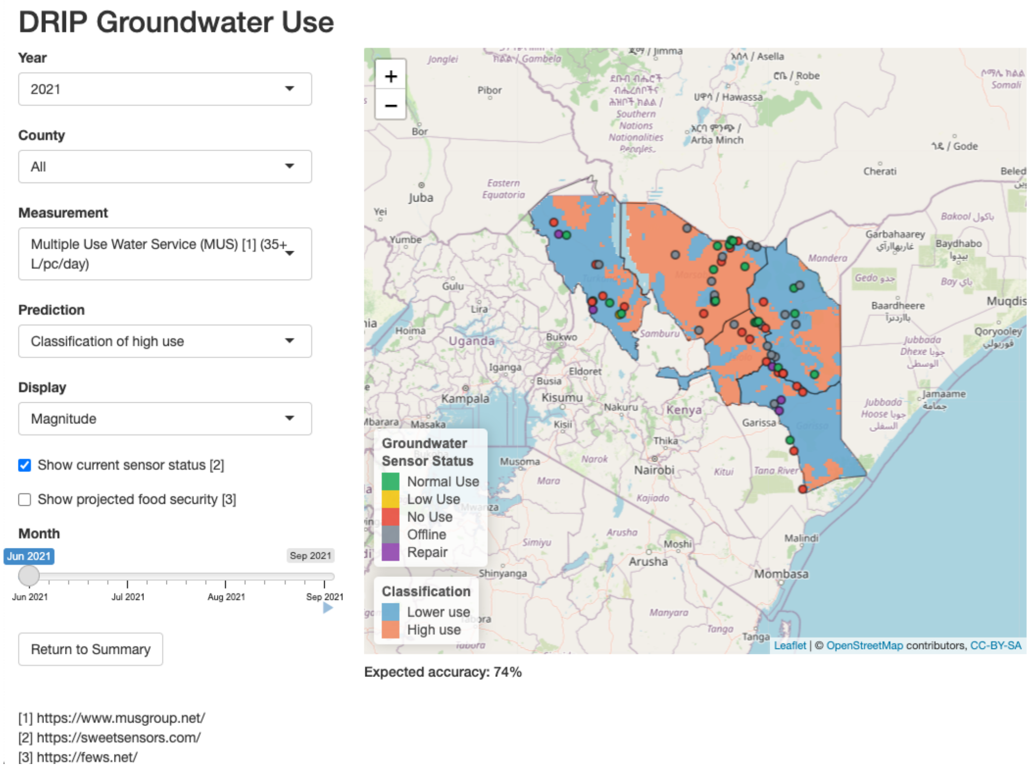
Graphical Abstract

Estimating groundwater use and demand in arid Kenya through assimilation of satellite data and in-situ sensors with machine learning toward drought early action

Katie Fankhauser, Denis Macharia, Jeremy Coyle, Styvers Kathuni, Amy McNally, Kimberly Slinski, Evan Thomas



Highlights

Estimating groundwater use and demand in arid Kenya through assimilation of satellite data and in-situ sensors with machine learning toward drought early action

Katie Fankhauser, Denis Macharia, Jeremy Coyle, Styvers Kathuni, Amy McNally, Kimberly Slinski, Evan Thomas

- Supporting local groundwater management would improve resilience to drought.
- Kenya groundwater use estimated with sensors, satellite data, and machine learning.
- Historical use was modeled with up to 75% cross-validated accuracy.
- Forecasts for the 2021 dry season indicated up to 80% external validation accuracy.

Estimating groundwater use and demand in arid Kenya through assimilation of satellite data and in-situ sensors with machine learning toward drought early action

Katie Fankhauser[a], Denis Macharia[a,b,c], Jeremy Coyle[a,d], Styvers Kathuni[d], Amy McNally[e,f,g], Kimberly Slinski[e,h], Evan Thomas[a,d,*]

[a] *Mortenson Center in Global Engineering, University of Colorado Boulder, 4001 Discovery Drive, Boulder, 80303, Colorado, USA*
[b] *Regional Centre for Mapping of Resources for Development, Nairobi, Kenya*
[c] *Colorado, USA*
[d] *SweetSense Inc., USA*
[e] *NASA Goddard Space Flight Center, Greenbelt, Maryland, USA*
[f] *Science Applications International Corporation, Reston, Virginia, USA*
[g] *US Agency for International Development, Washington, DC, USA*
[h] *University of Maryland Earth Systems Science Interdisciplinary Center, College Park, Maryland, USA*

*Corresponding author
Email address: evan.thomas@colorado.edu (Evan Thomas)

Abstract

Groundwater is an important source of water for people, livestock, and agriculture during drought in the Horn of Africa. In this work, areas of high groundwater use and demand in drought-prone Kenya were identified and forecasted prior to the dry season. Estimates of groundwater use were extended from a sentinel network of 69 in-situ sensored mechanical boreholes to the region with satellite data and a machine learning model. The sensors contributed 756 site-month observations from June 2017 to September 2021 for model building and validation at a density of approximately one sensor per 3,700 km^2. An ensemble of 19 parameterized algorithms was informed by features including satellite-derived precipitation, surface water availability, vegetation indices, hydrologic land surface modeling, and site characteristics to dichotomize high groundwater pump utilization. Three operational definitions of high demand on groundwater infrastructure were considered: 1) mechanical runtime of pumps greater than a quarter of a day (6+ hrs) and daily per capita volume extractions indicative of 2) domestic water needs (35+ L), and 3) intermediate needs including livestock (75+ L). Gridded interpolation of localized groundwater use and demand was provided from 2017 to 2020 and forecasted for the 2021 dry season, June - September 2021. Cross-validated skill for contemporary estimates of daily pump runtime and daily volume extraction to meet domestic and intermediate water needs was 68%, 69%, and 75%, respectively. Forecasts were externally validated with an accuracy of at least 56%, 70%, 72% for each groundwater use definition. The groundwater maps are accessible to stakeholders including the Kenya National Drought Management Authority (NDMA) and the Famine Early Warning Systems Network (FEWS NET). These maps represent the first operational spatially-explicit sub-seasonal to seasonal (S2S) estimates of groundwater use and demand in the literature. Knowledge of historical and forecasted groundwater use is anticipated to improve decision-making and resource allocation for a range of early warning early action applications.

Keywords: drought, groundwater, early warning, early action, machine learning, remote sensing, Kenya

1. Introduction

Drought is one of the most persistent, expansive, and damaging of natural disasters. Coupled with weak institutional and civil systems, experience of drought leads to regional food insecurity, water insecurity, disease, and violent conflict for billions of people globally (Wilhite and Glantz, 1985). The Horn of Africa is susceptible to natural exposure and societal vulnerability to drought (Liebmann et al., 2017; NDMA, 2015). As a slow onset disaster, public and private organizations may not effectively mobilize resources in response (Cabot Venton, 2018). An advantage, however, is that the risk of drought can be monitored and forecasted in advance, allowing for the implementation of early warning early action systems (Funk and Shukla, 2020).

Drought early warning is an established field with existing indices that rely on climatological and land surface modeling or observing causal factors. For example, the United States National Integrated Drought Information System (NIDIS) [drought.gov] produces current condition and 3-month drought outlooks from weather and climate, soil moisture, streamflow, and rainfall and snowpack data informed by qualitative expert opinion. In lower resource settings, such as Sub-Saharan Africa, observational data is sparse and satellite-derived and land surface modeled estimates of climate and the hydrologic cycle become the main inputs in drought monitoring. The Famine Early Warning Systems Network (FEWS NET) [fews.net] makes the link between drought and food insecurity in low- and middle-income countries explicit.

