio, F. A., Rudnick, D., Heeren, D., McCabe, M., Ziliani, M., Jin, Z., Guan, K., Pan, M., Gates, J., and Wardlow, B.: The role of topography, soil, and remotely sensed vegetation condition towards predicting crop yield, Field Crops Research, 252, 107 788, https://doi.org/10.1016/j.fcr.2020.107788, 2020.

630　Funk, C. and Budde, M. E.: Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe, Remote Sensing of Environment, 113, 115–125, https://doi.org/10.1016/j.rse.2008.08.015, 2009.

　　　Gourlay, S., Kilic, T., and Lobell, D. B.: A new spin on an old debate: Errors in farmer-reported production and their implications for inverse scale - Productivity relationship in Uganda, Journal of Development Economics, 141, 102 376, https://doi.org/10.1016/j.jdeveco.2019.102376, 2019.

635　Grayson, R. and Blöschl, G.: Spatial patterns in catchment hydrology : observations and modelling, Cambridge Univ. Press, 2001.

　　　Gregorutti, B., Michel, B., and Saint-Pierre, P.: Correlation and variable importance in random forests, Statistics and Computing, 27, 659–678, https://doi.org/10.1007/s11222-016-9646-1, 2016.

　　　Gu, Y., Wylie, B. K., Howard, D. M., Phuyal, K. P., and Ji, L.: NDVI saturation adjustment: A new approach for improving cropland performance estimates in the Greater Platte River Basin, USA, Ecological Indicators, 30, 1–6, https://doi.org/10.1016/j.ecolind.2013.01.041,
640　　2013.

　　　Guilpart, N., Grassini, P., van Wart, J., Yang, H., van Ittersum, M. K., van Bussel, L. G. J., Wolf, J., Claessens, L., Leenaars, J. G. B., and Cassman, K. G.: Rooting for food security in Sub-Saharan Africa, Environmental Research Letters, 12, 114 036, https://doi.org/10.1088/1748-9326/aa9003, 2017.

　　　GYGA: Global Yield Gap and Water Productivity Atlas, http://www.yieldgap.org/, 2020.

645　Hao, Z. and Singh, V. P.: Drought characterization from a multivariate perspective: A review, Journal of Hydrology, 527, 668–678, https://doi.org/10.1016/j.jhydrol.2015.05.031, 2015.

　　　Harrison, L., Michaelsen, J., Funk, C., and Husak, G.: Effects of temperature changes on maize production in Mozambique, Climate Research, 46, 211–222, https://doi.org/10.3354/cr00979, https://pubs.er.usgs.gov/publication/70036288, 2011.

　　　Hatfield, J. L. and Prueger, J. H.: Temperature extremes: Effect on plant growth and development, Weather and Climate Extremes, 10, 4–10,
650　　https://doi.org/10.1016/j.wace.2015.08.001, https://www.sciencedirect.com/science/article/pii/S2212094715300116, 2015.

　　　Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler,

I., Mantel, S., and Kempen, B.: SoilGrids250m: Global gridded soil information based on machine learning, PLOS ONE, 12, e0169 748, https://doi.org/10.1371/journal.pone.0169748, 2017.

655 Jain, S.: An empirical economic assessment of impacts of climate change on agriculture in Zambia, The World Bank, https://doi.org/10.1596/1813-9450-4291, 2007.

Jayne, T., Chamberlin, J., Traub, L., Sitko, N., Muyanga, M., Yeboah, F. K., Anseeuw, W., Chapoto, A., Wineman, A., Nkonde, C., and Kachule, R.: Africa's changing farm size distribution patterns: the rise of medium-scale farms, Agricultural Economics, 47, 197–214, https://doi.org/10.1111/agec.12308, 2016.

660 Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., Timlin, D. J., Shim, K.-M., Gerber, J. S., Reddy, V. R., and Kim, S.-H.: Random Forests for Global and Regional Crop Yield Predictions, PLOS ONE, 11, e0156 571, https://doi.org/10.1371/journal.pone.0156571, 2016.

Jin, Z., Azzari, G., Burke, M., Aston, S., and Lobell, D.: Mapping Smallholder Yield Heterogeneity at Multiple Scales in Eastern Africa, Remote Sensing, 9, 931, https://doi.org/10.3390/rs9090931, 2017.

