Predicting Rainy Season Onset in the Ethiopian Highlands for Agricultural Planning^①

JONATHAN LALA

Department of Civil and Environmental Engineering, University of Wisconsin–Madison, Madison, Wisconsin

SEIFU TILAHUN

Faculty of Civil and Water Resources Engineering, Bahir Dar Institute of Technology, Bahir Dar, Ethiopia

PAUL BLOCK

Department of Civil and Environmental Engineering, University of Wisconsin–Madison, Madison, Wisconsin

(Manuscript received 5 March 2020, in final form 3 June 2020)

ABSTRACT

The Kiremt rainy season is the foundation of agriculture in the Ethiopian Highlands and a key driver of economic development as well as the instigator of famines that have plagued the country's history. Despite the importance of these rains, relatively little research exists on predicting the season's onset; even less research evaluates statistical modeling approaches, in spite of their demonstrated utility for decision-making at local scales. To explore these methods, predictions are generated conditioned on three definitions of onset, at three lead times, using partial least squares (PLS) regression and random forest classification. Results illustrate moderate prediction skill and an ability to avoid false onsets, which may guide planting decisions; however, they are highly sensitive to how onset is defined, suggesting that future prediction approaches should additionally consider local agricultural definitions of onset.

1. Introduction

Ethiopia's main rainy season, the Kiremt, occurs during the boreal summer and is responsible for 65%–95% of total annual rainfall in the country, making it the primary driver of agricultural production (Segele et al. 2015). Agricultural planning, livestock herding, and reservoir management all rely on these rains, largely affecting national welfare. The tragically re-occurring droughts that have plagued East Africa's most populous country for centuries are most often associated with a failure of the Kiremt rains (Lanckriet et al. 2014). Past disasters and future climate uncertainty have prompted the Ethiopian government to specifically identify early warning systems and development insurance as key strategies in the country's National Adaptation Plan (Government of Ethiopia 2019).

While progress in this area is evident (Drechsler and Soer 2016; Ewbank et al. 2019), the need to identify additional tools to specifically tailor local decision-making is significant.

Year-to-year variability in rainy season precipitation is a major challenge for farmers in East Africa, yet proper management may bring about positive gains or avoided losses. For example, in Kenya, effective use of weather information was found to potentially increase gross margins for maize by up to 70% for perfect information and up to 24% using sea surface temperature (SST)-based forecasts (Hansen et al. 2009). Moreover, estimates indicate that drought early action measures could have saved humanitarian agencies \$1.2 billion (U.S. dollars) over a 15-yr period in the Tigray and Somali regions of Ethiopia (Cabot Venton 2018). Finally, although climate change is projected to increase total precipitation in the Nile basin for most of the year, projections for the core months of the Kiremt are much less clear (Ferguson et al. 2018). This motivates the continued development of proactive sectoral information conditioned on predictions.

^① Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JHM-D-20-0058.s1>.

Corresponding author: Jonathan Lala, jonalala@hotmail.com

Although many studies analyze the hydroclimatology of East Africa, focusing either on the bimodal precipitation regime common in Tanzania and Kenya or the unimodal regime in most of Ethiopia (e.g., [Hastenrath et al. 2011](#); [Moron et al. 2013](#); [Yang et al. 2015](#)), relatively less attention is paid to intraseasonal characteristics, despite their contribution to seasonal variability ([Berhane and Zaitchik 2014](#); [Nicholson 2017](#)). Of the recent studies that do address advances in predicting intraseasonal Kiremt characteristics, most notably [MacLeod \(2018\)](#), they primarily utilize dynamic model predictions at coarse resolution, potentially resulting in a mismatch with localized agricultural decision-making ([Wambui 2019](#)). In contrast, statistical models conditioned with local observations have the potential to capture local variability and small-scale interactions ([Alexander et al. 2019](#)). This may be especially evident where topography is highly varied and hydrologic conditions are heterogeneous, as is the case in the Ethiopian highlands ([Zhang et al. 2016](#); [Alexander et al. 2019](#)). Additionally, statistical prediction models may better capture extremes, when changes from business-as-usual may be most valuable ([Nicholson 2017](#)). A skillful prediction of Kiremt onset, therefore, coupled with effective dissemination of information, may allow farmers to plant crops in a timely manner, guide pastoralists in their search for pasture, and facilitate strategic reservoir operations.

In contrast to seasonal total precipitation, onset definition is much less clear. Definitions range from arbitrary thresholds, to precipitation anomalies, to non-precipitation-based phenomena, such as the behavior of local flora and fauna or changes in atmospheric patterns. These framings, based on observations, can produce widely varying onset dates—from as early as April to as late as July—necessitating an analysis of their strengths and weaknesses, particularly in the context of water management and agricultural planning. This paper presents development and analysis of statistical onset prediction models and their application to multiple onset definitions. Application of forecasts to early-season planting—particularly for maize—is also considered. The Koga watershed in the Ethiopian highlands is selected for demonstration due to its status as an intensive agricultural region and because of the presence of a major reservoir at its outlet.

