Operational maize yield forecasts for Sub-Saharan Africa using Earth observation data and machine learning

Donghoon Lee[a,*], Frank Davenport[a], Shraddhanand Shukla[a], Greg Husak[a], Chris Funk[a], Laura Harrison[a], Amy McNally[b,c,d], James Rowland[e], Michael Budde[e], Jim Verdin[c]

[a] *Climate Hazards Center, Department of Geography, University of California, Santa Barbara, California, United States*
[b] *NASA Goddard Space Flight Center, Greenbelt, Maryland, United States*
[c] *U.S. Agency for International Development, Washington, D.C., United States*
[d] *Science Applications International Corporation, Reston, Virginia, United States*
[e] *U.S. Geological Survey, Earth Resources Observation and Science Center, Sioux Falls, South Dakota, United States*

Abstract

Food insecurity continues to grow in Sub-Saharan Africa (SSA). In 2019, chronically malnourished people numbered nearly 240 million, or 20% of the population in SSA. Globally, numerous efforts have been made to anticipate potential droughts, crop conditions, and food shortages to foster improved food insecurity early warning and risk management. To support this goal, we develop an Earth Observation (EO) and machine-learning-based operational, subnational maize yield forecast system and evaluate its out-of-sample forecast skills during the growing seasons for Kenya, Somalia, Malawi, and Burkina Faso. In general, forecast skills improve substantially during the vegetative growth period (VP) and gradually during the reproductive development period (RP). Thus, mid-season assessment can provide effective early warning months before harvest. Skillful forecasts (Nash Sutcliffe Efficiency (NSE) > 0.6 and Mean Absolute Percentage Error (MAPE) $< 20\%$) appear approximately two dekads after the VP; for example, they appear in May in Kenya and Somalia, January in Malawi, and July in Burkina Faso.

*Corresponding author
Email address: donghoonlee@ucsb.edu (Donghoon Lee)

During model development, skillful EO features are also identified, such as the number of dry days and extremely high and low temperatures in early VP. Compared to monthly standard EO features, sub-monthly (dekadal), non-standard, and serial EO features significantly improve forecast skills by +0.3 NSE and −10% of MAPE, demonstrating the ability to more precisely and effectively capture favorable or detrimental crop development conditions. Finally, skillful forecasts and practical utility are demonstrated in the recent normal and dry years in each region. Overall, the developed yield forecasting system can provide skillful predictions during the growing season, supporting regional and international agricultural decision-making processes, including informing food-security planning and management, thereby mitigating food shortages caused by unfavorable climate conditions.

Keywords: earth observation, machine learning, food security, maize yield, operational forecast, Sub-Saharan Africa

Highlights

Maize yield forecasts for Sub-Saharan Africa using Earth observation data and machine learning

- A subnational maize yield forecasting system is developed using Earth observation and machine learning
- Skillful forecasts appear during the growing seasons for Kenya, Somalia, Malawi, and Burkina Faso
- Sub-monthly and specific features significantly improve forecast accuracy
- Operationally actionable forecasts are demonstrated in the recent dry and normal years

Maize yield forecasts for Sub-Saharan Africa using Earth observation data and machine learning

Abstract

Food insecurity continues to grow in Sub-Saharan Africa (SSA). In 2019, chronically malnourished people numbered nearly 240 million, or 20% of the population in SSA. Globally, numerous efforts have been made to anticipate potential droughts, crop conditions, and food shortages in order to improve early warning and risk management for food insecurity. To support this goal, we develop an Earth Observation (EO) and machine-learning-based operational, subnational maize yield forecast system and evaluate its out-of-sample forecast skills during the growing seasons for Kenya, Somalia, Malawi, and Burkina Faso. In general, forecast skills improve substantially during the vegetative growth period (VP) and gradually during the reproductive development period (RP). Thus, mid-season assessment can provide effective early warning months before harvest. Skillful forecasts (Nash Sutcliffe Efficiency (NSE) > 0.6 and Mean Absolute Percentage Error (MAPE) $< 20\%$) appear approximately two dekads after the VP; for example, skillful forecasts appear in May in Kenya and Somalia, January in Malawi, and July in Burkina Faso. During model development, effective EO features are also identified, such as precipitation and available water during VP, and dry days and extreme temperatures in early VP. Compared to monthly standard EO features, sub-monthly (dekadal), non-standard, and serial EO features significantly improve forecast skills by $+0.3$ NSE and -10% of MAPE, demonstrating the ability to more precisely and effectively capture favorable or detrimental crop development conditions. Finally, skillful forecasts and practical utility are demonstrated in the recent normal and dry years in each region. Overall, the developed yield forecasting system can provide skillful predictions during the growing season, supporting regional and international agricultural decision-making processes, including informing food-security planning and management, thereby mitigating food shortages caused by unfavorable climate conditions.

Preprint submitted to Global Food Security *March 26, 2022*

Keywords: earth observation, machine learning, food security, maize yield, operational forecast, Sub-Saharan Africa

1. Introduction

Globally, recurring droughts, conflicts, and extreme weather have contributed to severe food shortages and acute food insecurity. In 2019, 750 million people, or 10% of the world population, were exposed to severe levels of food insecurity, and 135 million people were classified as acutely food insecure (FAO, 2021; FSIN, 2020). In Sub-Saharan Africa (SSA), the chronically undernourished population reached 240 million people, or 20% of the population (FAO, 2021). In 2016 and 2017, approximately 13 million people faced severe hunger, primarily in East Africa, as a result of multi-season droughts (Funk et al., 2018).

Operational drought and food security early warning systems have demonstrated their value in facilitating effective humanitarian actions by identifying the most food-insecure populations in need of assistance (Basso et al., 2013; Chipanshi et al., 2015; Fritz et al., 2019; Guimarães Nobre et al., 2019; Nakalembe et al., 2021). For instance, the Famine Early Warning Systems Network (FEWS NET) has developed a number of analytical tools and conducted diagnostic analyses and scenario assessments to assist international relief agencies and national governments in planning for and responding to humanitarian crises (Funk et al., 2019). Recently, the Group on Earth Observations Global Agricultural Monitoring Initiative (GEOGLAM) Crop Monitor alerted Ugandan authorities to an impending crop failure due to severe drought in 2017. This early warning triggered the country's disaster risk financing fund, which assisted approximately 150,000 people and saved $2.6 million during the crisis (Nakalembe, 2018). More global agricultural monitoring systems and operational assessments are reviewed in Nakalembe et al. (2021).

Using EO data, numerous efforts have been made to provide early warning of droughts contributing to food insecurity (Funk et al., 2018; Shukla et al., 2020, 2021). Recently, due to its increased spatial and temporal coverage, open access, and near-real-time operation, EO data has been extensively used for large-scale, subnational-level crop yield monitoring and forecasting. For example, Davenport et al. (2019) developed and evaluated the out-of-sample performance of end-of-season maize yield forecasts for Kenya and Somalia

using five EO products. Laudien et al. (2020) used weather and sea surface temperature data to develop a statistical within-season forecast of maize yields at the subnational level for all of Tanzania with a lead-time of 6 weeks. Schwalbert et al. (2020) developed in-season municipality-scale soybean yield forecasts in Brazil using vegetation indices, temperature, and precipitation data. Machine learning is also widely used for crop yield forecasting, as it is capable of modeling non-linear relationships between EO data and crop yield, and of processing big data efficiently (Paudel et al., 2021).

Despite extensive EO data and modeling frameworks (Basso and Liu, 2019; Schauberger et al., 2020), additional understanding is needed to maximize the utility of EO data and enhance the functionality and accuracy of EO-based crop yield forecast models, ultimately fostering improved in-season food security assessments in food-insecure regions and smallholder areas (Davenport et al., 2019). For instance, while seasonal or monthly accumulations or averaged of EO data are frequently used to monitor and forecast crop conditions (Funk et al., 2019), there has been relatively little investigation of yield forecasting at a finer temporal resolution, such as a dekad (i.e., a period of 10 consecutive days), which can provide a better representation of crop-growth dynamics and phenology (FAO, 2019). Similarly, the utility of various properties derived from the same EO product has been demonstrated in modeling crop development, including the number of dry days and the frequency of extremely high and low temperatures (Basso et al., 2013; Cairns et al., 2013; Laudien et al., 2020). Quantifying the benefits of these perspectives enables a better understanding of the characteristics of EO data and increases its utility for diagnostic analyses and yield forecasting.

While current monitoring tools and scenario assessments are actively used to anticipate potential droughts, crop conditions, and food shortages, there is a need for additional operational large-scale subnational grain yield forecasting systems to foster improved food insecurity early warning, particularly for food-insecure countries in Sub-Saharan Africa (SSA). A more accurate and reliable forecast system will enable decision-makers to monitor crop development and overall crop condition throughout the growing season, ultimately informing decision-making processes related to early disaster response and mitigation measures that reduce food insecurity.

In this study, we outline and describe a machine learning- and Earth Observation (EO)-based, subnational maize yield forecasting system for four countries in SSA— Kenya, Somalia, Malawi, and Burkina Faso. Specifically, we perform a number of different forecasting experiments with the goal of

developing an operational machine learning-supported grain forecasting system. We measure skills of crop yield models by their out-of-sample performance, which is worthwhile for two reasons. First, sample sizes of yield data are typically small, resulting in overfitted models. Second, crop yield models are frequently used to make deterministic and probabilistic predictions about future yields. It, therefore, seems natural to assess their capacity to describe out-of-sample yields (Norwood et al., 2004). Along with evaluating out-of-sample forecast skills, we specifically investigate 1) the most skillful EO features, 2) the quantitative benefits of the finer temporal resolution and non-standard EO data information, and 3) operational forecast performances for recent years. While some technical and decision-making processes remain to be improved, this will provide the groundwork for a more reliable and practical African grain-yield forecasting and food-security warning system with enhanced resolution and spatial coverage.