665 Johnson, D. M.: A comprehensive assessment of the correlations between field crop yields and commonly used MODIS products, International Journal of Applied Earth Observation and Geoinformation, 52, 65–81, https://doi.org/10.1016/j.jag.2016.05.010, 2016.

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., Wilkens, P. W., Singh, U., Gijsman, A. J., and Ritchie, J. T.: The DSSAT cropping system model, European Journal of Agronomy, 18, 235–265, https://doi.org/10.1016/S1161-0301(02)00107-7, https://www.sciencedirect.com/science/article/pii/S1161030102001077?via%3Dihub, 2003.

670 Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., and Wood, E. F.: Four decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons, Advances in Water Resources, 109, 236–252, https://doi.org/10.1016/j.advwatres.2017.09.010, 2017.

Karthikeyan, L., Chawla, I., and Mishra, A. K.: A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses, Journal of Hydrology, 586, 124 905, https://doi.org/10.1016/j.jhydrol.2020.124905, 2020.

675 Keating, B. A., Carberry, P. S., Hammer, G. L., Probert, M. E., Robertson, M. J., Holzworth, D., Huth, N. I., Hargreaves, J. N. G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J. P., Silburn, M., Wang, E., Brown, S., Bristow, K. L., Asseng, S., Chapman, S., McCown, R. L., Freebairn, D. M., and Smith, C. J.: An overview of APSIM, a model designed for farming systems simulation, European Journal of Agronomy, 18, 267–288, https://doi.org/10.1016/S1161-0301(02)00108-9, https://www.sciencedirect.com/science/article/pii/S1161030102001089?via%3Dihub, 2003.

680 Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., and Shearer, S.: Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield, Computers and Electronics in Agriculture, 153, 213–225, https://doi.org/10.1016/j.compag.2018.07.016, 2018.

Kintisch, E.: How a 'Godzilla El Nino shook up weather forecasts, Science, 352, 1501–1502, https://doi.org/10.1126/science.352.6293.1501, 2016.

685 Koster, R. D., Walker, G. K., Collatz, G. J., and Thornton, P. E.: Hydroclimatic Controls on the Means and Variability of Vegetation Phenology and Carbon Uptake, Journal of Climate, 27, 5632–5652, https://doi.org/10.1175/jcli-d-13-00477.1, 2014.

Kristjanson, P., Neufeldt, H., Gassner, A., Mango, J., Kyazze, F. B., Desta, S., Sayula, G., Thiede, B., Förch, W., Thornton, P. K., and Coe, R.: Are food insecure smallholder households making changes in their farming practices? Evidence from East Africa, Food Security, 4, 381–397, https://doi.org/10.1007/s12571-012-0194-z, 2012.

690   Lehmann, C. E. R. and Parr, C. L.: Tropical grassy biomes: linking ecology, human use and conservation, Philosophical Transactions of the Royal Society B: Biological Sciences, 371, 20160 329, https://doi.org/10.1098/rstb.2016.0329, 2016.

Lobell, D. B.: The use of satellite data for crop yield gap analysis, Field Crops Research, 143, 56–64, https://doi.org/10.1016/j.fcr.2012.08.008, 2013.

Lobell, D. B. and Burke, M. B.: On the use of statistical models to predict crop yield responses to climate change, Agricultural and Forest

695   Meteorology, 150, 1443–1452, https://doi.org/10.1016/j.agrformet.2010.07.008, 2010.

Lobell, D. B. and Field, C. B.: Global scale climate–crop yield relationships and the impacts of recent warming, Environmental Research Letters, 2, 014 002, https://doi.org/10.1088/1748-9326/2/1/014002, 2007.

Lobell, D. B., Ortiz-Monasterio, J. I., and Falcon, W. P.: Yield uncertainty at the field scale evaluated with multi-year satellite data, Agricultural Systems, 92, 76–90, https://doi.org/10.1016/j.agsy.2006.02.010, 2007.

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

Luo, Q.: Temperature thresholds and crop production: a review, Climatic Change, 109, 583–598, https://doi.org/10.1007/s10584-011-0028-6, 2011.

Maggio, G., Sitko, N. J., and Ignaciuk, A.: Cropping system diversification in Eastern and Southern Africa: Identifying policy options to

705   enhance productivity and build resilience, https://doi.org/10.22004/ag.econ.288953, https://ideas.repec.org/p/ags/faoaes/288953.html, 2018.