2. Data and methods

The approach adopted here includes determining observed onset annually according to three definitions, correlating onset date with large-scale preseason climate drivers that may serve as predictors, and producing

season-ahead onset date predictions at four different lead times using quantitative and qualitative methods. Comparisons are then made with farmers' planting decision calendar to demonstrate the utility of an onset forecast.

a. Onset definition

Onset date is characterized according to three definitions: a threshold-based definition (*threshold definition*), a definition based on precipitation anomalies relative to the long-term average (*yearly definition*), and a definition based on precipitation anomalies relative to April–July long-term averages (*window definition*). Daily precipitation values from 1981 to 2019 are taken from the Climate Hazards Group Infrared Precipitation with Station dataset (CHIRPS; [Funk et al. 2014](#)), based on 0.05° satellite imagery and corrected with in situ gauge data; this dataset has been shown to demonstrate very low bias over northwestern Ethiopia ([Dinku et al. 2018](#)). To evaluate the potential predictive skill based on a statistical approach, a primarily rainfed agricultural area upstream of the Koga reservoir (11.05°–11.35°N, 37.00°–37.35°E) is considered. Onset date is calculated at each 0.05° grid independently across the area, with the median value from all grid cells for each year retained for forecast training and validation ([Fig. 1](#)).

The threshold definition ([Segele and Lamb 2005](#)) is defined as a 3-day accumulation of 20 mm or more after 1 April, with no dry spells (<0.1 mm each day for at least eight days) in the next 30 days. Although this is seemingly arbitrary, it is useful for local agronomy, acting as a proxy for soil moisture ([MacLeod 2018](#)), and it was specifically developed for the relatively wet areas of northwest Ethiopia considered in this study. In contrast, two precipitation anomaly-based definitions are also considered. The yearly definition ([Dunning et al. 2016](#)) is calculated by taking the cumulative daily precipitation anomaly, starting 1 January of each year, relative to the long-term daily average. The global minimum of this cumulative time series for each year is defined as the onset date. As this definition makes no assumption of when the season occurs—in contrast to farmers, who are unlikely to plant outside of a given window regardless of precipitation conditions—a similar definition was adapted from [MacLeod \(2018\)](#) in which cumulative anomalies are calculated, starting 1 April, relative to the April–July window only. This window definition, along with the yearly definition, is more likely to avoid false onsets (i.e., significant precipitation followed by a long dry spell), since the cumulative precipitation anomaly will be offset by the subsequent dry spell. [Figure 2](#) illustrates each onset definition for an example year.

FIG. 1. Overview of the study area (red box), showing mean onset date according to the yearly definition.

In addition to onset date, criteria for planting dates are also determined to test the utility of an onset forecast. Planting criteria typically rely on soil moisture, for which a threshold-type definition would be most apt (MacLeod 2018). The threshold onset definition of

Segele and Lamb (2005) is specifically designed for moist areas of Ethiopia that experience Kiremt rains (roughly north of 7°N), such as Koga, but farmers do not have perfect foresight about dry spells occurring in the 30 days after the rainfall threshold is reached.

FIG. 2. Daily precipitation (left axis) and cumulative precipitation anomaly (right axis) for an example year, showing long-term yearly mean (green horizontal line, 4.1 mm day^{-1}), long-term window mean (red horizontal line, 6.0 mm day^{-1}), corresponding cumulative anomalies (measured by right axis), and onsets for each definition.

Maize is among the first crops planted in order to maximize growing season (Liben et al. 2015); however, farmers run the risk of planting during a “false onset,” in which a wet period is followed by a prolonged dry spell, which may reduce yield or require replanting (Kipkorir et al. 2007; Tadross et al. 2009). Discussions with farmers located within or near the Koga study area indicated a typical maize planting date around mid-to-late May most years, which corresponds with mean onset date (Fig. 1); however, dates varied widely, ranging from early May to early July in a survey of approximately 2500 farmers for the 2018 cropping season. A smaller survey of the 2019 cropping season found maize planting dates centered around early June for that year.

Considering the range of planting dates, the suitability of a threshold definition for planting purposes, and a lack of foresight regarding false onsets, three different planting date criteria are developed: A, four days following a wet spell of 20 mm in three days on or after 1 May; B, four days following a wet spell of 20 mm in three days on or after 15 May; and C, four days following a wet spell of 50 mm in four days and at least one wet day (>0.1 mm) in the following three days on or after 1 April. Criteria A and B are taken from the threshold onset definition (Segele and Lamb 2005), but altered to eliminate foreknowledge of dry spells and to restrict planting dates to May or later, reflecting the farmer surveys. Criterion C is inspired by similar threshold methods in Tigray (Araya et al. 2012) and southeastern Africa (Tadross et al. 2009), but has slightly increased precipitation thresholds to account for the relatively wetter climate of the Koga region, as well as requiring an additional rainy day within the three days following a wet spell. Thus, the three criteria serve to reflect three sets of risk aversion, with farmers preferring to maximize growing season choosing criterion A, risk-averse farmers choosing to wait for a longer spell of reliable rain choosing C, and the rest choosing B.

b. Climate signals

Historically, seasonal forecasts of East African precipitation have primarily relied on SST and sea level pressure (SLP) from the Atlantic–Mediterranean, Indian, and Pacific Oceans, although recent research has also highlighted the value of atmospheric variables (Nicholson 2017), some of which are considered in this study (e.g., geopotential height and zonal wind). Moisture transport to the Ethiopian Highlands is primarily sourced from the Atlantic Ocean, Indian Ocean, and Mediterranean Sea (Viste and Sorteberg 2013; Segele et al. 2015). A trajectory analysis conducted by Jury (2011) confirmed an Indian Ocean–Red Sea–Mediterranean Sea cyclonic behavior responsible for major floods in the region.