Drought develops from a deficit of precipitation and water storage inadequate to support livelihoods. The four domains of drought – meteorological, hydrological, agricultural, and socioeconomic – compound from an initial lack of rainfall and lowered water availability to adverse impacts on land productivity and livestock and, ultimately, to water and food insecurity from diminished household returns (Wilhite and Glantz, 1985). The primary rainfall and agricultural season in East Africa has seen a marked decrease in precipitation in recent decades (Nicholson et al., 2018), leading to slower and lower recharge of surface water and diminished flow rates in large rivers. Moreover, the increase in surface temperatures from climate change, generates higher evapotranspiration, aggravating the decline in precipitation and intensifying aridification.

The arid and semi-arid lands (ASALs) of Kenya have faced regular drought since at least 2016 with below average rainfall placing 18 million people at risk (UNICEF, 2017; KNBS, 2019). During times of low precipitation and decreased surface water capacity, groundwater becomes a critical source of water for domestic, livestock, and agricultural needs (Thomas et al., 2019). Pastoralist communities, in particular, are dependent on predictable access to water, without which they are likely to continue to experience inequitable rates of poverty and poor health (NDMA, 2015). Functional and strategic groundwater distribution, then, may lower exposure and vulnerability to drought for affected populations.

For the region, a number of databases are available that address water supply (Hofste et al., 2019; Senay et al., 2013) and demand (Hofste et al., 2019; McNally et al., 2019; Senay et al., 2015). Although these data products often incorporate some information about groundwater, they rely on Gravity Recovery and Climate Experiment (GRACE) satellite data at low resolution (10000 km^2 - 150000 km^2) (Landerer and Swenson, 2012) or mechanistic hydrological modeling without other primary data. AQUASTAT (FAO, 2021) is a rich resource of water data, but it is only aggregated from annual questionnaires in tabulated form by country every 5 years, limiting its usability at finer spatial scales and its relevance to real- or near-time response. Therefore, a geospatially explicit dataset that gives estimates of local (<50 square kilometers) and acute (sub-seasonal) realized groundwater demand represents an important missing indicator of water resource availability and withdrawal that would improve interpretation of other water and drought indices in the region.

The Drought Resilience Impact Platform (DRIP) combines early detection with proactive groundwater management to ensure water availability with various stakeholders and decision support tools (Thomas et al., 2021). A principal feature of the platform is remote monitoring of a sentinel network of mechanical groundwater boreholes in drought-prone regions of East Africa. The sensors have been used to identify non-functionality (Thomas et al., 2021), improve water service delivery and increase water infrastructure uptime (Nagel et al., 2015; Wilson et al., 2017), demonstrate an inverse relationship between rainfall and borehole use (Thomas et al., 2019), and examine a framework of groundwater management for drought resilience (Turman-Bryant et al., 2019). The current study explicitly addresses how monitoring data from the sensors can be used to quantify groundwater and improve delivery of drought early warning services and actions for the entire region.

Although groundwater is an essential water source, especially during the dry season, during the 2016-2017 drought, 55% of the pumps needed to access groundwater were non-functional in Kenya (UNICEF, 2017). The ability to direct limited resources to repair, maintain, or site the most critically needed boreholes based on projected use prior to drought would enable responsible stewardship of water resources and community resilience to drought. Interpreting use within known limitations of existing infrastructure may also help direct water resources other than groundwater, such as emergency water trucking. Furthermore, observing and predicting trends in groundwater use could be an important indicator of developing drought itself.

By leveraging additional remote sensing and in-situ data sources, we developed gridded estimates for historical and forecasted groundwater use in five prominent ASAL counties in Kenya. A statistical machine learning model extrapolates recorded borehole operation to locations that are not instrumented. The final product categorizes areas into high vs. low groundwater use based on definitions of pump runtime and per capita volume extraction. This paper describes development of the DRIP groundwater service and presents the first gridded maps of acute groundwater use and demand in this context. It is expected to inform ongoing local management and planning for water and food insecurity during drought in Kenya and demonstrate this for East Africa.

2. Methods

2.1. Context

This work builds upon efforts of the Kenya Resilient Arid Lands Partnership for Integrated Development (Kenya RAPID) program. In-situ data was collected at all mechanized groundwater pumps identified as critical drought infrastructure and localized groundwater estimates were developed for five ASAL counties: Garissa, Isiolo, Marsabit, Turkana, and Wajir (Figure 1). Collectively the five counties cover approximately 260,000 square kilometers - or nearly half of the total surface area of Kenya - and suffer from water and food insecurity, high poverty rates, and limited access to basic services (NDMA, 2015). Despite several moderate to high productivity sedimentary aquifers underlying the area, only 17% of renewable groundwater resources are being utilized (Mumma et al., 2011). The ASALs are expected to experience more frequent and greater duration dry periods, likely exacerbating existing disadvantage in the absence of sustainable development and dedicated drought risk reduction.

The Government of Kenya has committed to ending drought emergencies (EDE) by 2022. The National Drought Management Authority (NDMA), established in 2011, was directed to address drought proactively by confronting vulnerability and structural causes (NDMA, 2015). Cooperation between local county water officials and NDMA officials led to the selection of strategic boreholes for drought response that obtain financial and operational support. These boreholes in the five programmatic counties received sensor installation and monitoring. Previously, we observed that the strategic boreholes were more likely to be non-functional and susceptible to demand elasticity (Turman-Bryant et al., 2019). Drought triggers the mobilization of financial and material resources for maintenance of NDMA priority boreholes and the execution of contingency plans. In years where precipitation is higher, national and local resources for assessing drought risk and repair are withheld or reallocated, leading to an increase in the number and duration of malfunctions. The capacity to use strategic boreholes to mitigate future droughts is diminished by a smaller inventory of operating sites and increasing strain on remaining infrastructure. This suggests that resources should consistently be directed to critical boreholes and to areas with restricted surface water availability or where demonstrable demand is already high.

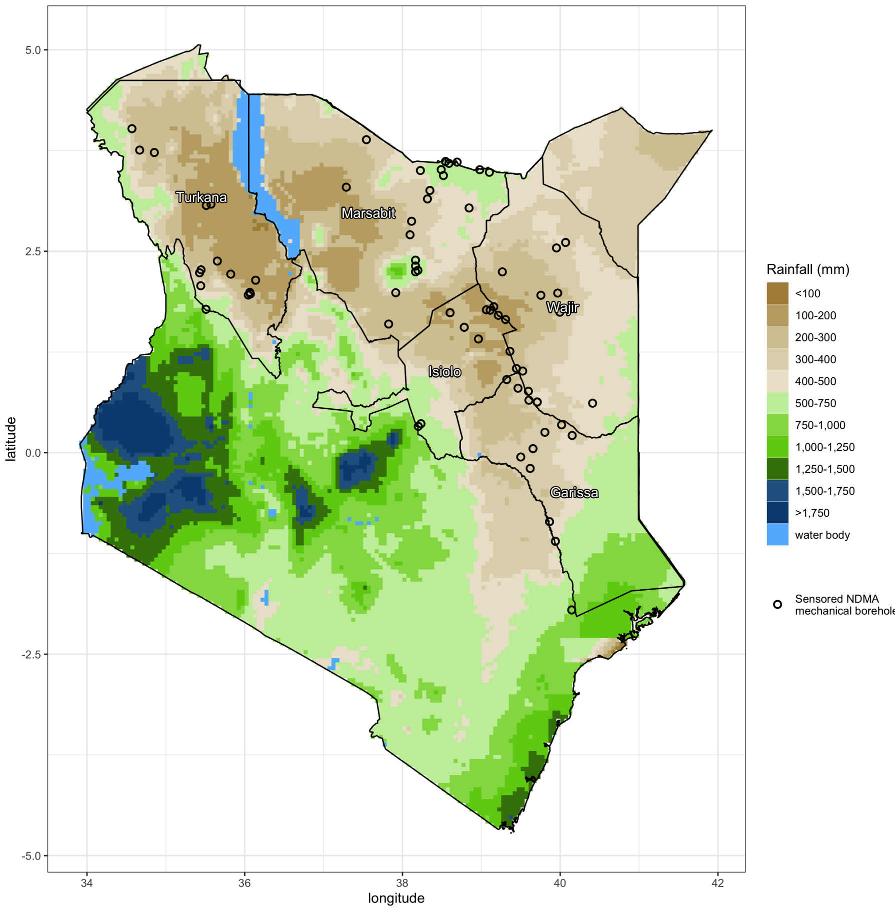


Figure 1: Average annual precipitation, 1981 - 2019. The locations of sensor-monitored NDMA (National Drought Management Authority) strategic drought emergency boreholes are shown in the five programmatic arid and semi-arid counties of Kenya. Data source: Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS) v2.0

2.2. Data

2.2.1. In-situ data

From 2016 to present, 238 mechanized groundwater pumps in Kenya have been monitored by satellite- and cellular-connected sensors provided by SweetSense Inc.