2. Data

2.1. Maize yield data

Maize production is critical for food security in SSA; eastern and southern Africa consume 85% of maize produced for food (Shiferaw et al., 2011). We obtain maize yield data from the FEWS NET's Data Warehouse (FDW). The FDW updates crop data on a regular basis using agricultural reports from individual countries. Specifically, we obtain district-level maize-harvested areas and quantity-produced data for Kenya's long-rain season (Mar-Aug), Somalia's Gu season (Mar-Aug), Malawi's main season (Oct-Jul), and Burkina Faso's main season (May-Dec) (Table 1). The FEWS NET's crop reporting unit is typically based on administrative boundaries at the time of data collection. However, administrative boundaries can change over time, so historical crop reporting units are frequently inconsistent with current boundaries. We correct these spatial inconsistencies by aggregating or disaggregating time-series crop data according to the changes in administrative boundaries of individual countries. We then calculate maize yields using the corrected harvested areas and quantity produced.

Since all EO data are available beginning in 1981, we use maize-yield data from 1982. Finally, we select districts with more than 14-year records, yielding a total of 150 districts, with an average of 31 years; Kenya (46 administrative level-1 districts with an average of 34 years), Somalia (32

Table 1: Subnational maize yield data used in this study.

Country	Crop season	Administrative level	Number of districts	Reference year	Record years
Kenya	Long-rain season (Mar-Aug)	Level 1	46	2013	34 years
Somalia	Gu season (Mar-Aug)	Level 2	32	1990	21 years
Malawi	Main season (Oct-Jul)	Level 2	27	2003	32 years
Burkina Faso	Main season (May-Dec)	Level 2	45	2001	34 years

*The reference year refers to the year in which administrative boundaries were established.

administrative level-2 districts with an average of 21 years), Malawi (27 administrative level-2 districts with an average of 32 years), and Burkina Faso (45 administrative level-2 districts with an average of 34 years) (Table 1). The long-term mean and coefficient variation of yield data are illustrated in Figure 1 and Figure S1, respectively. Additional processing of the maize yield data is described in the supplementary materials S1.

2.2. Crop-growing season

The growing season period is obtained from the Joint Research Centre's Anomaly Hotspots of Agricultural Production (ASAP) database (Rembold et al., 2019) (`https://mars.jrc.ec.europa.eu/asap/index.php`, last access: 2 Dec 2021). In ASAP, the mean growing season period is defined by a satellite-derived phenology calculated from the long-term average of 10-day MODIS normalized difference vegetation index (NDVI) data. Specifically, we use subnational level start of season (SOS) values, which are defined as the time at which the NDVI reaches 25% of the seasonal ascending amplitude and verified using FAO data. This SOS is moved by two dekads earlier to indicate the phenological start of crop-development processes, such as sowing and germination. In our analysis, vegetative growth period (VP) is defined as 6 dekads (60 days) after SOS, whereas reproductive development period (RP) is defined as the time following VP. The SOS used in this study is depicted in Figure S2.

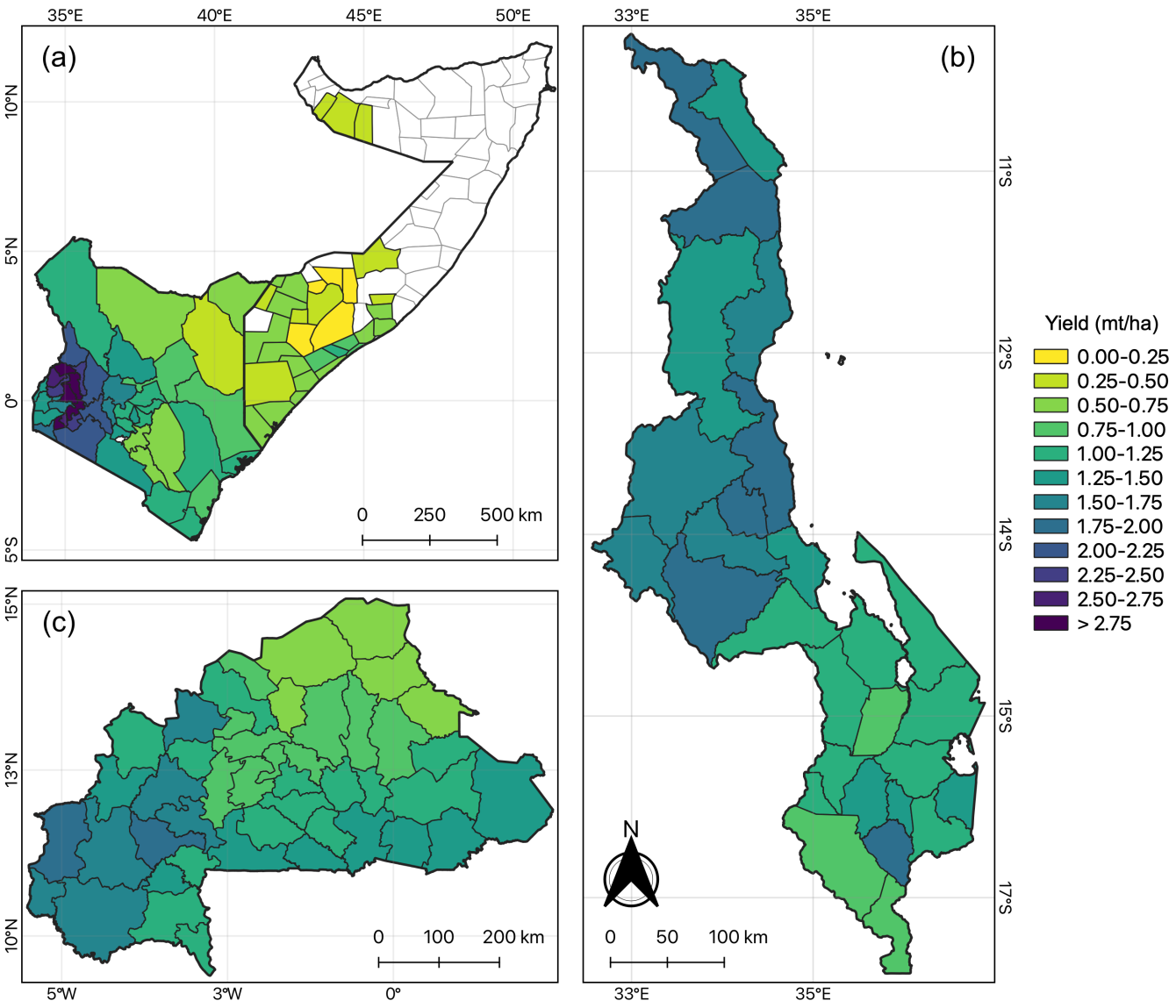

Figure 1: Long-term mean yield of subnational maize data in (a) Kenya (Long rain season) and Somalia (Gu season), (b) Malawi (Main season), and (c) Burkina Faso (Main season).

2.3. EO data sets

In this study, we use four operational EO data sets: precipitation, reference evapotranspiration (ETo; short grass), NDVI, and temperature (Table 2). Supplementary materials S2 describe the quality and details of EO products. These data sets were chosen because they provide operationally usable data with high temporal and spatial resolution with short delays and low latency. We apply a cropland mask (Fritz et al., 2015) and take spatial averages of each EO data on any cropland areas across districts.

We also generate additional EO variables from the four EO data sets. For example, available water is simply calculated by subtracting ETo from precipitation using the same concept as "climatic water balance" (Piedallu et al., 2013), except that reference evapotranspiration is used instead of potential evapotranspiration. The available water on a daily and dekadal basis reflects the surface water content, which is critical for rainfed crops. The growing degree day (GDD) is the most common temperature index for determining the plant development and growth stage (Davenport et al., 2018; Westgate et al., 2004). The GDD is used in certain crops to assist in planning crop-management decisions, such as irrigation and pesticide application timing, as well as harvest scheduling (Hicks et al., 2004). Here, the GDD is calculated based on the daily maximum and minimum temperatures (see supplementary materials S3).

2.4. Non-standard and serial EO features

We generate 8 "standard" EO features: dekadal (10 consecutive days) accumulated precipitation, ETo, available water, and mean of daily minimum, maximum, and average temperature, accumulated GDD, and maximum NDVI. In addition to these commonly used EO features, we also generate 25 additional "non-standard" EO features from the same data sets. These 25 features are not used often in crop modeling and forecasting literature, and we include them here to provide a more thorough examination of which specific features of an EO data set are optimal for crop-yield forecasting.

For precipitation, ETo, and available water, we calculate the number of days in each dekad when the daily record exceeds the district-specific long-term (1986-2015) 95, 90, 80, and 50 percentiles or subceeds long-term 5, 10, and 20 percentiles. Additionally, the number of dry days (i.e., days of zero precipitation) is included, as it has been identified as a critical predictor in Tanzania (Laudien et al., 2020) and Sub-Saharan Africa (Cairns et al., 2013). Laudien et al. (2020) also noted that high temperatures accelerate

the development rate, resulting in a shortened growing season, decreased grain size, and consequently, declined yields. To account for the detrimental effects of excessively high or low temperatures, we include the number of days when the maximum temperature exceeds the 95 percentile or the minimum temperature subceeds 5 percentile. Finally, the accumulated GDD anomaly to the long-term (1986-2015) average is included. Table S1 contains the abbreviations and descriptions for all 33 EO features.