Mason, N. M. and Myers, R. J.: The effects of the Food Reserve Agency on maize market prices in Zambia, Agricultural Economics, 44, 203–216, https://doi.org/10.1111/agec.12004, 2013.

McNaughton, K. and Jarvis, P.: Effects of spatial scale on stomatal control of transpiration, Agricultural and Forest Meteorology, 54, 279–302, https://doi.org/10.1016/0168-1923(91)90010-n, 1991.

710   Mueller, N. D., Gerber, J. S., Johnston, M., Ray, D. K., Ramankutty, N., and Foley, J. A.: Closing yield gaps through nutrient and water management, Nature, 490, 254–257, https://doi.org/10.1038/nature11420, 2012.

Mulenga, B. P., Wineman, A., and Sitko, N. J.: Climate Trends and Farmers' Perceptions of Climate Change in Zambia, Environmental Management, 59, 291–306, https://doi.org/10.1007/s00267-016-0780-5, 2016.

Ng'ombe, J. N.: Technical efficiency of smallholder maize production in Zambia: a stochastic meta-frontier approach, Agrekon, 56, 347–365,

715   https://doi.org/10.1080/03031853.2017.1409127, 2017.

Niu, G.-Y., Yang, Z.-L., Mitchell, K. E., Chen, F., Ek, M. B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., and Xia, Y.: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements, Journal of Geophysical Research, 116, https://doi.org/10.1029/2010jd015139, 2011.

Paliwal, A. and Jain, M.: The Accuracy of Self-Reported Crop Yield Estimates and Their Ability to Train Remote Sensing Algorithms,

720   Frontiers in Sustainable Food Systems, 4, https://doi.org/10.3389/fsufs.2020.00025, 2020.

Peichl, M., Thober, S., Meyer, V., and Samaniego, L.: The effect of soil moisture anomalies on maize yield in Germany, Natural Hazards and Earth System Sciences, 18, 889–906, https://doi.org/10.5194/nhess-18-889-2018, 2018.

Petersen, L.: Real-Time Prediction of Crop Yields From MODIS Relative Vegetation Health: A Continent-Wide Analysis of Africa, Remote Sensing, 10, 1726, https://doi.org/10.3390/rs10111726, 2018.

725   Sadri, S., Pan, M., Wada, Y., Vergopolan, N., Sheffield, J., Famiglietti, J. S., Kerr, Y., and Wood, E.: A global near-real-time soil moisture index monitor for food security using integrated SMOS and SMAP, Remote Sensing of Environment, 246, 111 864, https://doi.org/10.1016/j.rse.2020.111864, 2020.

Saxton, K. E. and Rawls, W. J.: Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions, Soil Science Society of America Journal, 70, 1569, https://doi.org/10.2136/sssaj2005.0117, 2006.

730 Scanlon, T. M., Caylor, K. K., Manfreda, S., Levin, S. A., and Rodriguez-Iturbe, I.: Dynamic response of grass cover to rainfall variability: implications for the function and persistence of savanna ecosystems, Advances in Water Resources, 28, 291–302, https://doi.org/10.1016/j.advwatres.2004.10.014, 2005.

Schauberger, B., Archontoulis, S., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Elliott, J., Folberth, C., Khabarov, N., Müller, C., Pugh, T. A. M., Rolinski, S., Schaphoff, S., Schmid, E., Wang, X., Schlenker, W., and Frieler, K.: Consistent negative response of US crops to high 735 temperatures in observations and crop models, Nature Communications, 8, https://doi.org/10.1038/ncomms13931, 2017.

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

Sheffield, J.: A simulated soil moisture based drought analysis for the United States, Journal of Geophysical Research, 109, 740 https://doi.org/10.1029/2004jd005182, 2004.

Steward, P. R., Dougill, A. J., Thierfelder, C., Pittelkow, C. M., Stringer, L. C., Kudzala, M., and Shackelford, G. E.: The adaptive capacity of maize-based conservation agriculture systems to climate stress in tropical and subtropical environments: A meta-regression of yields, Agriculture, Ecosystems & Environment, 251, 194–202, https://doi.org/10.1016/j.agee.2017.09.019, 2018.

Sutanto, S. J., van der Weert, M., Wanders, N., Blauhut, V., and Van Lanen, H. A. J.: Moving from drought hazard to impact forecasts, Nature 745 Communications, 10, 1–7, https://doi.org/10.1038/s41467-019-12840-z, https://www.nature.com/articles/s41467-019-12840-z, 2019.