(sweetsense.space). Pressac brand wireless current transformers measure the frequency and duration of borehole use by logging the electrical current (amperage) delivered to the mechanical pump over time. Data is transmitted daily to a central database and web application (Thomas et al., 2021). Sub-hourly sensor data was aggregated to mean daily pump runtime and volume by month.

NDMA selected the EDE strategic boreholes to prioritize for operations during drought. Further, data was subset to months during Kenya's main dry season each year – June through September – as the annual period when drought is of biggest concern and deliberately surveilled. In-situ measurements from representative NDMA EDE boreholes were included in the sample when the mechanical borehole and sensor were reasonably known to be operational. Borehole and sensor status were assessed from maintenance surveys and the availability of daily data.

The final sample included 69 unique mechanical borehole sites (Figure 1) and 565 site-month observations from June 2017 to September 2020. The four months of the 2021 dry season (June - September 2021) contributed 191 site-month observations from 51 unique sites for external validation.

In order to provide comprehensive mapping of groundwater resources for different stakeholders and applications, three separate, but complementary, binary high groundwater use classifications were made from the sensor data. Each specification informs an aspect of groundwater accessibility, including physical infrastructure and population-based need. They are applied to the project region to be considered singularly or in comparison.

The three definitions of categorical high groundwater use by month translated from the sensor data are: 1) mean daily pump runtime over six hours per day (a quarter day) and mean daily per capita volume extracted 2) over 35 liters per person per day (L/pc/d) and 3) over 75 liters per person per day (L/pc/d). From the electrical current delivered to the pump, the number of hours a pump was switched on per day was measured and taken as pump runtime (Thomas et al., 2021). Estimated flow rate at the pump in cubic meters per hour was collected during site visits, and taking the product of runtime and yield provided estimates of volume of water extracted per day.

The high use threshold for pump runtime was set at approximately the observed mean while high per capita volume of over 35 L/pc/d was meant to capture basic needs on the multiple-use water services (MUS) ladder. Between 20 and 50 L/pc/d, users are expected to be able to meet most domestic and some livestock and personal agricultural needs (MUS Group, 2013). Additionally, a threshold of over 75 L/pc/d was considered to capture the higher water demands of pastoralists with livestock and represent an intermediate service level on the MUS ladder. Thus, locations of groundwater use at or above the high runtime or high volume thresholds indicate, respectively, areas where boreholes are expected to run longer than average or the population is relying on groundwater to meet domestic and livestock needs, presumably because other water resources are unavailable.

2.2.2. Remote sensing data

The variables taken or derived from satellite data were elevation (RCMRD Geoportal, 2015); precipitation (Funk et al., 2015); surface water availability (Senay et al., 2013); baseflow, runoff, soil moisture, and actual evapotranspiration (Case et al., 2014); Normalized Difference Vegetation Index [NDVI] (USGS FEWS NET Data Portal, 2017); greenness vegetation fraction [GVF] (Vargas et al., 2015); and population density (Stevens et al., 2015). More detail about the actual nature of the features is given below in section 2.4.

Data was retrieved from January 2016 to May 2021 to cover the period of sensor-recorded observations of borehole use and accommodate lagging of time-varying covariates for forecasting. Different temporal resolutions were ultimately averaged or summed by month. Native spatial resolution of the satellite data products ranged from 250 m to 50 km, and values were extracted at point coordinates of the sensors or centroids of a 0.05° x 0.05° ($\sim$30 km^2) grid over the five counties. Higher resolution data were resampled using bilinear interpolation to agree with the reference grid.

Gridded population data at 1 km resolution was retrieved from WorldPop [worldpop.org] (Stevens et al., 2015; Lloyd et al., 2019) to scale sensor-derived groundwater volume extraction by local population numbers. Volume per capita has been accepted as a proxy for water demand in previous studies (McNally et al., 2019).

2.3. Supervised machine learning

Supervised machine learning (ML) models are empirical. They optimize model fit to achieve the most similar outcomes to input training data without regard to deterministic or mechanistic explanations. Predictions are derived from the fit applied to new observations. Ensemble ML (also known as Super Learning or model stacking) executes multiple models (”learners”) on the same data and selects the optimal combination of them through cross-validation (van der Laan et al., 2007). Our library consisted of 19 candidate learners: nine XGBoost algorithms with different hyperparameter selections (Chen and Guestrin, 2016); logistic regression; a Bayesian generalized linear model (Gelman et al., 2008); Ridge regression, LASSO regression, and another elastic net regularization with an alpha parameter equal to 0.5 (Friedman et al., 2010); three k-Nearest Neighbors learners (Mouselimis, 2021) considering varying numbers of neighbors; Random Forests (Breiman, 2001); and a null model.

Covariates, or features, are the independent variables that have an assumed impact on the outcome (see section 2.4). Given the inclusion of algorithms in the ensemble that perform feature selection in addition to a pre-screening procedure, a large number of features, including those that are correlated, can be considered without overfitting the model. Overfitting occurs when a model describes training data well, but cannot generalize to new data. Feature selection and pre-screening reduce the number of features to improve parsimony and cross-validation retains independence between training and testing datasets so that the model can predict new observations of groundwater use with accuracy. Furthermore, ensemble models are proven to perform as well as or better than any one candidate learner in the ensemble and minimizing cross-validated risk controls for overfitting even when a large number of learners are considered (van der Laan et al., 2007; Polley and Laan, 2010).

The purpose of modeling in this study was twofold: to extend runtime and volume predictions spatially and forward in time. To provide predictions across the region at a resolution of approximately 30 km^2, a model trained and validated at instrumented boreholes was fit on independent explanatory features observed at centroid point coordinates of a $0.05° \times 0.05°$ grid over the five programmatic counties. To extend the predictions forward in time, forecasting up to four months was achieved by lagging the features by the same number of months and individual model runs. Since the effect of climatological phenomenon on groundwater supply and demand may be delayed itself, we provided the model with all lagged features available at the respective forecast period and allowed the machine learner to screen and select the best performing lag for each time varying feature separately. For each outcome, the corresponding final model yielded a predicted probability that a certain centroid during a particular month was in a high use category as defined above. To establish binary high vs. low use, a threshold probability was selected based on a trade-off between true positive (high use borehole confirmed by the sensor record) and true negative (low use borehole confirmed by the sensor record) rates (see section 2.5).

To replicate operational conditions for the groundwater use forecasts and evaluate performance, we constrained model building between 2017 and 2020 and projected groundwater use forward for the 2021 dry season, June through September. Reviewing anticipated groundwater use alongside current infrastructure functionality prior to the dry season should inform investments and actions taken by local water officers and national decision-makers to reduce drought risk. Moreover, we evaluated the skill of our forecasts, which serve as a source of external validation for future estimates.

Modeling and feature design of contemporary groundwater estimates respected an internal five-fold stratified cross-validation structure. In cross-validation the observed data is sequentially partitioned into independent training and testing subsets. Five equally sized subsamples were generated randomly with all time series observations from one sensor being grouped in the same partition. The model was trained with four of the subsets and tested on the remaining, held-out subsample. This was repeated a total of five times ("folds") with each of the subsamples being used once as the testing dataset. Cross-validation allows for the calculation of performance statistics and the ability to generalize the model to new data.