Table 2: A list of Earth observation data sets utilized to forecast yield. The "Features" column describes how we use each product in our statistical models.

Variable	Product	Spatial and temporal resolution, and operational delay	Features[a]
Precipitation	CHIRPS (Climate Hazards Center InfraRed Precipitation with Station data) version 2 (Funk et al., 2015)	0.05° (∼5 km), 1981 - present, Daily data, 3-weeks delay; preliminary data is released 2 days after the end of each pentad (5 day period)	PACU, PDRY, PA95, PA90, PA80, PA50, PB05, PB10, PB20, *AACU, *AA95, *AA90, *AA80, *AA50, *AB05, *AB10, *AB20
ETo	NOAA Reference ET (ETo) Monitoring Data set (uses MERRA2 atmospheric reanalysis) (Hobbins, 2016)	0.125° (∼12km) 1980 - present Daily data About10-days delay	EACU, EA95, EA90, EA80, EA50, EB05, EB10, EB20, *AACU, *AA95, *AA90, *AA80, *AA50, *AB05, *AB10, *AB20
NDVI[b]	NOAA Climate Data Record Advanced Very High-Resolution Radiometer (AVHRR) version 5	0.01° (∼1 km) 1981 - present Daily data	NDVI
	USGS eMODIS temporally smoothed NDVI C6	0.002° (∼250m) 2002 - present Dekadal[c] data 3 dekads of delay (interim data with cloud masks are available)	
Temperature	NOAA CPC global daily maximum and minimum surface air temperature	0.5° (∼50 km) 1979 - present Daily data 1-day delay	TMAX, TMIN, TAVG, TA95, TB05, AGDD, GDDA

[a] The features with asterisk are generated using both precipitation and ETo data sets.
[b] Aggregation of two NDVI data sets is represented in the supplementary materials S2.
[c] Three dekads are used to represent each month. The first two dekads each have ten days. The final dekad has the remaining days in the month.

Along with "non-standard" features at each lead-time, we also consider "serial" features that accumulate the values of an individual feature over the course of the forecast season. For example, if the lead-time is 3 dekads from the starting point, we consider accumulated precipitation during 1, 2, 3, 1-2, 2-3, and 1-3 lead-dekads. This results in 33 features at the beginning of the lead-time (18 lead-dekad) and 5,643 serial features at the end of the lead-time (1 lead-dekad) (Figure 2).

3. Methods

In this study, we develop dekadal maize yield forecast models for a period of 18 dekads lead-time beginning from the pre-season by taking the following steps (Figure 2):

1) Apply a leave-one-out cross-validation (LOOCV) for all years and then select the most skillful EO features in each lead-time.
2) Develop a forecast model using the selected EO features and the first 70% of yield data (year_0 to year_t) to predict a yield in year_{t+1}.
3) Iterate step 2 for the last 30% of data and measure out-of-sample forecast skills.

Therefore, an individual forecast model is constructed and validated for each lead-time, resulting in 18 lead models. The lead models are then re-ordered based on their performances. For instance, if the D13 lead model outperforms the D12 lead model, we keep it instead of the D12 lead model, assuming the D12 lead model selected ineffective EO features (Figure 2). This framework is applied on a district basis for all countries. Also, a long-term trend of yield data is considered by detrending the yield data linearly before the model development and retrending the predicted yield values.

The forecast time horizon is determined using the FEWS NET's crop growing calendar and the actual SOS for districts. While SOS begins with the start of the crop calendar in all districts in Malawi (October), it begins earlier in some districts in Kenya, Somalia, and Burkina Faso (Figure S2). Reflecting this, our forecasts begin in pre-season (i.e., one month earlier) in Kenya and Somalia (February), and in Burkina Faso (April).

3.1. Extremely Randomized Trees (ERT) model

The ERT model is a decision tree-based ensemble regression algorithm that uses randomization and iterations to reduce the chances of overfitting

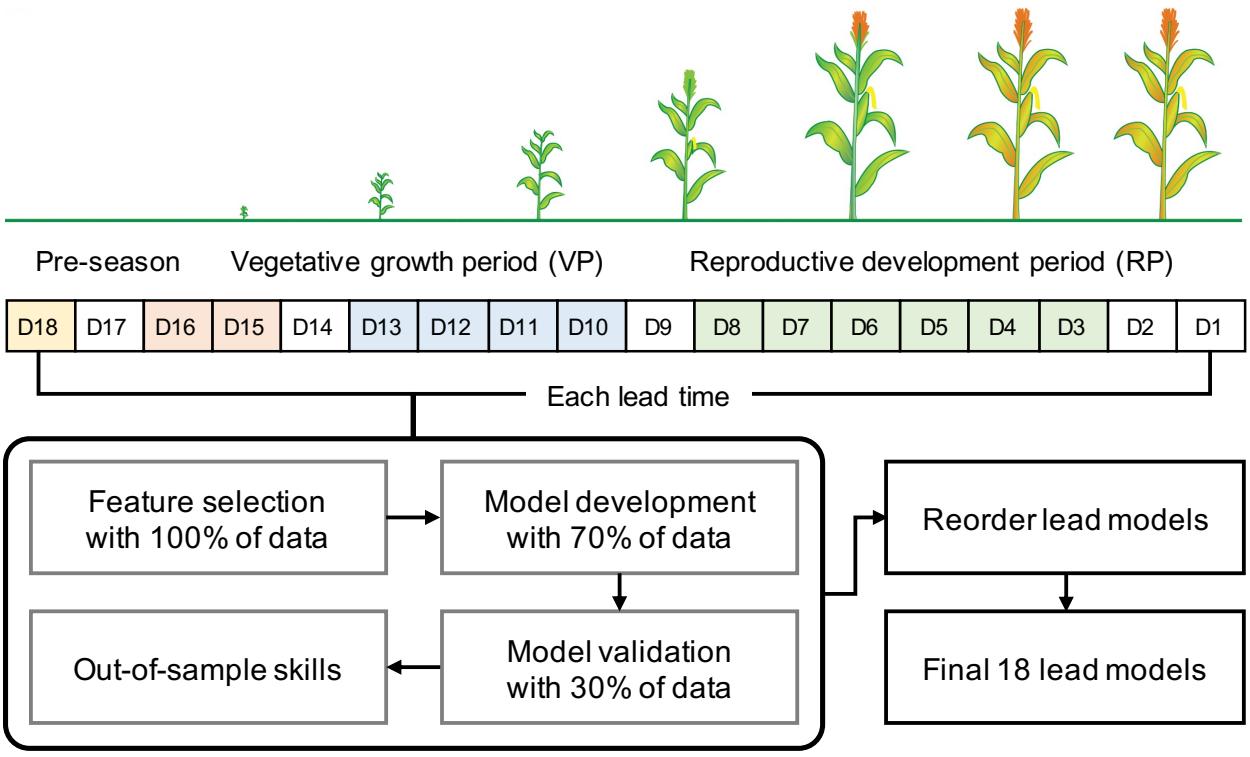

Figure 2: Conceptual framework of forecast model development. The colored lead-time represents the concept of serial features, such as accumulated precipitation during D18, D16-D15, D13-10, and D8-D3. This results in a total of 5,643 possible features during 18 lead dekads. While a forecast time horizon is fixed by country (e.g., February-July for all districts in Kenya), the actual calendar dekads of VP and RP vary by district according to the individual SOS.

models to training data. ERT is very similar to the widely used Random Forest (RF) machine learning algorithm, but has some differences to reduce model bias and variance while also increasing computational efficiency (Geurts et al., 2006). Both algorithms randomly select a subset of predictors for each split; however, in ERT, thresholds are drawn at random for each candidate feature and the best of these is chosen as the splitting rule. The resulting forest contains more variable, but less correlated, trees than those in the RF (Geurts et al., 2006), which means that ensemble predictions are more resilient to individual model errors. Another difference from RF is that ERT retains all of the original samples instead of sampling them with replacement, and this can reduce model bias.

We use the ERT model without parameterization for feature selection (Step 1). When developing a more optimal forecast model (Step 2), hyperparameter tuning is used to achieve high prediction accuracy (see supplementary materials S4).

3.2. Skill scores

We employ two different skill scores: Nash Sutcliffe Efficiency (NSE) and Mean Absolute Percentage Error (MAPE). The NSE is nearly identical to the coefficient of determination and is calculated as one minus the ratio of the prediction error variance to the observation variance, indicating how well the model performs against long-term mean predictions. NSE = 1.0 is considered to be the perfect fit, NSE $>$ 0.75 is considered to be a very good fit, NSE = 0.64 to 0.74 is considered to be a good fit, NSE = 0.5 to 0.64 is considered to be a reliable fit, and NSE $<$ 0.5 is considered to be an unsatisfactory fit (Moriasi et al., 2007).

Additionally, the MAPE is used to evaluate model performance. While the NSE is used to select features and reorder lead models, we used MAPE exclusively for exploration and visualization because it is an intuitive and unit-independent metric that can be used in both low- and high-yield regions.

3.3. Feature selection

In order to handle the large number of feature candidates, we apply a two-step feature selection method with LOOCV for all years. We begin by applying the Boruta algorithm (Kursa and Rudnicki, 2010) to select ten to thirty candidate features. In contrast to conventional feature selection methods, which employ a minimal optimal strategy for a small subset of features with a minimal error, the Boruta algorithm employs an all-relevant feature

selection strategy. Specifically, the Boruta algorithm compares the feature importance (i.e., mean decrease impurity) to the randomized original features and uses a binomial distribution to determine which features truly improve performance. If the Boruta algorithm produces fewer than 30 features for any reason, we calculate the Gini importance (Breiman, 2001) and then select the best 30 features.