Selected oceanic and atmospheric variables (NCEP–DOE Reanalysis 2 dataset; Kanamitsu et al. 2002) from January to April (prior to onset) at a semimonthly time step, that may serve as skillful predictors based on the literature described above, are correlated with onset date across 1981–2019. The month(s) with the highest correlation between onset date and potential predictor is retained. Further reduction in the number of predictor variables is achieved using the generalized cross-validation score (GCV; Craven and Wahba 1978), balancing model error and the number of selected predictors, defined as

$$\text{GCV} = \frac{\sum_{i=1}^N \frac{e_i^2}{N}}{\left(1 - \frac{m}{N}\right)^2}, \quad (1)$$

where e_i is the model residual, N is the number of data points, and m is the number of predictors. The set of predictors with the lowest GCV score for each onset definition and lead time is retained (see Table 2 for the full set of retained predictors and Table S1 in the online supplement for the full set of candidate predictors).

c. Prediction models

Two model types are adopted for predicting onset: partial least squares (PLS) regression (Wold et al. 1984), producing a quantitative ensemble prediction of the exact onset date, and random forest classification (Breiman 2001), providing a qualitative categorical (early/normal/late) prediction. Predictions are generated using both methods for four different issue dates—15 March, 1 April, 15 April, and 1 May—to understand the trade-off between lead time and forecast accuracy. As an example, following Table 2, the March 15 issue date for the yearly definition uses four predictor variables: 1) SLP averaged over the month of January and over the coordinates 35°–40°N, 0°–10°E; 2) 500-mb (1 mb = 1 hPa) geopotential height averaged over the month of February and over the coordinates 10°S–5°N, 20°–40°E; 3) 1000-mb geopotential height averaged over 1–14 March and over the coordinates 35°–40°N, 0°–10°E; and 4) 1000-mb geopotential height averaged over 1–14 March and over the coordinates 10°–40°N, 20°–45°E. Note that for a 1 April issue date, predictors 3 and 4 are averaged over the entire month of March, and for a 15 April or 1 May issue date, a fifth predictor (500-mb zonal wind averaged over the coordinates 5°–15°N, 5°–20°E and over 1–14 April or 1–30 April, respectively) is included. Thus, for each predictor, there is a time series with a single value for each year from 1981 to 2019. This set of predictors is then used in the PLS regression or random forest classification. The prediction is still

calculated in the case of an early onset for training purposes; hence “percent onsets before issue date” is included as an independent skill metric (see Table 3).

PLS regression is calculated based on the z scores of the selected climate predictors, unique for each onset definition. Predictors and responses are then decomposed according to

$$X = T\mathbf{P}^T, \quad (2)$$

$$Y = U\mathbf{Q}^T, \quad (3)$$

where X is the set of predictors; Y is the set of responses; T and U are projections of X and Y , respectively; and $\mathbf{P}$ and $\mathbf{Q}$ are the loading matrices of X and Y , respectively (Ng 2013). The decompositions are designed to maximize the covariance between X and Y , solving the multicollinearity problem, similar to principal component analysis, while also explicitly considering Y in the estimation of components (Abdi 2003). Terms explaining more than 10% of the variance in the input data are retained, similar to Kaiser’s rule (Kaiser 1960). Given the relatively small number of years ($n = 39$) from which to fit the models, and given the statistical independence in onsets between years, a hindcast across 1981–2019 is constructed in a drop-one cross-validation mode to evaluate performance. Ensemble predictions for each year are formed by bootstrapping, with replacement, from the residuals of the hindcast. Additionally, a qualitative prediction is generated from the deterministic model output, classifying as early, normal, or late onset based on the predicted date relative to the tercile of historical onset dates.

Separately, a random forest qualitative prediction is calculated using the same input data as the PLS regression and same classification criteria. Random forests are generated from a set of decision trees, with each tree generated by bagging (i.e., randomly selecting, without replacement, from examples in the training set); the general form is given as

$$\{h(\mathbf{x}, \Theta_k), k = 1, \dots\}, \quad (4)$$

where h is the classifier function, $\mathbf{x}$ is the set of inputs, and $\{\Theta_k\}$ is a set of independent but identically distributed random vectors. Classifications are made from the votes of each tree for the most popular class at input $\mathbf{x}$ (Breiman 2001).

Skill is measured using mean absolute error in days and rank probability skill score (RPSS; Wilks 1995) for the quantitative forecasts and correct classification rate and extreme miss rates (i.e., forecast of early when observed late and vice versa) for the qualitative forecasts. The RPSS provides a skill metric for ensemble forecasts

by comparing the probabilities of categorical predictions with respect to a reference forecast, such as climatology. The rank probability score (RPS) is first calculated as

$$\text{RPS} = \sum_{i=1}^n \left(F_{\text{fct}_i} - F_{\text{obs}_i} \right)^2, \quad (5)$$

where F_{fct_i} and F_{obs_i} are the cumulative probabilities of the forecasts and observations, respectively, through category i , and n is the number of categories. The RPSS is then calculated as

$$\text{RPSS} = \left(1 - \frac{\text{RPS}_{\text{forecast}}}{\text{RPS}_{\text{climatology}}} \right) \times 100\%. \quad (6)$$

An RPSS of 100% indicates a perfect forecast, whereas scores below zero indicate inferior skill relative to climatology. For this study, the median RPSS over all hindcast years is taken as the relevant score. To capture statistically significant trends, the dates separating categories (early, normal, or late) are shifted each year by the slope of the trend line, uniquely for each onset definition.

3. Results

The Kiremt season generally observes a gradient of increasingly later onset as one moves northeast; this is true of all definitions of onset, although the anomaly methods tend to define onset later overall than the threshold method does (Fig. 3). Variability in onset date follows a similar pattern among definitions, with eastern areas experiencing more variability than western areas. However, variability also dramatically increases at the southern and eastern edges of the Kiremt region, where a clear unimodal rainfall regime starts to transition into the bimodal regime of southern and eastern Ethiopia (Tsidu 2012).