All data management and analysis was conducted in R statistical computing software (R Core Team, 2020).

2.4. Features

The following variables are conventionally expected to influence drought and water insecurity and were included in our study: precipitation, land surface temperature, vegetation indices, soil moisture profile, evapotranspiration, other hydrologic system dynamics such as baseflow and runoff, population density, geographic location (longitude, latitude, and elevation), and seasonality (Funk and Shukla, 2020).

The Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) v2.0 (Funk et al., 2015) is a 30+ year quasi-global rainfall dataset that incorporates satellite imagery with available in-situ station data to create 0.05° x 0.05° gridded rainfall time series. However, station data is inconsistent and inadequate over Africa. The Trans-African Hydro-Meteorological Observatory [TAHMO] (tahmo.org) is gradually enhancing weather data collection with the installation of low-cost stations across the continent (Giesen et al., 2014). Currently, they operate over 500 stations in 25 countries, including 130 stations in Kenya. To leverage the advantages of both precipitation datasets — the coverage and legacy of CHIRPS with the fine resolution in-situ data of TAHMO — we used inverse distance weighting (IDW) to interpolate station data and combined the output with the satellite estimates at monthly time-scale through a Simple Bias Adjustment (SBA) merging method. The interpolation was done using a maximum correlation distance of 250 km, achieving a minimum and maximum number of stations within this distance of two and ten, respectively. If there were no stations in the vicinity to correct the satellite estimate, the underlying CHIRPS grid value was retained, resulting in a spatially consistent bias-corrected gridded product. We adopted this method from the ENACTS merging process (Dinku et al., 2014).

Gridded maximum daily temperature from the Global Telecommunication System (GTS) of the World Meteorological Organization (WMO) was retrieved and averaged by month (Physical Sciences Laboratory). NDVI from the Collection 6 Moderate Resolution Imaging Spectroradiometer (MODIS) instrument on the Aqua satellite was taken from processed data distributed by the United States Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center (USGS FEWS NET Data Portal, 2017). GVF from the Visible Infrared Imaging Radiometer Suite (VIIRS) instrument aboard the Suomi National Polar-orbiting Partnership (NPP) and National Oceanic and Atmospheric Administration (NOAA)-20 satellites (Vargas et al., 2015) was based on the differences between the enhanced vegetation index (EVI) for bare soils and dense vegetation.

Land surface modeling (LSM) is a method to extend the coverage of hydrologic observations in data sparse contexts, such as remote and developing locations. The Noah-Multiparameterization (Noah-MP) LSM relies on water and energy balances to physically describe land-atmosphere interaction processes under multiple specifications (Niu et al., 2011). Mean daily soil moisture (m^3/m^3) at depths of 0 - 10 cm, 10 - 40 cm, 40 - 100 cm, and 100 - 200 cm and total monthly actual evapotranspiration (mm), baseflow (mm), and runoff (mm) were derived from local instances of the Noah-MP LSM run per month (Case et al., 2014).

Population density was downloaded from WorldPop (worldpop.org) gridded population data from combined geospatial and census data (Stevens et al., 2015; Lloyd et al., 2019). Data is provided only through 2020; for 2021, we applied a population growth rate of 2.3% to the 2020 data (The World Bank, 2021). Population density was used as an independent variable when modeling pump runtime, but for per capita volume predictions, it was only used to scale the outcome and not as a model feature.

Other variables were included based on a demonstrable or logical relationship to groundwater use and demand. Public water management in Kenya is deregulated to the county level, leading to substantial differences in institutional and operational practices; thus, a county variable was added to describe unmeasured variability in groundwater supply and use due to administrative differences. Generally, determinants of the magnitude of groundwater use will be supply, demand, and access. Regionally interpolated relative surface water availability (Senay et al., 2013) was included to express the inverse correlation between rainfall, and by extension surface water supply, and borehole use (Thomas et al., 2019; Thomson et al., 2019). From a data inventory collated by Acacia Water as part of the Kenya RAPID program (kenyarapid.acaciadata.com/map/13) we also considered the presence within 10 km of floodplains, lakes, springs, or river basins. Proximity to installed water infrastructure, such as dams, sand pans, and other boreholes, may limit the use at any one groundwater source, and, therefore, counts of operational infrastructure within 10 km were created. Calculated domestic water demand in 2015 provides an estimate of volume per square kilometer per day ($m^3/km^2/d$) by sub-county required for adequate human consumption (Tolk et al., 2016). Social and economic networks can be demonstrated by access to main roads, markets, or towns, or, in this case, a count of the number of these amenities within 10 km.

Livestock are one of the largest consumers of water in the ASALs. Pastoralists commonly direct their herds to boreholes along migration routes so the presence of a major livestock migration route within 10 km was added as a feature. Estimated water demand for livestock ($m^3/km^2/d$) at the sub-county level was also included (Tolk et al., 2016).

We generated several variables from the sensor data to describe typical use behavior of boreholes in the region. One, we linearly regressed mean daily pump runtime or volume on static site characteristics, including longitude, latitude, elevation, county, water demand, proximity to natural water sources, other water infrastructure, and amenities, and livestock movement. Two, to leverage the sensor network and characterize the behavior of neighboring boreholes, we computed mean use at the five closest, by Euclidean distance, sensored sites as well as an interpolated layer of mean use at the maximally correlated site despite geographic location.

An advantage of using an ensemble learner and cross-validation is the ability to consider many features, models, and settings at once with less constraint on assuring a parsimonious model arising from a priori selection of the best set of variables and estimator (Polley and Laan, 2010). Thus, many social, environmental, and economic dimensions of drought risk and vulnerability were included. Ultimately, the ML was provided approximately 30 features from which each algorithm was built after screening.

See Appendix A for a detailed summary of all model features.

2.5. Statistical Evaluation

Model performance is evaluated by Receiver Operating Characteristic (ROC) curves and area under the curve (AUC), cross- and externally validated accuracy and rates of misclassification, and reliability diagrams. ROC curves are a performance measurement for classification problems. The ROC curve is a plot of sensitivity (y-axis) – the proportion of correctly identified instances of high use – against one minus specificity (x-axis) – the proportion of wrongly identified cases of high groundwater use among observed low use. The AUC demonstrates how good the model is at discriminating between high and low use. Determining which threshold to set when translating predicted probabilities to categories is always a trade-off between sensitivity and specificity. Youden's J statistic identifies the closest point on the ROC curve to the uppermost left corner of the plot and assumes both are equally important.

Accuracy is the percent number of sites where predicted high use correctly corresponds to observed high use after assigning predicted probabilities to above or below a set threshold. The null or no information rate (NIR) is the expected accuracy given random assignment based on the prevalence of the outcome in the data alone. Instances of incorrectly identified high use are evaluated with the false positive rate (FPR) and false negative rate (FNR) rate. The FPR is the proportion of wrongly identified cases of high groundwater use among observed low use. The FNR is relative to the number of misidentified low use among actual instances of high use.

Reliability is assessed from a linear relationship between predicted probability and observed frequency of high use relative to the magnitude of the probability. A perfectly reliable model exhibits a 1:1 relationship between the two.

3. Results

3.1. High resolution maps of groundwater use and demand

The ensemble machine learning models allowed for the creation of high spatial and temporal resolution maps of groundwater use and demand. For each dry month June 2017 - September 2020 contemporary estimates of pump runtime and volume extracted at existing or hypothetical groundwater infrastructure are available at approximately 30 km^2 resolution. The most recent dry season, June - September 2021, was forecasted and validated with ensuing observed data.