After selecting a subset of features, we use forward feature selection to determine the optimal set of EO features with the best performance. Given the emphasis on minimizing error variance in our forecast model framework, we use the maximum NSE as a criterion for selecting the optimal set of features. This criterion is also used to reorder the lead models, but only using model development data in order to preserve the fixed order and out-of-sample independence.

4. Results

4.1. Out-of-sample forecast skills

The out-of-sample forecast skill scores (NSE and MAPE) are illustrated in Figure 3. In general, both skill scores are improved with a reduction in lead-time in all countries. Specifically, the NSE significantly increases until about two dekads after VP, and then gradually improves during the RP. A decrease in skill with decreasing lead-time rarely occurs (e.g., NSE in February in Somalia), except when an inefficient lead model is retained with a higher NSE during the model development period (first 70% of yield data) but not during the out-of-sample validation period (last 30% of yield data). In terms of NSE in the two dekads after VP, Somalia and Malawi show "good" skills (NSE = 0.65 to 0.74), and Kenya and Burkina Faso show "satisfactory" skills (NSE = 0.55).

The skill scores are spatially represented in each region (Figures S3-S5). In Kenya and Somalia, districts with poor NSE scores (< 0.2) appear during the pre-season (February), but not during the VP (March and April), with the exception of a few districts in central Kenya and Somalia. The highest MAPE ($> 90\%$) is found in eastern Kenya, where yields are low and variable (Figure 1a and S1a), but are also improved by lead-time. This strong collinearity between forecast accuracy and district mean yield is also observed in (Davenport et al., 2019). Since the end of May, our forecast has become consistently skillful (NSE > 0.5 and MAPE $< 30\%$) for the majority of districts (Figure S3).

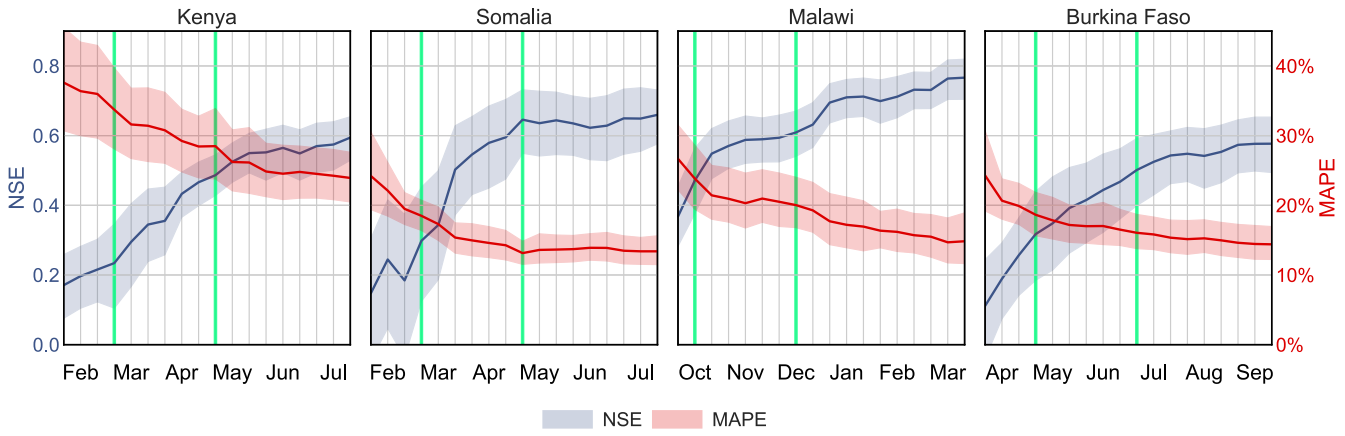

Figure 3: Forecast skill scores (NSE and MAPE) of all districts in each country; lines and shades represent national mean and 90% confidence intervals, respectively. The higher (lower) value of NSE (MAPE) indicates better forecast performance. The vertical green lines mark the national median vegetative growth period (VP).

In October, Malawi's southern region presented low NSE scores (< 0.2), but these rapidly improved in November. Although the most southerly district (Nsanje district) exhibits an abnormally high percentage of error, increasing the overall variation of MAPE scores in Malawi (Figure 3), the low MAPE scores ($< 25\%$) across the country indicate relatively high accuracy. In general, similar to the case of Kenya and Somalia, forecasting skill improves significantly across Malawi during the early VP (October-November) and gradually improves thereafter (Figure S4). While more precise forecasts can be achieved with a shorter lead-time, reliable forecasts (NSE > 0.6 and MAPE $< 25\%$) can be obtained as early as November.

Finally, in Burkina Faso, poor forecasting skills in the pre-season (April) significantly improve during the VP (May and June) (Figure S5). In most districts, reliable forecasts (NSE > 0.5) begin in June; however, in five districts in the western and northern regions, negative NSE scores did not improve until September. This is also due to the fact that, during the model development period, lead models are reordered with ineffective features. As with Malawi, the country has consistently low and stable MAPE scores ($<$ 20%), except in the far north (Oudalan province), where yield data are highly variable (CV $> 80\%$) (Figure S1c).

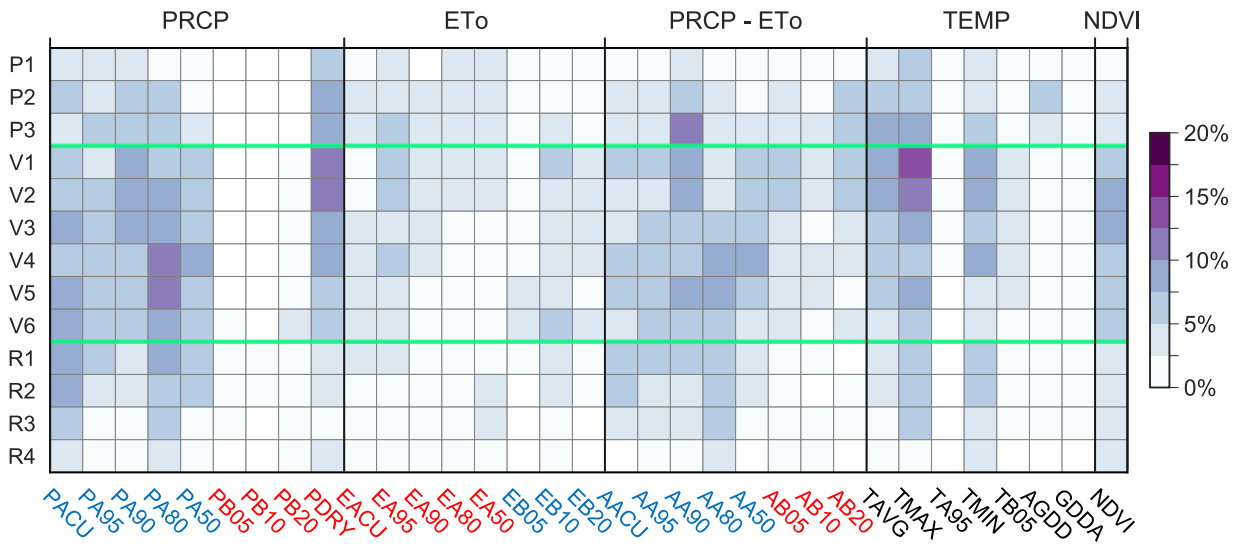

Figure 4: Frequency of EO features selected at R4 (4th dekad of reproductive development period) across all countries. The "P," "V," and "R" represent pre-season, vegetative growth, and reproductive development periods, respectively. Lead-time outlined by green lines are vegetative growth periods. Blue and red-colored feature names denote "wetness" and "dryness" features, respectively.

4.2. Skillful EO features

The model with a shorter lead-time appears to select EO features that have better overall forecasting skill, as the most skillful forecasts tend to be those started after the vegetative growth period when the agroclimatic conditions for the growing season are more determined. We illustrate the frequency of EO features selected by a particular lead model at the 4th dekad of reproductive development period (R4) across all countries (Figure 4) and in each country (Figure S6). Here, we aggregate the R4 lead models based on the growing season of an individual district, which is defined by SOS (Figure S2), so the actual calendar dekads of the R4 lead models could be different regionally.

Certain features are commonly chosen across all countries. Unsurprisingly, precipitation-related features are among the most commonly selected (Figure 4). Many of these areas have primarily rainfed cropping systems. In particular, accumulated precipitation and features of high precipitation percentiles (e.g., PA80 "Number of days precipitation over 80 percentile") are frequently selected during the pre-season and VP in Kenya and Somalia,

as well as during the VP and early RP in Malawi and Burkina Faso (Figure 4 and Figure S6). Low-precipitation features are rarely chosen, besides the PDRY "Number of dry days" which is frequently selected around early VP in Kenya, Malawi, and Burkina Faso. This result is supported by the fact that the amount of grain produced is dependent on the rate and duration of dry matter accumulation, such as consecutive dry days (Laudien et al., 2020), heat stress at high temperatures (Cairns et al., 2013), and soil water deficit and extreme degree days (Basso and Ritchie, 2014).

Features related to available water (PRCP − ETo) are frequently selected throughout the VP, such as AA90 "Number of days available water above 90 percentile" in Kenya and Somalia. This indicates that sufficient water is required to avoid moisture stress during the vegetative development. This important matter of dry conditions in early VP is particularly well captured in Malawi and Kenya by highly selected "dryness" features, such as dry days, high evaporative demand, and low-available water features. In Somalia, the districts in the southern region are irrigated, which may explain why "dryness" features are rarely selected with the exception of EA95 "Number of days evapotranspiration above 95 percentile."