For Koga, onset dates differed notably depending on the definition; early or late onset generally agrees between definitions, however, the specific dates vary widely (Fig. 4, Table 1). Generally, the threshold definition produces the earliest onsets whereas the window definition estimates the latest onsets.

a. Climate signal results

In addition to SLP, which has been historically popular in seasonal forecasts, atmospheric variables—particularly geopotential height and zonal wind—are also found to be important contributors to onset across definitions (Table 2). Correlations with onset tend to be slightly stronger among the anomaly definitions, which have an intrinsic connection with total seasonal precipitation.

FIG. 3. (left) Mean onset dates and (right) standard deviations for the (top) threshold, (middle) yearly, and (bottom) window definitions in northeastern Ethiopia, 1981–2019, showing the Koga study area (red box).

FIG. 4. Historical onset dates and trends for three definitions.

In contrast, the threshold definition relies more strongly on specific weather events (e.g., a single period of heavy rainfall), which are more difficult to forecast given the long lead times considered in this paper. Because of collinearity between predictors, correlations are not necessarily indicative of the strongest predictors, however; for example, 200-mb geopotential height over the Sahara has the lowest correlation coefficient with threshold onset, but the highest PLS coefficient.

SST is notably absent from all onset definitions, suggesting that surface forcings may play a lesser role in onset. Contrastingly, pressure variables (SLP and geopotential height) are featured as predictors in all onset definitions, with the Mediterranean and Red Sea featured as key regions. Predictors located in these regions exhibit high PLS coefficient values for the threshold and window methods, suggesting primacy in terms of influencing onset and confirming the findings of Jury (2011). For the yearly method, low-level pressures in the Atlantic and Pacific tend to dominate, although their influence is less stark relative to other variables, for which PLS coefficients are of similar magnitude.

b. Quantitative forecast results

PLS regression-based predictions demonstrate improvement for all definitions and all issue dates relative to climatology (Table 3). The GCV score for the window definition was lowest in the absence of April climate signals; hence, only a 15 March and 1 April forecast date are presented.

All definitions display only modest reduction in error and low median RPSS scores for a 15 March issue date, but performance in both metrics increases consistently with later issue dates. Median RPSS for the five earliest and five latest onset dates generally—but not always—increases with later issue dates, suggesting that although later issue dates result in more accurate forecasts, uncertainty bounds are not substantially reduced.

This is particularly true for the threshold method, in which RPSS scores for late years actually drop for a 1 May issue date (Table 3) and for which error bars do not noticeably narrow with later issue dates (Fig. 5). Later issue dates also correspond to occasional missed onsets, in which the forecast is issued after onset occurred. This is also a notable problem for the threshold definition, in which 5% of historical onsets occurred before 15 April and 23% occurred before 1 May. In contrast, no onsets were observed before 1 April, for any definition, over the entire study period.

In contrast to the threshold definition, the yearly definition demonstrates greater skill in nearly all metrics, while reducing the risk of issuing a forecast after onset has occurred (Table 3, Fig. 6). Confidence intervals also tend to narrow, albeit slightly, with later forecast dates, which is less apparent for the threshold definition.

The window definition demonstrates the least skill among definitions, and confidence intervals only marginally narrow with a later issue date (Fig. 7). Predictions from later issue dates (15 April and 1 May) are also conditioned on the same set of pre-April predictors, as potential additional predictors from April were eliminated by GCV, thus resulting in no change to the forecast.

c. Qualitative forecast results

Qualitative forecasts for the threshold definition show general improvement with later issue dates, particularly by reducing the rate of extreme misses (Fig. 8).

TABLE 1. Pearson's correlation coefficient (above diagonal) and mean absolute difference (days; below diagonal) between onset dates for each definition.

	Threshold	Yearly	Window
Threshold	—	0.73	0.56
Yearly	6.7	—	0.56
Window	18.3	16.7	—

TABLE 2. Predictor variables, correlation with onset, and PLS coefficients.

Variable		Region	Latitude	Longitude	Month(s)	Correlation	PLS coefficient			
Threshold	SLP						15 Mar	1 Apr	15 Apr	1 May
Window	SLP	Southern Ethiopia	0°–5°N	40°–45°E	February–March	0.43	6.46	3.58	2.05	1.62
	Geopotential height	Southern Ethiopia	0°–5°N	40°–45°E	March	0.53	2.99	4.27	3.58	2.53
		Eastern Mediterranean (500 mb)	30°–35°N	25°–35°E	February	–0.42	–6.70	–5.03	–7.14	–7.99
		North Africa (200 mb)	20°–30°N	30°–40°E	March	0.56	1.61	5.32	5.60	5.59
Yearly	Geopotential height	Sahara (1000 mb)	20°–30°N	10°–30°E	April	0.38	—	—	7.92	8.96
		Eastern Mediterranean (500 mb)	30°–40°N	25°–35°N	January–February	–0.57	–10.09	–9.12	—	—
	Zonal wind	Western Pacific (1000 mb)	15°S–5°N	150°–175°E	February	0.39	6.26	5.20	—	—
		Central Sahara (1000 mb)	15°–20°N	0°–10°E	March	–0.52	–2.46	–5.62	—	—
Yearly	SLP	North Atlantic	35°–40°N	0°–20°W	January	0.45	8.14	7.44	6.07	6.15
	Geopotential height	Eastern Mediterranean (500 mb)	30°–45°N	20°–40°E	February	–0.51	–7.82	–5.76	–5.98	–5.86
		Equatorial Pacific (1000 mb)	10°S–5°N	130°–160°W	March	–0.54	—	–6.71	–6.74	–6.73
	Zonal wind	Mediterranean/Red Sea (1000 mb)	10°–40°N	20°–45°E	March	0.54	5.33	5.17	5.11	4.23
Central Sahara (500 mb)		5°–15°N	5°–20°E	April	–0.41	—	—	–4.39	–5.18	

Correct classification also tends to increase increases with later issue dates for the PLS regression (from 35% to 54%) between 15 March and 1 May, but the improvement is marginal for the random forest method. Overall, the PLS method seems to outperform the random forest method in both correct classification and avoiding extreme misses.