Figure 2 demonstrates how predicted probability output from the model was translated to high use for domestic needs ($>$ 35 L/pc/d). Red intensity indicates a higher probability of high use and ultimate high use class assignment is outlined. These maps may also suggest the certainty of high use classification. At a probability between 80-100%, we are relatively more confident groundwater use will surpass the high use threshold; if it is closer to 40-60%, we recognize that the model has identified one use category, but it may disagree with observed values on occasion.

Site-level accuracy for each 2021 month (72%, 74%, 77%, and 70%) showed that performance was relatively stable over different forecast periods, although it was lowest in September 2021 when the model had a lead of four months. High use was overestimated (i.e., FPR $>$ FNR) in June and September and underestimated in July and August (i.e., FNR $>$ FPR). When considering all years, specificity was high relative to sensitivity (Figure 4), indicating that generally the rate of false positives was lower than the rate of false negatives despite forecast lead time.

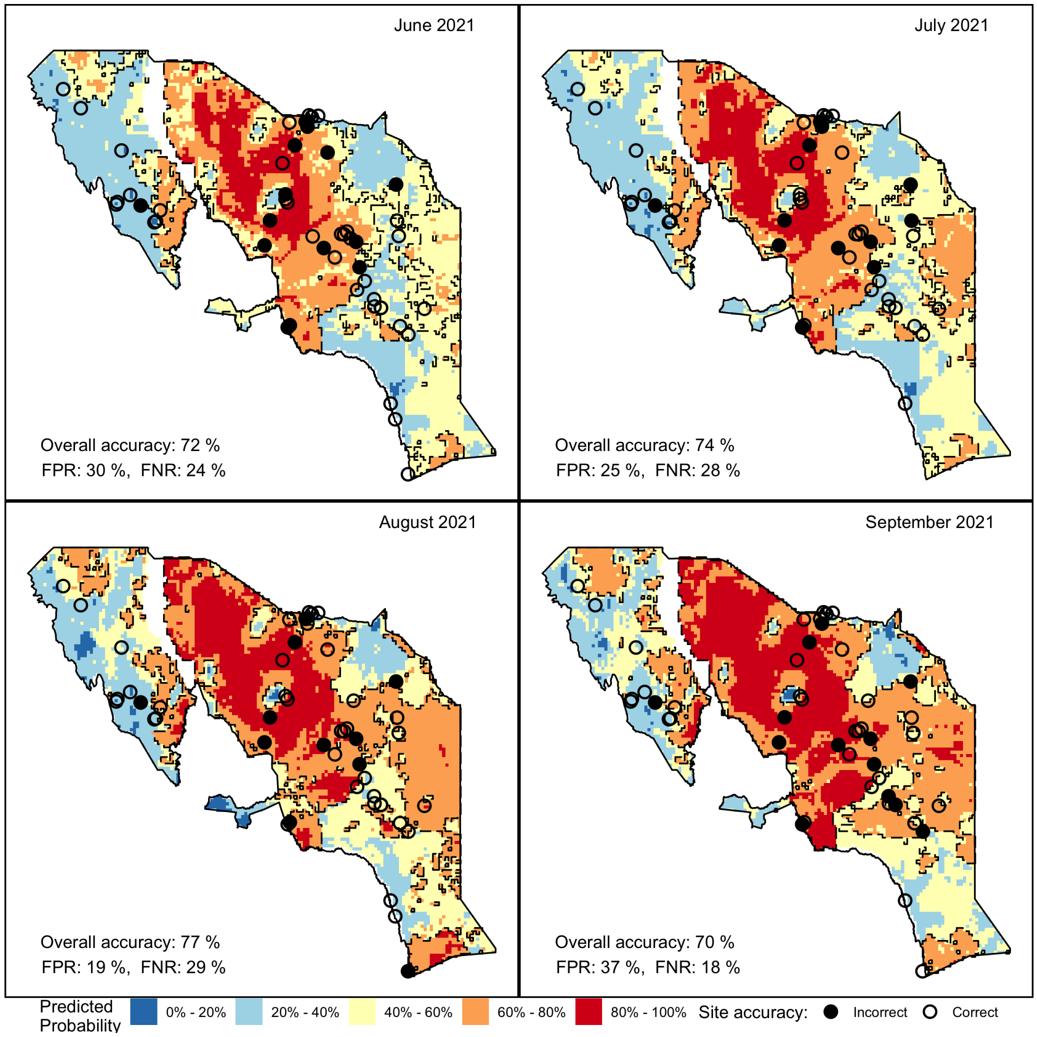


Figure 2: Predicted probability of groundwater demand for daily domestic water needs (>35 L per person per day) forecasted for the 2021 dry season. Categorization of high demand is outlined and accuracy is given for each dry season month. Rates of false positives (FPR) and false negatives (FNR) are based on site-level observations.

A similar demonstration was made for intermediate groundwater volume (> 75 L/pc/d) (Figure 3). Areas of northern Turkana County and southwestern Garissa County showed diminished need for higher volumes indicative of livestock watering. This was somewhat unexpected as these areas do support livestock. However, these areas are also places of violent conflict. So while there may be a high need for water, we anticipate some groundwater points in these region are demonstrating lower user due to insecurity limiting access. There is unmet demand for water and rehabilitating water points or providing alternative water sources could serve to improve resource-based conflict in the area. Site accuracy was high (73-80%) and mis-identification of false negatives was more likely than false positives across the season.

Groundwater demand increased as the dry season progressed and was consistently highest in Marsabit and Isiolo counties for both domestic and livestock needs. Here only per capita volume was shown, but categories of pump runtime displayed similar patterns.

All maps are accessible via a web-based platform at drip.shinyapps.io/groundwater. We presented only the pre-drought season 2021 volume per capita forecasts, but through the application, users can view historical and projected groundwater use. Other utilities include comparison of trends by month or year, viewing the two-class categorization of high groundwater use informed by the predicted probability, displaying the predictions as absolute magnitude or difference from averages, filtering by county, and seeing the status of the sensored boreholes.

For the purposes of this paper, we constrained model building prior to 2021 in order to externally validate the forecasts. In practice, the model and application will be updated per month during the dry season with an approximate one-week delay, at which time the concurrent groundwater estimates and forecasts roll forward one month. The model building procedure is repeated with the additional time series so the size and variability of the data will increase and we expect performance to improve over time with a longer data record.

3.2. Performance of gridded contemporary and forecasted estimates

The models achieved high statistical performance (Table 1, Figure 4). Estimation of volume per capita for domestic and intermediate needs proceeded the best.

The AUC for determining and forecasting high groundwater use ranged from 0.703 to 0.714, from 0.746 to 0.756, and from 0.778 to 0.787 for daily pump runtime, domestic volume per capita, and intermediate volume per capita, respectively (Table 1). Regardless if groundwater use is defined by mechanical requirements of the borehole or by water volume, our models contain a 70% or greater probability of accurately discriminating between low and high use. Ultimately, after classifying high use from predicted probabilities, the sensitivity for contemporary estimates, by use definition, was 58%, 57%, and 55% with corresponding specificity of 75%, 82%, and 88%. The rates for the one- to four-month forecasts were similar and can be read from Figure 4A. Since specificity is greater than sensitivity, it is less likely that a site predicted to exhibit high groundwater use would be incorrectly identified and there are less false positives than false negatives in the predictions. Thus, our models