Thomas, E., Jordan, E., Linden, K., Mogesse, B., Hailu, T., Jirma, H., Thomson, P., Koehler, J., and Collins, G.: Reducing drought emergencies in the Horn of Africa, Science of The Total Environment, 727, 138 772, https://doi.org/10.1016/j.scitotenv.2020.138772, 2020.

Troy, T. J., Kipgen, C., and Pal, I.: The impact of climate extremes and irrigation on US crop yields, Environmental Research Letters, 10, 054 013, https://doi.org/10.1088/1748-9326/10/5/054013, 2015.

750 Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I. J. M., Stahl, K., Hannaford, J., Di Baldassarre, G., Teuling, A. J., Tallaksen, L. M., Uijlenhoet, R., Hannah, D. M., Sheffield, J., Svoboda, M., Verbeiren, B., Wagener, T., Rangecroft, S., Wanders, N., and Van Lanen, H. A. J.: Drought in the Anthropocene, Nature Geoscience, 9, 89–91, https://doi.org/10.1038/ngeo2646, http://refhub.elsevier.com/S0012-8252(18) 30351-9/rf0730, 2016.

Vergopolan, N., Chaney, N. W., Beck, H. E., Pan, M., Sheffield, J., Chan, S., and Wood, E. F.: Combining hyper-resolution land surface 755 modeling with SMAP brightness temperatures to obtain 30-m soil moisture estimates, Remote Sensing of Environment, 242, 111 740, https://doi.org/10.1016/j.rse.2020.111740, 2020.

Waldman, K. B., Vergopolan, N., Attari, S. Z., Sheffield, J., Estes, L. D., Caylor, K. K., and Evans, T. P.: Cognitive Biases about Climate Variability in Smallholder Farming Systems in Zambia, Weather, Climate, and Society, 11, 369–383, https://doi.org/10.1175/wcas-d-18-0050.1, 2019.

760 Wanders, N., Bachas, A., He, X. G., Huang, H., Koppa, A., Mekonnen, Z. T., Pagán, B. R., Peng, L. Q., Vergopolan, N., Wang, K. J., Xiao, M., Zhan, S., Lettenmaier, D. P., and Wood, E. F.: Forecasting the Hydroclimatic Signature of the 2015/16 El Niño Event on the Western United States, Journal of Hydrometeorology, 18, 177–186, https://doi.org/10.1175/jhm-d-16-0230.1, 2017.

Williams, C. A., Reichstein, M., Buchmann, N., Baldocchi, D., Beer, C., Schwalm, C., Wohlfahrt, G., Hasler, N., Bernhofer, C., Foken, T., Papale, D., Schymanski, S., and Schaefer, K.: Climate and vegetation controls on the surface water balance: Synthesis of evapotranspiration 765 measured across a global network of flux towers, Water Resources Research, 48, https://doi.org/10.1029/2011wr011586, 2012.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resources Research, 47, https://doi.org/10.1029/2010wr010090, 2011.

770   Wylie, B. K., Pastick, N. J., Picotte, J. J., and Deering, C.: Geospatial data mining for digital raster mapping, GIScience and Remote Sensing, 56, 406–429, https://doi.org/10.1080/15481603.2018.1517445, 2019.

Xia, Y., Ek, M. B., Peters-Lidard, C. D., Mocko, D., Svoboda, M., Sheffield, J., and Wood, E. F.: Application of USDM statistics in NLDAS-2: Optimal blended NLDAS drought index over the continental United States, Journal of Geophysical Research: Atmospheres, 119, 2947–2965, https://doi.org/10.1002/2013jd020994, 2014.

775   Yonts, C., Melvin, S., and Eisenhauer, D.: Predicting the last irrigation of the season, 2008.

Zargar, A., Sadiq, R., Naser, B., and Khan, F. I.: A review of drought indices, Environmental Reviews, 19, 333–349, https://doi.org/10.1139/a11-013, 2011.

Zhao, Y., Vergopolan, N., Baylis, K., Blekking, J., Caylor, K., Evans, T., Giroux, S., Sheffield, J., and Estes, L.: Comparing empirical and survey-based yield forecasts in a dryland agro-ecosystem, Agricultural and Forest Meteorology, 262, 147–156, 780   https://doi.org/10.1016/j.agrformet.2018.06.024, 2018.