Relative to the threshold definition, qualitative forecasts for the yearly definition have more correct classifications (54%–70% for the PLS regression and 31%–62% for the random forest) as well as fewer extreme misses (Fig. 9). This complements the results of the quantitative forecasts, for which the yearly method is also generally superior.

Finally, the window method performs moderately well in classifying categories for both the PLS and random forest methods, although both result in several extreme misses (Fig. 10). Notably, the random forest method slightly outperforms the PLS method in correct classification rates (60% versus 56%, respectively, for 15 March, and 60% versus 59%, respectively, for 1 April).

Performance between post-PLS regression classification and direct classification via random forests was mixed; classifying post-PLS slightly outperformed the random forest model for the threshold and yearly definitions and reduced the rate of extreme misses for the window definition, however, the random forest model was superior in correct classification for the window definition. In all three definitions, however, random forest classification results in more extreme misses. Thus, there may be scope for considering both models simultaneously.

d. Application to planting dates

Planting dates vary widely from year to year, with criterion A usually the earliest and criterion C the latest (Fig. 11). The spread agrees well with the surveys, which centered 2018 planting dates around late May with several weeks to either side and centered 2019 planting dates around early June.

For early-season planting, a primary hazard to farmers is a false onset, in which farmers plant during a wet spell that is followed by a dry spell (Kipkorir et al. 2007; Tadross et al. 2009). Indeed, farmers in the Koga area have specifically requested onset forecasts in order to properly prepare for planting. To test the utility of the forecasts to this end, a planting date occurring greater than 7 days prior to onset was defined as a “false start”; planting based directly on the forecast was used to compare to the naïve planting criteria. For example, in 1997, early wet spells triggered planting dates on 14 May, 3 June, and 9 June for criteria A, B, and C,

TABLE 3. Average prediction error (days) and rank probability skill scores (RPSS) for each forecast (by onset definition and issue date) using PLS regression, including percent reduction in error over climatology (prediction of mean onset date) and percent of onsets occurring before issue date.

Definition	Average climatology error (days)	Issue date	Average prediction error (days)	Reduction in error	Onsets before issue date	Median RPSS	Median RPSS (early)	Median RPSS (late)
Threshold	16.4274	15 Mar	13.1017	20%	0%	31.56%	55.90%	40.64%
		1 Apr	12.3523	25%	0%	35.16%	79.19%	53.74%
		15 Apr	11.4669	30%	5%	49.38%	84.86%	54.55%
		1 May	10.7199	35%	23%	49.15%	86.72%	29.93%
Yearly	15.5464	15 Mar	12.0016	23%	0%	25.39%	68.25%	0.55%
		1 Apr	10.5222	32%	0%	40.37%	90.48%	36.41%
		15 Apr	10.0029	36%	5%	50.28%	95.95%	58.46%
		1 May	9.7161	38%	15%	54.98%	93.50%	69.74%
Window	13.6963	15 Mar	10.9614	20%	0%	14.23%	45.39%	4.78%
		1 Apr	10.0559	27%	0%	32.95%	83.35%	8.83%

respectively, but onsets actually occurred on 26 May, 28 May, and 24 June for the threshold, yearly, and window methods, respectively. Thus, criterion A was considered a false start according to all onset definitions, whereas criteria B and C were only considered false starts for the window definition. For years in which there was a false start given a chosen onset definition and planting criterion, forecasts were found to successfully avoid false onsets by advising a later planting date (Table 4).

As expected, criterion A runs the most risk of false onset, while criterion C experiences the least risk. For all onset definitions and forecast issue dates, at least some false onsets are avoided; later issue dates tend to avoid more false onsets except in the case of the window definition. Because the window definition tends to specify later onsets in general, it has a very high rate of false onsets, suggesting that it is less useful of an onset definition for maize farmers, although it may be

useful for midseason crops or for other sectors. In contrast, the yearly method substantially reduces the number of false onsets, but this reduction is not monotonic with successive issue dates. The threshold definition, however, results in at least a 50% reduction in false onsets relative to any naïve planting criteria for a 1 May issue date.

4. Conclusions and discussion

This study presents an evaluation of Kiremt onset predictions using statistical methods. Three definitions of onset and three lead times are proposed, applying both a PLS regression and random forest classification method. Predictions illustrate moderate skill by improving over climatology, although later prediction issue dates (e.g., 15 April and 1 May) result in missing some onsets (occurring prior to the issue date). Ensemble predictions using PLS regression show moderate skill in both average

FIG. 5. Ensemble hindcasts, threshold definition, for 15 Mar (leftmost bar) to 1 May (rightmost bar) issue dates, along with observed onset date (black line and circles) and climatological mean onset date (blue line).

FIG. 6. Ensemble hindcasts, yearly definition, for 15 Mar (leftmost bar) to 1 May (rightmost bar) issue dates, along with observed onset date (black line and circles) and climatological mean onset date (blue line).

errors and RPSS values, especially for later issue dates. Qualitative predictions are classified using both random forests and PLS regression, however, neither was unilaterally superior across all onset definitions. Results compare favorably with dynamic model predictions.