In all countries, the TMAX "Mean of daily maximum temperature" is actively selected around the early VP. Although our temperature data set is surface air temperature, it is known that soil temperatures greater than 15°C and sufficient moisture during this emerging period promote uniform and rapid germination, resulting in improved growth and a reduced risk of fungus damage (Westgate et al., 2004). On the contrary, extreme temperatures during this period can harm plants and result in significant stand loss. In Kenya and Malawi, the TMIN "Mean of daily minimum temperature" is actively selected during the early VP stage, when the growth point is still underground, and the early leaf stage is strongly influenced by cold soil temperature (Kaspar and Bland, 1992). Additionally, the TAVG "Mean of averaged temperature" is frequently selected around the early VP stage in Somalia and Burkina Faso (Figure S6).

Maize is also sensitive to moisture stress during the RP, which is the stage of active grain filling. Specifically, maize that developed well during the VP but is stressed during the grain filling is more susceptible to stem and root rot infection (Sah et al., 2020). In our results, while the "wetness" features are frequently selected during early RP, the "dryness" features are rarely selected. The two-step feature selection procedure used in this study is more performance-oriented than importance-oriented. Whereas the latter

considers the significance of individual features, the former takes into account the performance of the grouped features. For instance, while NDVI clearly correlates with our yield data, it is frequently selected only in Somalia (Figure S6), possibly due to competition from other feature combinations, such as wetness features.

4.3. Improvements from non-standard and serial EO features

The proposed forecast framework makes a significant advancement by incorporating standard, non-standard, and serial EO features that can provide better information about specific temporal crop development. To quantify the improvements, we evaluate forecast skills using a variety of model settings. Specifically, we retain the same model framework and use only non-standard, or serial, or both non-standard and serial, EO features. Additionally, we use the same experimental setup with monthly aggregated features to compare performances of monthly and dekadal features (Figure 5; all cases are illustrated in Figure S6). Except for the first few dekads, the features are selected in a similar number in all lead-times and different feature settings; thus, changes in forecast skills are primarily due to more effective features.

Dekadal forecasts, on average, outperform monthly forecasts, demonstrating how temporally finer EO features can provide superior information (Figure 5). The dekadal data can more precisely represent the conditions of EO data (e.g., dry or wet) during crop-development periods, such as the dryness features selected in early VP (e.g., dry spell) (Figure 4). When both non-standard and serial features are used, dekadal forecasts provide approximately +0.16 of NSE and −6% of MAPE in all lead-times compared to monthly forecasts, and these gaps narrow with shorter lead-times.

The dekadal "non-standard" and "serial" features improve the forecast skills by +0.09 of NSE and −3% of MAPE and +0.05 of NSE and −1% of MAPE than those of the original dekadal features, respectively, indicating usefulness of additional unique information over serial characteristics of the original EO data (Figure 6S). Furthermore, results indicate that using both dekadal non-standard and serial features together can more precisely and effectively capture favorable or detrimental crop-development conditions. This is based on significantly improved forecasting skills, by +0.16 of NSE and −5% of MAPE, which is slightly higher than the combined improvements of the individual cases. Monthly forecasts exhibit the same pattern. Monthly non-standard and serial features, in particular, outperform dekadal, original

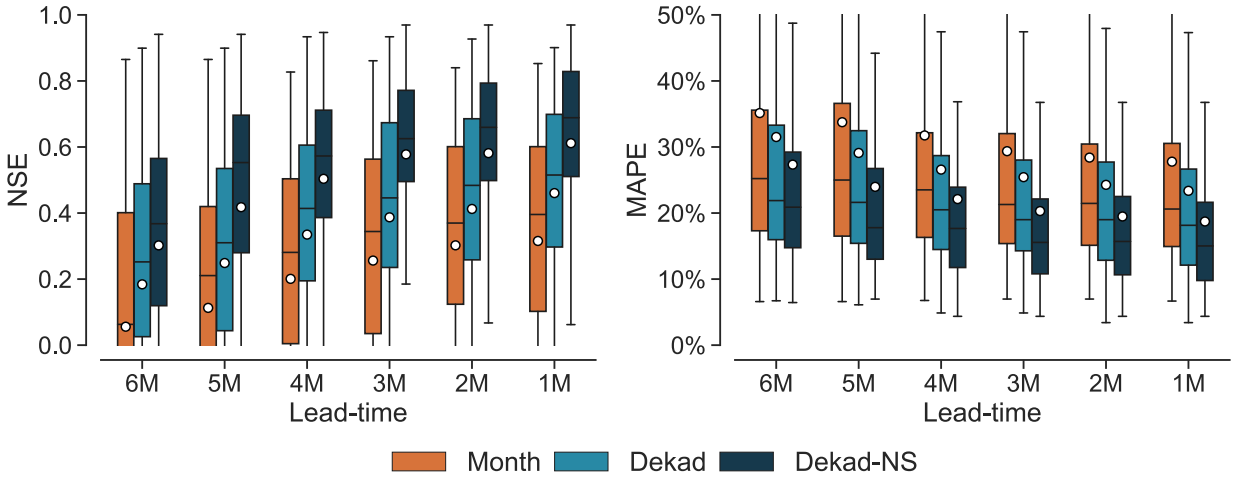

Figure 5: Boxplots of NSE (left) and MAPE (right) scores of monthly and dekadal forecasts of all countries at the end of each lead month. The higher (lower) value of NSE (MAPE) indicates better forecast performance. The model name with "NS" denotes the adoption of non-standard and serial features; otherwise, original features are adopted. The white dot represents a mean value of each boxplot.

features slightly. This means that data with additional and serial characteristics can provide comparable capabilities to data with a higher temporal resolution. Finally, when compared to monthly original features, dekadal non-standard and serial features significantly improve forecasting abilities by +0.3 NSE and −10% of MAPE.

4.4. Operational forecast performances for recent years

To evaluate the forecast system's operational performance, the out-of-sample forecast was tested for recent normal or dry years in each region (Figures 6-8). Here, the forecast models were trained until a year prior to the testing year as an actual practice, and we only consider 4 months from the growing season that forecast information to be useful and critical to management.

For Kenya and Somalia, 2015 (normal year) and 2011 (dry year) are evaluated (Figure 6). In 2015, our forecast slightly underestimated yields in low-yield regions (e.g., eastern and southern Kenya and Somalia) and slightly overestimated yields in high-yield regions (e.g., central and western Kenya). Overall, the 2015 forecast presents highly accurate predictions across the

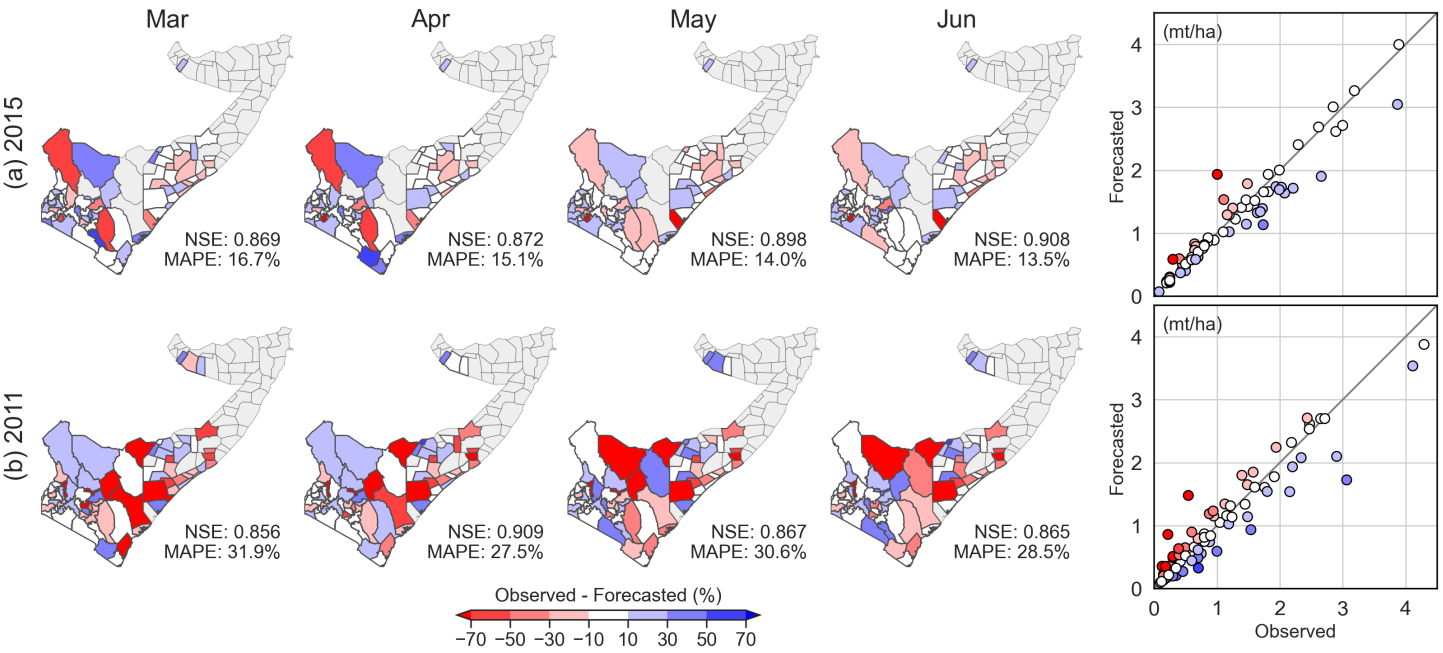

Figure 6: Out-of-sample forecasts for Kenya and Somalia in (a) 2015 (normal year) and (b) 2011 (dry year) at the end of March, April, May, and June. The scatter plots represent the same regional results as the June forecasts (i.e., each circle represent a district). In both maps and scatter plots, all colors represent percentages of errors on the same scale.

regions and especially in the low-yield regions. In comparison to the 2015 forecast, the 2011 forecast shows a reversal of trends in the low and high-yield regions. Although a high percentage of errors occurs in eastern Kenya, forecast skill is generally high (NSE $>$ 0.85). Given that 2011 was a year of severe drought and led to major food insecurity and famine in parts of East Africa (Haan et al., 2012), this performance is promising, as it indicates that crop-yield forecasts could have provided early indication of possible decline in food production. In both years, skillful forecasting (NSE $>$ 0.8) begins in March and continues throughout the growing season, particularly in the high-yield region (e.g., central and western Kenya). With the exception of a few districts in the low-yield region, our forecasts generally maintain the signs of forecasted yields, allowing for consistent management decisions.