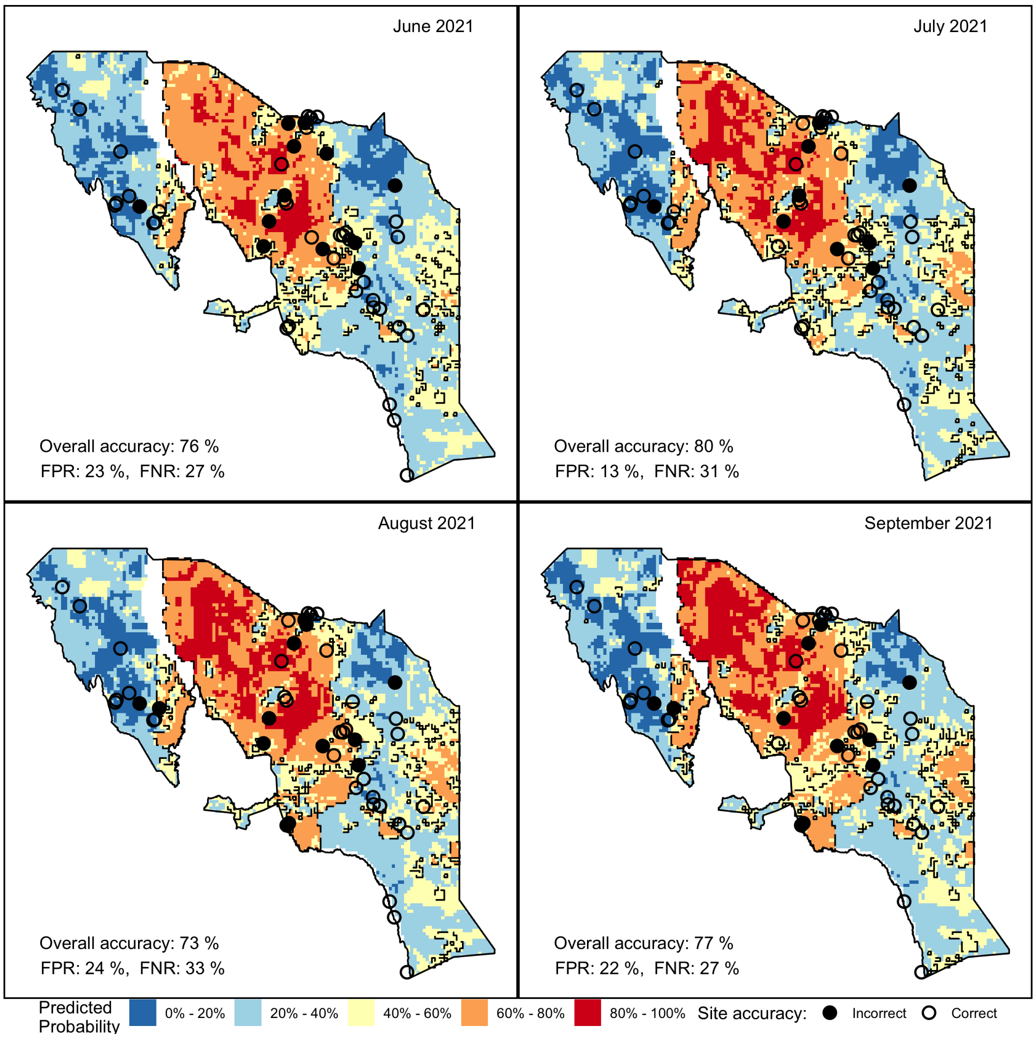


Figure 3: Predicted probability of groundwater demand for daily intermediate water needs (>75 L per person per day) forecasted for the 2021 dry season. Categorization of high demand is outlined and accuracy is given for each dry season month. Rates of false positives (FPR) and false negatives (FNR) are based on site-level observations.

are optimized to conserve material resources used to respond to high groundwater need, such as deploying pump repair teams, at the occasional expense of failing to capture some regions of high use.

Overall model accuracy (Table 1) of the contemporary categorical high runtime and basic and intermediate volume per capita was 67.6%, 69.4%, and 74.5%, respectively. A similar level of accuracy was observed for the one- to four-month forecasts. High predictive skill of volume per capita was demonstrated with external validation, indicating these models were not overfit. Accuracy was at least 70% and as high as 80% for forecasting per capita volume and represented significant improvement over the NIR. Predictive skill of pump runtime was lower when tested on held-out data.

Table 1: Accuracy statistics for contemporary and one- to four-month forecasts of categorical high pump runtime (6+ hr/d), volume per capita for basic needs (35+ L/pc/d), and volume per capita for intermediate needs (75+ L/pc/d). Sensor data from the 2021 dry season was reserved from model building to provide evidence of operational accuracy for the forecasts from external validation. AUC = Area Under the Curve; NIR = No Information Rate; p-value = significance value of difference between accuracy and NIR

Model	**AUC (95% CI)**	**Accuracy, % (p-value)**	**Validation Acc, %**
Pump Runtime, 6+ hr/d	–	NIR: 57.7	–
Contemporary	0.705 (0.663 - 0.748)	67.6 (<0.001)	–
1 mo forecast - Jun 2021	0.706 (0.663 - 0.748)	66.5 (<0.001)	56.0
2 mo forecast - Jul 2021	0.714 (0.672 - 0.756)	66.9 (<0.001)	58.7
3 mo forecast - Aug 2021	0.703 (0.661 - 0.746)	63.9 (0.002)	56.2
4 mo forecast - Sep 2021	0.708 (0.666 - 0.750)	65.8 (<0.001)	57.4
Volume Per Capita, 35+ L/pc/d	–	NIR: 51.5	–
Contemporary	0.756 (0.717 - 0.796)	69.4 (<0.001)	–
1 mo forecast - Jun 2021	0.756 (0.717 - 0.795)	68.7 (<0.001)	72.0
2 mo forecast - Jul 2021	0.748 (0.708 - 0.787)	66.7 (<0.001)	73.9
3 mo forecast - Aug 2021	0.749 (0.709 - 0.788)	68.3 (<0.001)	77.1
4 mo forecast - Sep 2021	0.746 (0.705 - 0.786)	69.0 (<0.001)	70.2
Volume Per Capita, 75+ L/pc/d	–	NIR: 58.4	–
Contemporary	0.787 (0.749 - 0.825)	74.5 (<0.001)	–
1 mo forecast - Jun 2021	0.785 (0.747 - 0.823)	74.3 (<0.001)	76.0
2 mo forecast - Jul 2021	0.783 (0.745 - 0.821)	74.5 (<0.001)	80.4
3 mo forecast - Aug 2021	0.778 (0.739 - 0.817)	73.8 (<0.001)	72.9
4 mo forecast - Sep 2021	0.781 (0.742 - 0.820)	75.0 (<0.001)	76.6