The variable performance of predictions considering onset definition and issue date suggests that onset is perhaps not best defined with a single method or model. Individual sectors—agriculture, pastoralism, reservoir management, etc.—certainly have varying definitions of onset, highlighting the importance of tailored decision-making for forecast-informed management decisions. Likewise, complex, heterogeneous geography or climate can further compound variation between onset definitions, motivating localized modeling approaches.

Although no significant correlation between onset date and Kiremt total seasonal precipitation for the

Koga region is evident, all definitions display a trend of increasingly early onset ($\sim 0.3\text{--}0.6$ days yr^{-1}). In fact, total precipitation during the core months of June–September has remained stationary, suggesting the Kiremt season is getting longer but not necessarily wetter. Correlations between onset and total annual precipitation, however, are significant (Pearson's r of -0.60 , -0.26 , and -0.51 , for the threshold, window, and yearly onset dates, respectively) in the Koga region, although total annual precipitation does not independently demonstrate a statistically significant increasing trend. Thus, precipitation earlier in the year has a greater effect on onset date than does precipitation during the core months of the Kiremt.

This study also highlights the importance of including atmospheric variables as seasonal predictors as suggested by [Nicholson \(2017\)](#). Geopotential height is found as an

FIG. 7. Ensemble hindcast, window definition, for 15 Mar (left bar) and 1 Apr (right bar) issue dates, along with observed onset date (black line and circles) and climatological mean onset date (blue line).

FIG. 8. Qualitative hindcasts, threshold definition, using (left) PLS regression and (right) random forest classification for 15 Mar (topmost value for each square) to 1 May (bottommost value) issue dates. Values indicate percentage of years for which hindcasts fell in each category; a perfect hindcast would contain 33% for each value in the diagonal (green) boxes and 0% in all other boxes.

important predictor for all onset definitions; zonal wind is also used to a lesser degree. SST is found to play a limited role in onset prediction relative to other variables; ENSO (i.e., SST in the South Pacific) is ruled out entirely by the GCV method, suggesting that more localized predictors play a larger role in onset variability. Pressure variables near the eastern Mediterranean and Red Sea are featured in all onset models at all lead times, which is consistent with the flood predictors of Jury (2011) and serves to demonstrate the blurring distinction of onset as a climate versus weather event.

In most metrics, forecasts using the yearly definition outperform those using the threshold definition. This may be attributable to its inherent connection with seasonal precipitation, which has been shown to be moderately to well predicted at similar lead times (e.g., Korecha and Barnston 2007; Alexander et al. 2019),

whereas the threshold definition is significantly more sensitive to particular precipitation events, making later issue dates more reliable. Thus, sectors that place importance on total seasonal precipitation, e.g., reservoir management, may prefer anomaly definitions whereas agriculture may favor threshold definitions as they are more closely related to soil moisture (MacLeod 2018). All onset definitions also tend to have higher RPSS scores for extreme events (defined as the earliest five and latest five historical onsets); the yearly definition's reach up to 96% for early and 70% for late. Hence, if abnormally early or late onsets are of particular interest, statistical forecasts, particularly using the yearly definition, may be of particular use.

The trade-off between lead time and performance for the threshold and yearly anomaly definitions also warrants consideration. Including additional climate

FIG. 9. Qualitative hindcasts, yearly definition, using (left) PLS regression and (right) random forest classification for 15 Mar (topmost value for each square) to 1 May (bottommost value) issue dates.

FIG. 10. Qualitative hindcasts, window definition, using (left) PLS regression and (right) random forest classification for 15 Mar (top value for each square) and 1 Apr (bottom value) issue dates.

information and issuing forecasts later (April or May) does not necessitate an improvement in performance. Mean absolute error and median RPSS consistently improve with later issue dates, but RPSS scores for extreme events do not monotonically increase, suggesting an improvement in accuracy but not precision, which may explain the relatively small changes in error bars in Figs. 5–7. Planting during a false onset, likewise, is less likely when using forecasts, but rates are only slightly reduced for later forecast issue dates. Missed onsets is also a telling statistic; although a 15 March or 1 April issue date may be inferior in terms of percent reduction in error or RPSS, these issue date always occurred prior to all onset dates (39 years), whereas 3%–23% of onsets have already occurred by the 15 April and 1 May issue dates. For the Koga case study, an 1 April yearly definition prediction is arguably the best-performing choice, striking a balance between lead time, forecast skill, and no missed onsets, however, this is not likely generalizable to other locations.

Regarding utility of these forecasts for agricultural planning, the risk of planting during a false onset is

substantially reduced when using later forecast issue dates. Although these later dates are at risk of being issued after onset has occurred, they still are issued before farmers will ever plant based on criteria A and B, and before all but two historical planting dates for criterion C. Issuing both earlier (March and April) and later (May) forecasts, therefore, may benefit farmers, as early forecasts can help farmers prepare for the season while later forecasts can inform specifically when to plant. The forecasts thus successfully reduce the risk of planting too early; however, planting too late also presents some risk to farmers by resulting in a shorter the growing season, which is not addressed in this study. This is particularly true for the window definition, which tends to have much later onset dates; this definition may be best suited for other crops that are planted later in the season.

To the authors' knowledge, this study is unique in examining statistical prediction models of Kiremt onset in Ethiopia. The results of this study compare similarly with dynamic model onset predictions (see

FIG. 11. Historical planting dates for each criterion.

TABLE 4. Number of historical false onsets based on naïve planting criteria and number of false onsets in same years when planting according to forecasts for various issue dates.