Malawi is evaluated using the 2015 (normal) and 2016 (dry) years (Figure 7). In 2015, skillful forecasts (NSE $>$ 0.77) began to appear in early VP (October) and gradually improved with lead-time. In a few districts in

southern Malawi, the yield is overestimated (MAPE > 60%), but this is also improved with lead-time. At the end of January, our forecast demonstrates a high degree of accuracy (NSE is 0.88 and MAPE is 11%). Malawi declared a state of emergency in the 2016-2017 season due to drought, and food insecurity was predicted in central and southern Malawi (FEWS NET, 2017a). Our forecast indicates fair skills (NSE > 0.6) from the early VP. Forecasting accuracy gradually improves with lead-time, but moderate overestimations persist until January in the low-yield region (i.e., southern Malawi). While the MAPE appears to be 85% overall, the absolute percentage of errors appears to be less than 20% without this low-yield region. Given that the actual harvest month is July, the skillful out-of-sample performances from October and November could provide reliable and timely support for agricultural and food-security management (Figure 3). However, a warning for potential overestimation in low-yield regions (i.e., southern Malawi) should be accompanied.

For Burkina Faso, 2014 (normal year) and 2016 (relatively dry year) are evaluated (Figure 8). In 2014, our forecast presents fair forecasts (NSE > 0.5) from the early VP (April), and skills improve significantly in May and are maintained with lead-time. The 2016 forecast presents even better performances (NSE > 0.72 in April). The Sahel region shows excessive errors in April and May, but they are also improved in June. In both years, our forecasts marginally underestimate yields across the country with substantially low errors from June (NSE > 0.7 and MAPE < 14%). This high accuracy might be attributed to the less variable and relatively consistent yield records in Burkina Faso. Based on this operational evaluation and the overall forecast skills (Figure 3), our forecasts can provide accurate predictions from June. Given that the harvest month in Burkina Faso is December, accurate forecasts from June show promising performance.

5. Concluding remarks

In this study, we develop an Earth observation (EO) and machine-learning-based subnational maize yield forecast system for Kenya, Somalia, Malawi, and Burkina Faso. The forecast system is designed to provide dekadal yield prediction throughout a growing season in each country to support various regional and international food-security assessment activities. Along with evaluating out-of-sample forecast skills, we also investigate the most skillful EO features, the quantitative benefits of the finer temporal resolution and

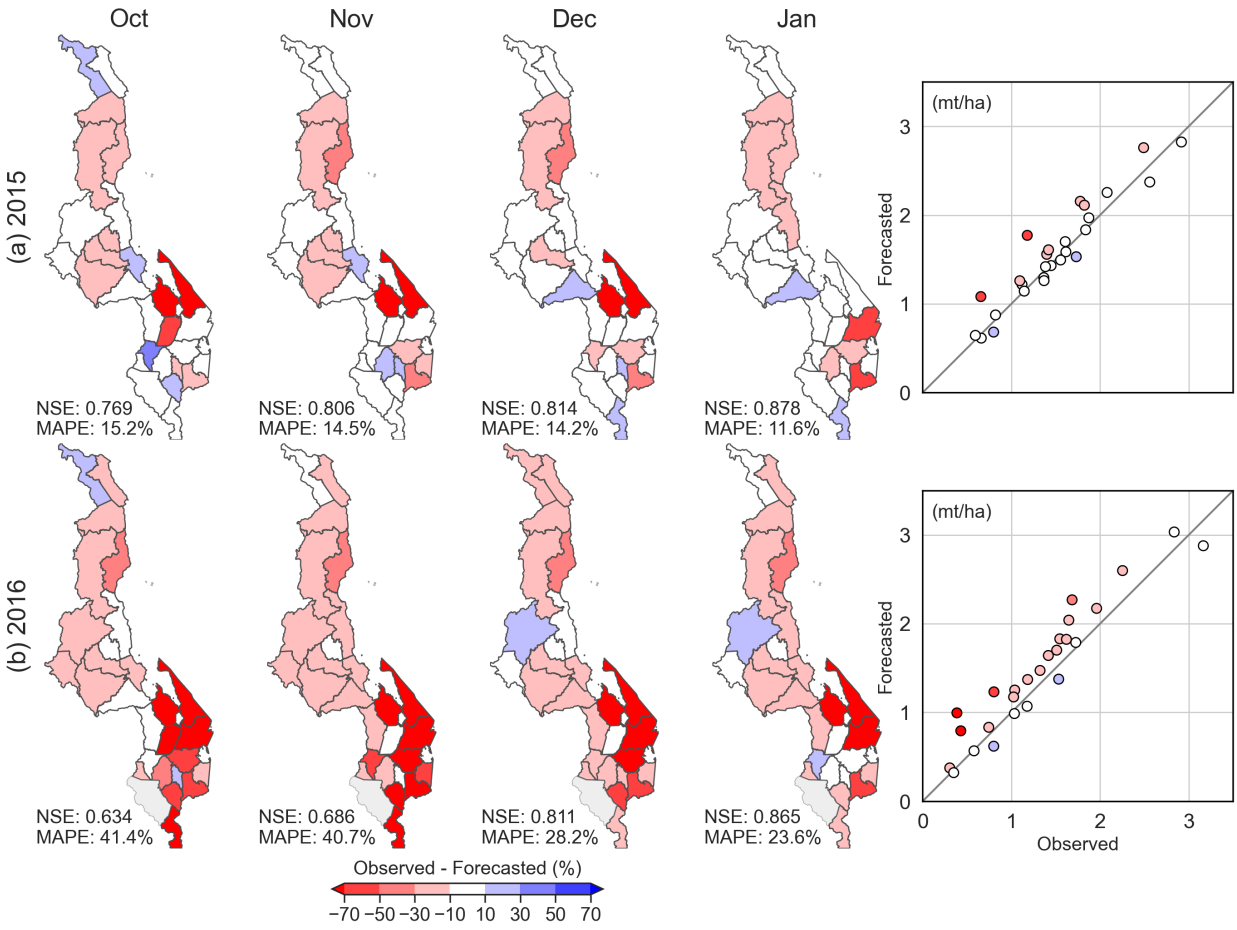

Figure 7: Same as Figure 6, but for Malawi in (a) 2015 (normal year) and (b) 2016 (dry year) at the end of October, November, December, and January. The scatter plots represent the same regional results as the January forecasts (i.e., each circle represent a district).

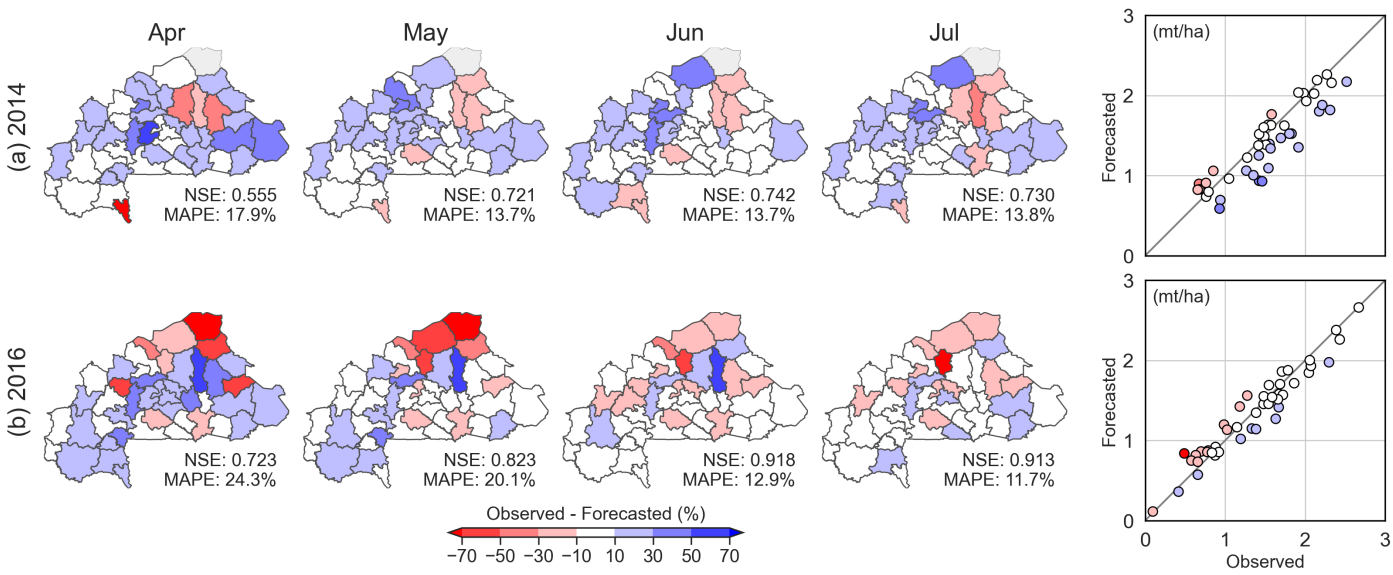

Figure 8: Same as Figure 6, but for Burkina Faso in (a) 2014 (normal year) and (b) 2016 (dry year) at the end of April, May, June, and July. The scatter plots represent the same regional results as the July forecasts (i.e., each circle represent a district).