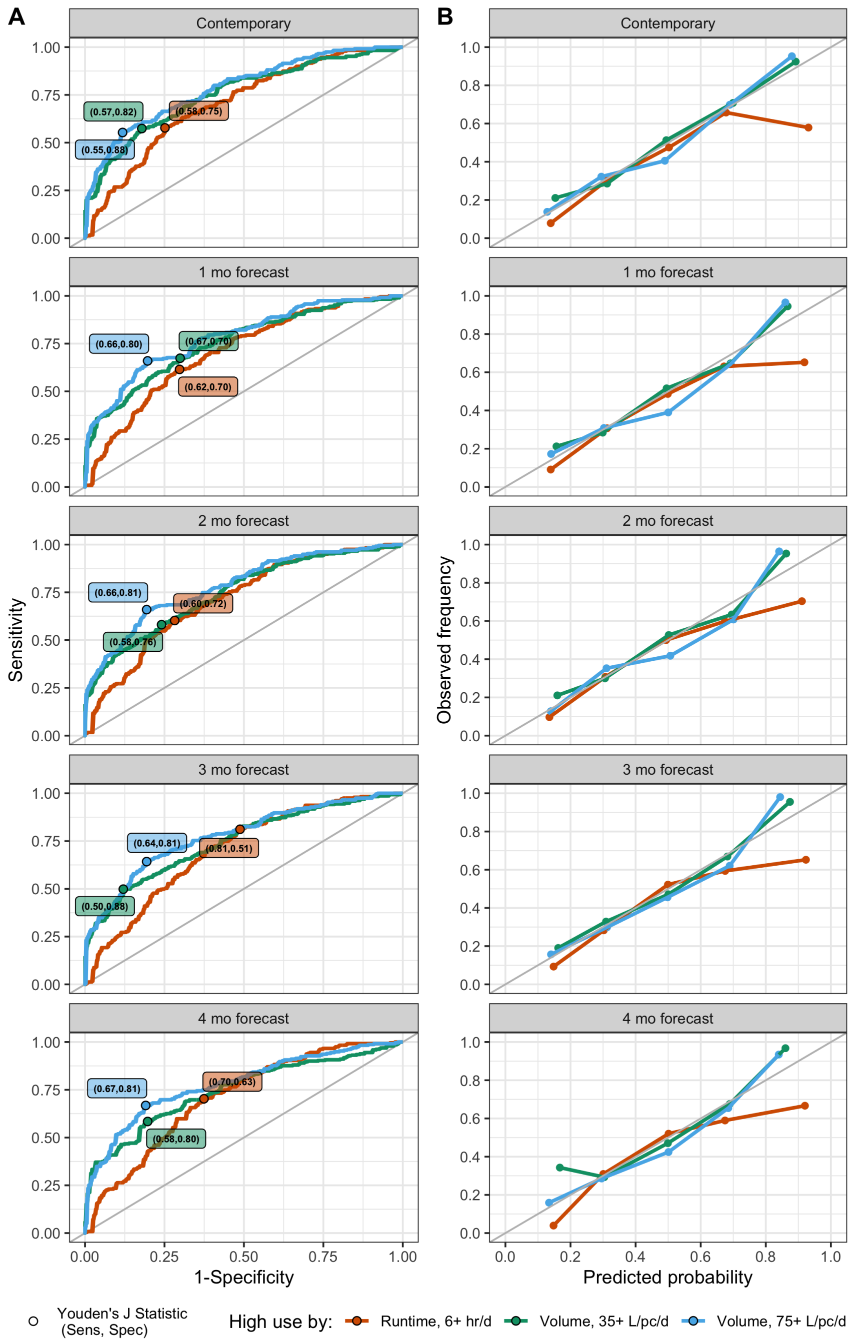


Figure 4: A) Receiver Operating Characteristic (ROC) curves and B) Reliability Diagrams. Performance given for pump runtime and volume per capita and by models used to estimate contemporary and forecasted groundwater use predictions.

Volume per capita predictions had the best reliability, exhibiting a consistent linear relationship between predicted probability and observed relative frequency (Figure 4B). Modeling of domestic volume per capita groundwater use assigned a probability of high use comparable to the occurrence of actual high use at each magnitude of probability except at the highest frequencies of use, when the probability was underestimated. At a lead time of four months, high use was also underestimated at the lowest frequencies of use. The model for groundwater use for intermediate water needs underestimated use at the highest frequencies and overestimated probability at median frequencies. Prediction of pump runtime tended to overestimate the probability of high use.

3.3. Contribution of model features

Variable importance plots (Figure 5) indicate the relative information gained from the independent explanatory features. Variable importance is measured by the ratio in error, using negative log likelihood loss, between the full ensemble model and a model on modified data where the respective features have been excluded from model building. Features are plotted against the risk ratio, where model error without the features is divided by the error obtained from the full model. See Appendix A for which features are described by each theme. For example, in the one-month lagged models aimed at forecasting groundwater use in June 2021, removing administrative features (county, longitude, latitude, elevation, and proximity to towns, markets, and roads) increased model error by about 8% for pump runtime, between 4% and 5% for basic per capita volume, and just over 1% for intermediate per capita volume, indicating that, collectively, these variables were less informative for forecasting groundwater volume per capita needs.

The plots demonstrate patterns in what forces the groundwater use predictions for each use definition and lead time. In general, features related to typical, neighboring borehole behavior were not instructive of site usage. At a lag of one month, features related to or suggestive of water availability at the surface – such as precipitation, water bodies, evapotranspiration, hydrology, and vegetation greenness – better informed volume per capita than pump runtime. With greater lagged observations, i.e. as the hypothetical dry season progressed, the signal between features and per capita volume forecasts became less clear; although for pump runtime, many of the features shared a similar level of importance throughout. Models with longer lead (three and four months) demonstrated, with a risk ratio less than one, some feature groups were damaging to model performance. This could be suggestive of some overfit or correlated variables within the group; however, since overall performance and accuracy of the forecasts were acceptable we did not investigate this further with a systematic sensitivity analysis.

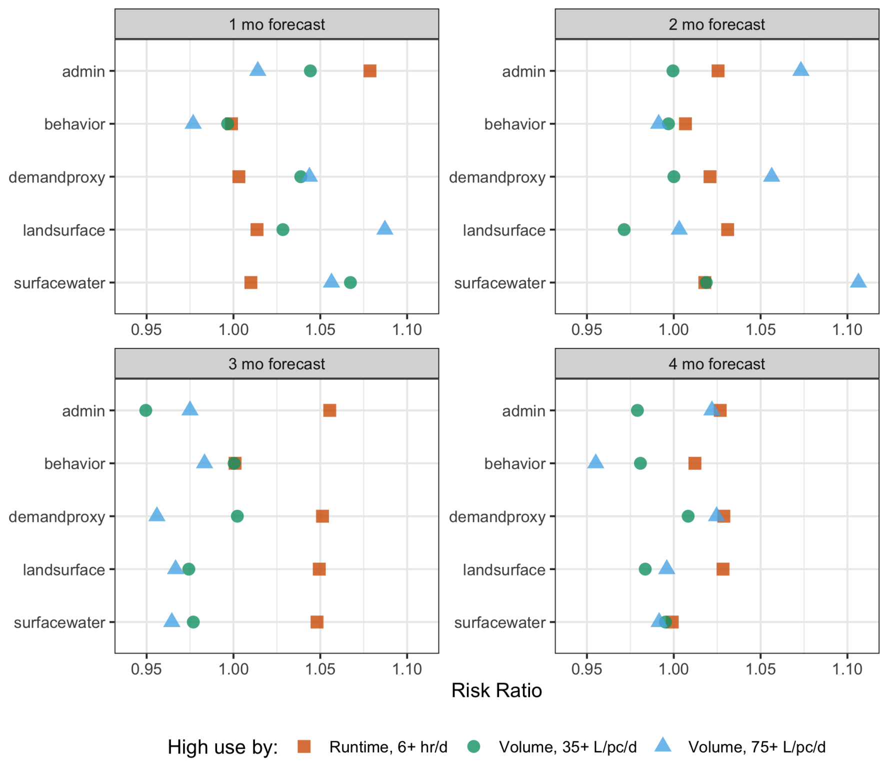


Figure 5: Variable importance plots for models predicting forecasted pump runtime and volume per capita. Covariates are grouped by theme; see Appendix A for which covariates are described by each theme.

4. Discussion

These maps represent the first operational spatially-explicit sub-seasonal to seasonal (S2S) estimates of groundwater use and demand in the literature. The integration of in-situ remote sensors with satellite data and hydrological land surface models through ensemble machine learning directly addresses an identified gap in population-based, near-real time, acute water monitoring (McNally et al., 2019) and supports other services for multi-dimensional drought early warning early action (Funk and Shukla, 2020).

The maps show areas of persistent high groundwater use, areas that develop a reliance on groundwater over the dry season, and locations where volume extraction suggests the requirement of groundwater to meet domestic needs for households and watering of livestock for pastoralists. Previous work (Thomas et al., 2019) has identified that pastoral and agropastoral households typically use a mix of groundwater and surface water to meet their needs, but have increased reliance on groundwater during dry seasons or in times of drought. Surface water or rainwater harvesting are preferred sources when available, with rural populations in Kenya reporting 34% less groundwater use during the wet season. The preference for surface water may be even more defined for livestock owners and the most water insecure populations (Thomson et al., 2019). Empirically we also observed an increase in borehole runtime and volume extraction during drought declared years and as the dry season advanced within years (see Section 4.1 below).