		Planting criteria		
		A	B	C
Threshold	Naïve planting	11	2	3
	15 Mar	5	2	0
	1 Apr	5	2	0
	15 Apr	5	2	0
	1 May	4	1	0
Yearly	Naïve planting	14	3	1
	15 Mar	6	2	0
	1 Apr	7	2	0
	15 Apr	9	3	0
	1 May	8	1	0
Window	Naïve planting	27	21	9
	15 Mar	10	10	5
	1 Apr	11	10	6

MacLeod 2018), although there are notable differences. The dynamic model (ECMWF SEAS5) provides a longer prediction lead time (1 February), but is coarser in resolution than CHIRPS (0.25° versus 0.05° , respectively), and exhibits a wide range of bias—from <7 days to >28 days using a window definition with CHIRPS data; the Koga region specifically illustrates a bias of 14–21 days, exceeding the bias stemming from the statistical forecast model (10–13 days)—for the Ethiopian Highlands. Also of note is the dynamic model's ability to forecast cessation with modest skill; a corresponding statistical forecast model was developed by the authors featuring similar and alternate oceanic–atmospheric predictor types, however, it produced only very limited predictive skill. While statistical methods can capture nuanced effects of climate signals at the local scale, they may also be inherently sensitive to small changes, evidenced by variations in predictors between issue dates for a given definition. The local focus of statistical models, moreover, limits their application to broad regions; evaluating the PLS model at other locations using the same set of predictors demonstrated limited skill, except at other areas within the western Ethiopian Highlands at similar elevations to Koga (Fig. S1). Consideration of both dynamical and statistical approaches may thus be advantageous, pairing early issue dates with location-specific predictions.

This study also considers qualitative forecasts, which may be uniquely received and acted on by individual stakeholders. Future research should therefore consider effective communication of forecasts, including the use of quantitative versus qualitative information, and extend these methods to agricultural and pastoral decision-making to better quantify expected value.

Acknowledgments. This material is based upon work supported by the National Science Foundation under Grant 1545874. The authors declare no conflicts of interest. The authors would also like to thank Berihun Adugna and Ezana Atsbeha for providing surveys from farmers concerning planting dates.

Data availability statement: All relevant input data and output scripts are available at <http://doi.org/10.5281/zenodo.3691948>.

REFERENCES

- Abdi, H., 2003: Partial least squares regression. *Encyclopedia of Social Sciences Research Methods*, M. Lewis-Beck, A. Bryman, and T. Futing, Eds., Sage, <https://doi.org/10.4135/9781412950589.n690>.
- Alexander, S., S. Wu, and P. Block, 2019: Model selection based on sectoral application scale for increased value of hydroclimate prediction information. *J. Water Resour. Plann. Manage.*, **145**, 04019006, [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001044](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001044).
- Araya, A., L. Stroosnijder, S. Habtu, S. D. Keesstra, M. Berhe, and K. M. Hadgu, 2012: Risk assessment by sowing date for barley (*Hordeum vulgare*) in northern Ethiopia. *Agric. For. Meteorol.*, **154–155**, 30–37, <https://doi.org/10.1016/j.agrformet.2011.11.001>.
- Berhane, F., and B. Zaitchik, 2014: Modulation of daily precipitation over East Africa by the Madden–Julian oscillation. *J. Climate*, **27**, 6016–6034, <https://doi.org/10.1175/JCLI-D-13-00693.1>.
- Breiman, L., 2001: Random forests. *Mach. Learn.*, **45**, 5–32, <https://doi.org/10.1023/A:1010933404324>.
- Cabot Venton, C., 2018: Economics of resilience to droughts: Ethiopia analysis. United States Agency for International Development, 43 pp., https://www.usaid.gov/sites/default/files/documents/1867/Ethiopia_Economics_of_Resilience_Final_Jan_4_2018_-_BRANDED.pdf.
- Craven, P., and G. Wahba, 1978: Smoothing noisy data with spline functions. *J. Numer. Math.*, **31**, 377–403, <https://doi.org/10.1007/BF01404567>.
- Dinku, T., C. Funk, P. Peterson, R. Maidment, T. Tadesse, H. Gadain, and P. Ceccato, 2018: Validation of the CHIRPS satellite rainfall estimates over eastern Africa. *Quart. J. Roy. Meteor. Soc.*, **144**, 292–312, <https://doi.org/10.1002/qj.3244>.
- Drechsler, M., and W. Soer, 2016: Early warning, early action: The use of predictive tools in drought response through Ethiopia's productive safety net programme. World Bank Policy Research Working Paper 7716, 42 pp., <https://doi.org/10.1596/1813-9450-7716>.
- Dunning, C. M., E. C. L. Black, and R. P. Allan, 2016: The onset and cessation of seasonal rainfall over Africa. *J. Geophys. Res. Atmos.*, **121**, 11 405–11 424, <https://doi.org/10.1002/2016JD025428>.
- Ewbank, R., C. Perez, H. Cornish, M. Worku, and S. Woldetsadik, 2019: Building resilience to El Niño-related drought: Experiences in early warning and early action from Nicaragua and Ethiopia. *Disasters*, **43**, S345–S367, <https://doi.org/10.1111/disa.12340>.
- Ferguson, C. R., M. Pan, and T. Oki, 2018: The effect of global warming on future water availability: CMIP5 synthesis. *Water Resour. Res.*, **54**, 7791–7819, <https://doi.org/10.1029/2018WR022792>.
- Funk, C. C., and Coauthors, 2014: A quasi-global precipitation time series for drought monitoring. U.S. Geological Survey Data Series 832, 4 pp., <http://pubs.usgs.gov/ds/832/>.