non-standard EO data information, and operational forecast performances for recent years. Results described in section 4 lead to the following conclusions:

1. In general, forecast skills improve substantially during the vegetative growth period (VP) and gradually during the reproductive development period (RP) (Figure 3). Skillful forecasts (NSE $>$ 0.6 and MAPE $<$ 20%) appear approximately two dekads after the VP: May in Kenya and Somalia, January in Malawi, and July in Burkina Faso. Poor forecast skills also appear in some districts where yield data is significantly low and variable (e.g., eastern Kenya and northern Burkina Faso) or where the model selects ineffective features (e.g., western Burkina Faso). With the exception of these low-skill districts, fair forecasts (NSE $>$ 0.5 and MAPE $<$ 25%) are available even in the early VP (March for Somalia and October for Malawi) and late VP (April for Kenya and June for Burkina Faso).
2. Effective EO features are examined based on the frequency of the selected features. In general, wetness features, such as high precipitation

and high available water, and dryness features, such as number of dry days, are frequently selected in VP. While not surprising, it is worth highlighting that indicators of sufficient water requirement during the vegetative development to avoid moisture stress serves as a valuable indicator of agricultural performance. Maximum and minimum temperatures are also frequently chosen, particularly during the early VP, indicating favorable germination conditions or damage to early leaf development.

3. We discovered that dekadal features outperform monthly features and that non-standard features outperform serial features slightly. When compared to monthly original EO features, the dekadal non-standard and serial features significantly improve forecasting skills by $+0.3$ of NSE and -10% of MAPE. This demonstrates the ability of temporally finer, non-standard EO features to more precisely and effectively capture favorable or detrimental crop-development conditions.
4. The operational performances are also evaluated for the recent normal and dry years in each region. While the overestimations appear in low-yield regions during severe drought years, such as 2011 for Kenya and 2016 for Malawi, the forecasts generally present skillful predictions (NSE > 0.8 and MAPE $< 20\%$) from the early VP in all regions. Although potential uncertainty should be stated in the low-yield region, particularly when drought is anticipated, this result also demonstrates the practical usefulness of the developed operational forecast system in general.

Numerous limitations may be worth investigating in future studies. For instance, we aggregated EO data on cropped areas using a static cropland mask corresponding to 2005. A more precise (i.e., crop-specific) and time-varying mapping of cropland location and extent could improve the accuracy of the aggregated EO data (Peterson and Husak, 2021; Zhang et al., 2019). Other practical EO data sets with data latency or shorter records, such as soil moisture or solar radiation, could be used to improve operational forecasting skills in the late part of the growing season. More advanced categorical or probabilistic yield forecasts should be investigated to improve forecasts, particularly in low-yield regions where food shortage is a critical issue.

The effect of model accuracy on food insecurity risk management may be heteroscedastic, depending on the agricultural and socioeconomic status of the location. Thus, it is worthwhile to tailor crop yield forecast models

in order to accurately forecast extremely low yield conditions, particularly in food-insecure regions where food shortage is a critical problem. A future study could consider more advanced categorical and probabilistic yield prediction techniques, such as weighting a model to accurately predict the bottom 20% percentile of yield records. Additionally, because our approach develops a model for each district, this weighted model could be used exclusively in food security districts at the higher risk of acute food insecurity (Shukla et al., 2021) or when only severe droughts or floods are anticipated, while maintaining overall accuracy for middle and high yield ranges.

In this study, we use only climate and vegetation index data to forecast maize yield. However, yields in some regions may be significantly influenced by factors that the model does not account for, such as local management (e.g., land preparation, sowing, irrigation, and fertilizer), natural phenomena (e.g., flood impacts, plant disease epidemics, and pest outbreaks), and political and socioeconomic changes (Laudien et al., 2020). For instance, pests and pathogens account for 30% of maize yield losses in SSA (Savary et al., 2019), and the risk of maize-specific, virus-induced food insecurity in SSA has increased (Mahuku et al., 2015). This underlines the critical need for incorporating local expert knowledge into operational forecast assessment. Therefore, the forecast should be part of an early warning system that integrates many information sources for the evaluation of the food-security situation (Laudien et al., 2020). The results presented here demonstrate that data-driven modeling can perform well in more humid regions where approaches like the Water Requirement Satisfaction Index (WRSI) can saturate. For example, despite a severe 2015-16 drought, 2015-16 WRSI show little water stress in Malawi, while, in fact, poor harvests led to crisis levels of insecurity (FEWS NET, 2017b).

The analyses and experiments conducted in this paper are intended to assess out-of-sample performances during the model validation period. In actual practice, overall predictability and utility might be enhanced by training and calibrating the models based on full yield records and providing probabilistic uncertainty for low-skill zones at a given lead-time. Overall, the developed yield forecasting system demonstrates its operational capability of providing skillful predictions during the growing season, as well as its potential to support regional and international agricultural decision-making processes, such as triggering relief mechanisms and informing food-security planning and management, thereby mitigating food shortages caused by unfavorable climate conditions.

Data statement

CPC Global Temperature data provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, via https://psl.noaa.gov/data/gridded/data.cpc.globaltemp.html (last access: 2 Dec 2021). CHIRPS precipitation data is available at: https://www.chc.ucsb.edu/data/chirps (last access: 2 Dec 2021). ETo data is available at https://psl.noaa.gov/eddi/globalrefet/ (last access: 2 Dec 2021). eMODIS NDVI C6 data is available at https://earlywarning.usgs.gov/fews/search/Africa (last access: 2 Dec 2021). The subnational maize yield data is subject to FEWS NET's data policy; please contact the corresponding author for data availability.

Acknowledgement

This work is supported by USGS grant #G21AC00026 and the NASA SERVIR program (Award #80NSSC20K0159). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s), including the U.S. Geological Survey. The authors acknowledge the Center for Scientific Computing at the California Nanosystems Institute at the University of California, Santa Barbara for the availability of high-performance computing resources and support. The authors also thank the Climate Hazard Center's technical editor, Juliet Way-Henthorne, for careful copyediting.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at [DOI].

References

Basso, B., Cammarano, D., Carfagna, E., 2013. Review of crop yield forecasting methods and early warning systems, in: Proceedings of the first meeting of the scientific advisory committee of the global strategy to improve agricultural and rural statistics, FAO Headquarters, Rome, Italy, pp. 1–56.

Basso, B., Liu, L., 2019. Seasonal crop yield forecast: Methods, applications, and accuracies. Adv.ances in Agronomy 154, 201–255. doi:10.1016/bs.agron.2018.11.002.

Basso, B., Ritchie, J., 2014. Temperature and drought effects on maize yield. Nature Climate Change 4, 233–233. doi:10.1038/nclimate2139.

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. doi:10.1023/A:1010933404324.

Cairns, J.E., Hellin, J., Sonder, K., Araus, J.L., MacRobert, J.F., Thierfelder, C., Prasanna, B.M., 2013. Adapting maize production to climate change in sub-saharan africa. Food Security 5, 345–360. doi:10.1007/s12571-013-0256-x.

Chipanshi, A., Zhang, Y., Kouadio, L., Newlands, N., Davidson, A., Hill, H., Warren, R., Qian, B., Daneshfar, B., Bedard, F., Reichert, G., 2015. Evaluation of the Integrated Canadian Crop Yield Forecaster (ICCYF) model for in-season prediction of crop yield across the Canadian agricultural landscape. Agricultural and Forest Meteorology 206, 137–150. doi:10.1016/j.agrformet.2015.03.007.

Davenport, F., Funk, C., Galu, G., 2018. How will East African maize yields respond to climate change and can agricultural development mitigate this response? Climatic Change 147, 491–506. doi:10.1007/s10584-018-2149-7.

Davenport, F.M., Harrison, L., Shukla, S., Husak, G., Funk, C., McNally, A., 2019. Using out-of-sample yield forecast experiments to evaluate which earth observation products best indicate end of season maize yields. Environmental Research Letters 14, 124095. doi:10.1088/1748-9326/ab5ccd.

FAO, 2019. Handbook on climate information for farming communities – What farmers need and what is available. Food and Agriculture Organization. URL: https://www.fao.org/publications/card/en/c/CA4059EN/.

FAO, 2021. Africa regional overview of food security and nutrition 2020: Transforming food systems for affordable healthy diets. Food and Agriculture Organization. URL: https://www.fao.org/documents/card/en/c/cb4831en/.

FEWS NET, 2017a. MALAWI Food Security Outlook June 2016-January 2017 Record high staple food prices driving acute food insecurity. Tech-

nical Report. FEWS NET. URL: https://fews.net/southern-africa/malawi/food-security-outlook/june-2016.

FEWS NET, 2017b. MALAWI Food Security Outlook October 2016-May 2017: The already large population of food insecure will continue to grow as the lean season peaks. Technical Report. FEWS NET. URL: https://reliefweb.int/report/malawi/malawi-food-security-outlook-october-2016-may-2017.

Fritz, S., See, L., Bayas, J.C.L., Waldner, F., Jacques, D., Becker-Reshef, I., Whitcraft, A., Baruth, B., Bonifacio, R., Crutchfield, J., Rembold, F., Rojas, O., Schucknecht, A., Van der Velde, M., Verdin, J., Wu, B., Yan, N., You, L., Gilliams, S., Mücher, S., Tetrault, R., Moorthy, I., McCallum, I., 2019. A comparison of global agricultural monitoring systems and current gaps. Agricultural Systems 168, 258–272. doi:10.1016/j.agsy.2018.05.010.

Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., Duerauer, M., Albrecht, F., Schill, C., Perger, C., et al., 2015. Mapping global cropland and field size. Global Change Biology 21, 1980–1992. doi:10.1111/gcb.12838.

FSIN, 2020. Global Report On Food Crises 2020. Technical Report. Food Security Information Network. URL: https://www.wfp.org/publications/2020-global-report-food-crises.

Funk, C., Davenport, F., Harrison, L., Magadzire, T., Galu, G., Artan, G.A., Shukla, S., Korecha, D., Indeje, M., Pomposi, C., Macharia, D., Husak, G., Nsadisa, F.D., 2018. Anthropogenic Enhancement of Moderate-to-Strong El Niño Events Likely Contributed to Drought and Poor Harvests in Southern Africa During 2016. Bulletin of the American Meteorological Society 99, S91–S96. doi:10.1175/BAMS-D-17-0112.1.

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J., 2015. The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. Scientific Data 2, 150066. doi:10.1038/sdata.2015.66.

Funk, C., Shukla, S., Thiaw, W.M., Rowland, J., Hoell, A., McNally, A., Husak, G., Novella, N., Budde, M., Peters-Lidard, C., Adoum, A., Galu, G., Korecha, D., Magadzire, T., Rodriguez, M., Robjhon, M., Bekele, E., Arsenault, K., Peterson, P., Harrison, L., Fuhrman, S., Davenport, F., Landsfeld, M., Pedreros, D., Jacob, J.P., Reynolds, C., Becker-Reshef, I., Verdin, J., 2019. Recognizing the Famine Early Warning Systems Network: Over 30 Years of Drought Early Warning Science Advances and Partnerships Promoting Global Food Security. Bulletin of the American Meteorological Society 100, 1011–1027. doi:10.1175/BAMS-D-17-0233.1.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Machine Learning 63, 3–42. doi:10.1007/s10994-006-6226-1.

Guimarães Nobre, G., Davenport, F., Bischiniotis, K., Veldkamp, T., Jongman, B., Funk, C.C., Husak, G., Ward, P.J., Aerts, J.C., 2019. Financing agricultural drought risk through ex-ante cash transfers. Science of The Total Environment 653, 523–535. doi:10.1016/j.scitotenv.2018.10.406.

Haan, N., Devereux, S., Maxwell, D., 2012. Global implications of Somalia 2011 for famine prevention, mitigation and response. Global Food Security 1, 74–79. doi:10.1016/j.gfs.2012.09.003.

Hicks, D.R., Thomison, P.R., Smith, C.W., Betran, J., 2004. Corn management, in: CW Smith, J. Betran, and ECA Runge (Ed.), Corn: Origin, History, Technology, and Production. John Wiley and Sons, Inc. Hoboken, NJ, pp. 481–522.

Hobbins, M.T., 2016. The variability of asce standardized reference evapotranspiration: A rigorous, conus-wide decomposition and attribution. Transactions of the ASABE 59, 561–576. doi:10.13031/trans.59.10975.

Kaspar, T.C., Bland, W.L., 1992. Soil temperature and root growth. Soil Science 154, 290–299. doi:10.1097/00010694-199210000-00005.

Kursa, M.B., Rudnicki, W.R., 2010. Feature Selection with the Boruta Package. Journal of Statistical Software 36, 1–13. doi:10.18637/jss.v036.i11.

Laudien, R., Schauberger, B., Makowski, D., Gornott, C., 2020. Robustly forecasting maize yields in tanzania based on climatic predictors. Scientific reports 10, 1–12. doi:10.1038/s41598-020-76315-8.

Mahuku, G., Lockhart, B.E., Wanjala, B., Jones, M.W., Kimunye, J.N., Stewart, L.R., Cassone, B.J., Sevgan, S., Nyasani, J.O., Kusia, E., Kumar, P.L., Niblett, C.L., Kiggundu, A., Asea, G., Pappu, H.R., Wangai, A., Prasanna, B.M., Redinbaugh, M.G., 2015. Maize Lethal Necrosis (MLN), an Emerging Threat to Maize-Based Food Security in Sub-Saharan Africa. Phytopathology 105, 956–965. doi:10.1094/phyto-12-14-0367-fi.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE 50, 885–900. doi:10.13031/2013.23153.

Nakalembe, C., 2018. Characterizing agricultural drought in the Karamoja subregion of Uganda with meteorological and satellite-based indices. Natural Hazards 91, 837–862. doi:10.1007/s11069-017-3106-x.

Nakalembe, C., Becker-Reshef, I., Bonifacio, R., Hu, G., Humber, M.L., Justice, C.J., Keniston, J., Mwangi, K., Rembold, F., Shukla, S., Urbano, F., Whitcraft, A.K., Li, Y., Zappacosta, M., Jarvis, I., Sanchez, A., 2021. A review of satellite-based global agricultural monitoring systems available for Africa. Global Food Security 29, 100543. doi:10.1016/j.gfs.2021.100543.

Norwood, B., Roberts, M.C., Lusk, J.L., 2004. Ranking crop yield models using out-of-sample likelihood functions. American Journal of Agricultural Economics 86, 1032–1043. doi:10.1111/j.0002-9092.2004.00651.x.

Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., Athanasiadis, I.N., 2021. Machine learning for large-scale crop yield forecasting. Agricultural Systems 187, 103016. doi:10.1016/j.agsy.2020.103016.

Peterson, S., Husak, G., 2021. Crop area mapping in southern and central malawi with google earth engine. Frontiers in Climate 3. doi:10.3389/fclim.2021.693653.

Piedallu, C., Gégout, J.C., Perez, V., Lebourgeois, F., 2013. Soil water balance performs better than climatic water variables in tree species distribution modelling. Global Ecology and Biogeography 22, 470–482.

Rembold, F., Meroni, M., Urbano, F., Csak, G., Kerdiles, H., Perez-Hoyos, A., Lemoine, G., Leo, O., Negre, T., 2019. ASAP: A new global early warning system to detect anomaly hot spots of agricultural production for food security analysis. Agricultural Systems 168, 247–257. doi:10.1016/j.agsy.2018.07.002.

Sah, R.P., Chakraborty, M., Prasad, K., Pandit, M., Tudu, V.K., Chakravarty, M.K., Narayan, S.C., Rana, M., Moharana, D., 2020. Impact of water deficit stress in maize: Phenology and yield components. Scientific Reports 10, 2944. doi:10.1038/s41598-020-59689-7.

Savary, S., Willocquet, L., Pethybridge, S.J., Esker, P., McRoberts, N., Nelson, A., 2019. The global burden of pathogens and pests on major food crops. Nature Ecology & Evolution 3, 430–439. doi:10.1038/s41559-018-0793-y.

Schauberger, B., Jägermeyr, J., Gornott, C., 2020. A systematic review of local to regional yield forecasting approaches and frequently used data resources. European Journal of Agronomy 120, 126153. doi:10.1016/j.eja.2020.126153.

Schwalbert, R.A., Amado, T., Corassa, G., Pott, L.P., Prasad, P., Ciampitti, I.A., 2020. Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. Agricultural and Forest Meteorology 284, 107886. doi:10.1016/j.agrformet.2019.107886.

Shiferaw, B., Prasanna, B.M., Hellin, J., Bänziger, M., 2011. Crops that feed the world 6. past successes and future challenges to the role played by maize in global food security. Food security 3, 307–327. doi:10.1007/s12571-011-0140-5.

Shukla, S., Arsenault, K.R., Hazra, A., Peters-Lidard, C., Koster, R.D., Davenport, F., Magadzire, T., Funk, C., Kumar, S., McNally, A., Getirana, A., Husak, G., Zaitchik, B., Verdin, J., Nsadisa, F.D., Becker-Reshef, I., 2020. Improving early warning of drought-driven food insecurity in southern Africa using operational hydrological monitoring and forecasting products. Natural Hazards and Earth System Sciences 20, 1187–1201. doi:10.5194/nhess-20-1187-2020.

Shukla, S., Husak, G., Turner, W., Davenport, F., Funk, C., Harrison, L., Krell, N., 2021. A slow rainy season onset is a reliable harbinger of drought in most food insecure regions in Sub-Saharan Africa. PLOS ONE 16, e0242883. doi:10.1371/journal.pone.0242883.

Westgate, M.E., Otegui, M.E., Andrade, F.H., 2004. Physiology of the Corn Plant, in: CW Smith, J. Betran, and ECA Runge (Ed.), Corn: Origin, History, Technology, and Production. John Wiley and Sons, Inc. Hoboken, NJ, pp. 235–272.

Zhang, Y., Chipanshi, A., Daneshfar, B., Koiter, L., Champagne, C., Davidson, A., Reichert, G., Bédard, F., 2019. Effect of using crop specific masks on earth observation based crop yield forecasting across Canada. Remote Sensing Applications: Society and Environment 13, 121–137. doi:10.1016/j.rsase.2018.10.002.