- Government of Ethiopia, 2019: Ethiopia's Climate Resilient Green Economy: National Adaptation Plan. United Nations Framework Convention on Climate Change, 147 pp., <https://www4.unfccc.int/sites/NAPC/Documents/Parties/NAP-ETH%20FINAL%20VERSION%20%20Mar%202019.pdf>.
- Hansen, J. W., A. Mishra, K. P. C. Rao, M. Indeje, and R. K. Ngugi, 2009: Potential value of GCM-based seasonal rainfall forecasts for maize management in semi-arid Kenya. *Agric. Syst.*, **101**, 80–90, <https://doi.org/10.1016/j.agry.2009.03.005>.
- Hastenrath, S., D. Polzin, and C. Mutai, 2011: Circulation mechanisms of Kenya rainfall anomalies. *J. Climate*, **24**, 404–412, <https://doi.org/10.1175/2010JCLI3599.1>.
- Jury, M. R., 2011: Meteorological scenario of Ethiopian floods in 2006–2007. *Theor. Appl. Climatol.*, **104**, 209–219, <https://doi.org/10.1007/s00704-010-0337-0>.
- Kaiser, H. F., 1960: The application of electronic computers to factor analysis. *Educ. Psychol. Meas.*, **20**, 141–151, <https://doi.org/10.1177/001316446002000116>.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643, <https://doi.org/10.1175/BAMS-83-11-1631>.
- Kipkorir, E. C., D. Raes, R. J. Bargerei, and E. M. Mugalavai, 2007: Evaluation of two risk assessment methods for sowing maize in Kenya. *Agric. For. Meteorol.*, **144**, 193–199, <https://doi.org/10.1016/j.agrformet.2007.02.008>.
- Korecha, D., and A. G. Barnston, 2007: Predictability of June–September rainfall in Ethiopia. *Mon. Wea. Rev.*, **135**, 628–650, <https://doi.org/10.1175/MWR3304.1>.
- Lanckriet, S., A. Frankl, E. Adgo, P. Termonia, and J. Nyssen, 2014: Droughts related to quasi-global oscillations: A diagnostic teleconnection analysis in North Ethiopia. *Int. J. Climatol.*, **35**, 1534–1542, <https://doi.org/10.1002/joc.4074>.
- Liben, M. F., S. C. Wortmann, and T. K. Fantaye, 2015: Risks associated with dry soil planting time in Ethiopia. *Int. J. Agron. Agric. Res.*, **6**, 14–23.
- MacLeod, D., 2018: Seasonal predictability of onset and cessation of the east African rains. *Wea. Climate Extremes*, **21**, 27–35, <https://doi.org/10.1016/j.wace.2018.05.003>.
- Moron, V., P. Camberlin, and A. W. Robertson, 2013: Extracting subseasonal scenarios: An alternative method to analyze seasonal predictability of regional-scale tropical rainfall. *J. Climate*, **26**, 2580–2600, <https://doi.org/10.1175/JCLI-D-12-00357.1>.
- Ng, K. S., 2013: A simple explanation of partial least squares. Tech. Rep., Australian National University, 10 pp., <http://users.cecs.anu.edu.au/~kee/pls.pdf>.
- Nicholson, S. E., 2017: Climate and climatic variability of rainfall over eastern Africa. *Rev. Geophys.*, **55**, 590–635, <https://doi.org/10.1002/2016RG000544>.
- Segele, Z. T., and P. J. Lamb, 2005: Characterization and variability of Kiremt rainy season over Ethiopia. *Meteor. Atmos. Phys.*, **89**, 153–180, <https://doi.org/10.1007/s00703-005-0127-x>.
- , M. B. Richman, L. M. Leslie, and P. J. Lamb, 2015: Seasonal-to-interannual variability of Ethiopia/Horn of Africa monsoon. Part II: Statistical multimodel ensemble rainfall predictions. *J. Climate*, **28**, 3511–3536, <https://doi.org/10.1175/JCLI-D-14-00476.1>.
- Tadross, M., and Coauthors, 2009: Growing-season rainfall and scenarios of future change in southeast Africa: Implications for cultivating maize. *Climate Res.*, **40**, 147–161, <https://doi.org/10.3354/cr00821>.
- Tsidu, G. M., 2012: High-resolution monthly rainfall database for Ethiopia: Homogenization, reconstruction, and gridding. *J. Climate*, **25**, 8422–8443, <https://doi.org/10.1175/JCLI-D-12-00027.1>.
- Viste, E., and A. Sorteberg, 2013: Moisture transport into the Ethiopian highlands. *Int. J. Climatol.*, **33**, 249–263, <https://doi.org/10.1002/joc.3409>.
- Wambui, C. 2019: Kenyan farmers trust tradition over tech to predict the weather. Thomas Reuters Foundation, 11 February, <https://news.trust.org/item/20190211065115-ywdmq/?source=spotlight>.
- Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences: An Introduction*. International Geophysics Series, Vol. 59, Academic Press, 467 pp.
- Wold, S., A. Ruhe, H. Wold, and W. J. Dunn, 1984: The collinearity problem in linear regression. The Partial Least Squares (PLS) approach to generalized inverses. *SIAM J. Sci. Statist. Comput.*, **5**, 735–743, <https://doi.org/10.1137/0905052>.
- Yang, W., R. Seager, M. A. Cane, and B. Lyon, 2015: The annual cycle of East African precipitation. *J. Climate*, **28**, 2385–2404, <https://doi.org/10.1175/JCLI-D-14-00484.1>.
- Zhang, Y., S. Moges, and P. Block, 2016: Optimal cluster analysis for objective regionalization of seasonal precipitation in regions of high-spatial-temporal variability: Application to western Ethiopia. *J. Climate*, **29**, 3697–3717, <https://doi.org/10.1175/JCLI-D-15-0582.1>.