The ability to discriminate between high and low groundwater use did not attenuate over the dry season despite lead time up to four months between predictions and predictive features. Groundwater needs become more acute as the dry season progresses, resulting in a stronger measurable signal. Increasingly, rains prior to the dry season provide temporary relief, but are insufficient to provide adequate soil moisture for vegetation and recharge of surface water for the duration of the dry season. Thus, characterizing the hydrologic productivity of the rainy season, which occurs one to four months in advance, can explain patterns in subsequent groundwater use.

With external validation, forecasts of pump runtime proved to be non-informative during the 2021 dry season. Thus, there is likely additional unexplained variability among sites and activities supported by boreholes. Previous inventories of mechanical groundwater infrastructure in this region indicate that there are substantial differences in withdrawal efficiencies due to size, manufacturer and operator, age, power source, and on-site storage (Thomas et al., 2019). Therefore, despite the introduction of an additional source of error when translating mechanical pump runtime to volume with a yield coefficient, predictive skill of volume per capita was highest, particularly for intermediate ($>$ 75 L/pc/d) use. Instead, we believe that flowrate helped control for differences in pump capacity and implementation. Moreover, a higher volume threshold likely further differentiated between boreholes used for critical drought mitigation during the dry season specifically and other EDE-targeted boreholes that might have been used more indiscriminately.

Predictors related to surface water availability – satellite and gauged precipitation, relative depth of freshwater water points, and proximity to other natural water bodies – were indicative of volume per capita model performance at one and two month lead times. Many of the other satellite-derived and modeled hydrological features had an inconsistent effect on model error. The change in relative variable importance by forecast lead time highlights that characterizing water resources and drought dynamically and 'on-the-ground' with respect to human behavior is complex.

The utility of a statistically driven model is the assimilation of satellite data, land surface models, and field data into accurate estimates that lead to actionable insights in the absence of a clear or complete deterministic understanding. This is especially relevant to Sub-Saharan Africa, a context of sparse monitoring data, limited capacity, and developing scientific consensus of the agro-hydrological mechanisms driving drought (Funk and Shukla, 2020).

4.1. Estimated impact of groundwater demand on drought risk and planning

The methodology presented here could not have proceeded without in-situ data, but we have demonstrated that low density sensors can integrate satellite data and mechanistic land surface modeling for local groundwater monitoring. The fine spatial and temporal resolution of our gridded groundwater use and demand maps has not been achievable before. Approximately 70 sensors were used to inform a total area of 260,000 km^2, or one sensor per 3,700 km^2 (1,500 mi^2). A small network of sentinel sensors provides economical and efficient means to ground-truth and expand the utility of earth observation data.

We made a pixel-to-point comparison between the gridded groundwater estimates and observed site data. One-to-one comparisons between gridded satellite data and station observations have been practiced in the literature (Dinku et al., 2014) and while they have been shown to represent area averages less well than pixel-to-pixel comparisons (Dinku et al., 2018), the method is relevant to our context. In the ASALs of northern Kenya residents may walk up to 10 km to collect water for domestic needs and pastoralists walk even greater distances in search of livestock watering. Generating estimates at 0.05° x 0.05°, roughly equal to 5 km x 5 km, is appropriate for describing water collection behavior then. Additionally, the use of high vs. low categories smooths differences in groundwater use attributable to unexplained site variability at this scale. We report high cross-validated accuracy at this density and resolution. Thus, the modeled groundwater use/demand pixels did characterize site-level observations well.

The distribution of sensors in the five program counties was not uniform, but this was by design with respect to drought emergency. NDMA and county officials identified strategic EDE boreholes for monitoring and water resource managers anticipate areas without sensors to be lower priority during drought response because of low population densities, uninterrupted water resources other than groundwater, or other mitigating factors. Despite these differences, similar hydrological, climatological, and socioeconomic conditions are captured at sensored sites since the ASALs share many common characteristics (NDMA, 2015). In counties with better distribution of sensors, such as Turkana, predictions were accurate across a range of different conditions. Moreover, in areas of predicted high use without existing EDE priority, our predictions may prompt re-evaluation of EDE resources and mobilization of additional in-situ data collection.

Gridded groundwater demand contains essential information for decision-makers in settings experiencing or expecting drought. The percent area and percent of population affected by reliance on groundwater each year can be enumerated from our maps (Table 2). We delineated trends in groundwater use over time that can be related to other meteorological and hydrological phenomenon. The national government of Kenya declared drought in 2017 after several years of low precipitation. An above-average long rains season in 2018 helped sustain improvements in water and food security, but the next year saw a return to drier conditions and higher temperatures and, subsequently, 2019 was another year of drought. However, rainfall toward the end of the year and during the 2020 long rains promoted recovery. Unfortunately, the following two rainy seasons were critically below average, and another national drought disaster was declared in 2021 (OCHA, 2021a). The ASALs, where now over 2.5 million people are facing water and food shortages, have been especially impacted (OCHA, 2021b). We see relative changes reflected in high groundwater demand during drought years and accurately forecasted higher demand prior to the 2021 dry season (Table 2).

Table 2: Frequency of predicted high groundwater use over northern Kenya by percent of total area and population affected each year at any point during the dry season (June - September) of that year. Population counts were taken from WorldPop unconstrained UN Census adjusted estimates (Lloyd et al., 2019).

Year	**Pump Runtime 6+ hr/d**		**Volume Per Capita 35+ L/pc/d**		**Volume Per Capita 75+ L/pc/d**	
	Area (%)	*Popl (%)*	*Area (%)*	*Popl (%)*	*Area (%)*	*Popl (%)*
2017	54.2	50.3	64.3	40.4	33.6	10.1
2018	46.5	40.6	59.8	30.1	40.3	12.9
2019	51.5	49.0	57.9	27.4	37.0	10.0
2020	49.6	43.6	57.6	28.1	40.4	11.7
2021	71.4	68.4	63.1	34.7	45.0	14.7

A review of the 2021 dry season before it begins (Table 3), provided national and county officials with the total and percent of their constituents expected to be living under circumstances contributing to high groundwater use, such as limited surface water and lowered access to other services. Percent change in groundwater use by the end of the dry season demonstrates how and where need changes. All counties were expected to experience an increase (positive change) in groundwater reliance as the dry season progressed with one exception. In Isiolo, the percent of the population expected to undertake high groundwater use to meet basic water needs decreased from June to September. Counties that experienced the greatest change, such as Wajir and Garissa, may engender more attention and resource allocation for early warning early action services.

Table 3: Impact of 2021 dry season. Total and percent of population predicted to experience high groundwater use at the start of the 2021 dry season (June) by county and the the percent increase in population expected to be relying on groundwater by the end of the dry season (September). Population counts were taken from WorldPop unconstrained UN Census adjusted estimates (Lloyd et al., 2019).

County	**Pump Runtime 6+ hr/d**			**Volume Per Capita 35+ L/pc/d**			**Volume Per Capita 75+ L/pc/d**		
	# thous.	*%*	*% chg*	*# thous.*	*%*	*% chg*	*# thous.*	*%*	*% chg*
All	1830	35.9	63.4	1050	20.7	61.2	424	8.3	49.5
Garissa	392	24.1	45.5	123	7.6	140.5	34.1	2.1	81.9
Isiolo	72.7	39.5	14.3	105	57.0	-5.2	59.1	32.1	33.1
Marsabit	179	49.2	40.0	202	55.8	10.0	187	51.4	9.1
Turkana	859	72.5	25.8	191	16.1	24.9	65.1	5.5	56.3
Wajir	326	18.8	207.7	432	24.9	94.8	78.5	4.5	138.0

Several actions are anticipated from these maps in advance of drought. When mapped together, sensor data indicating pumps needing repair where high usage is predicted would be an alert to prioritize maintenance services to these monitored sites. Where there are no sensors, this kind of assessment would need to proceed from the institutional knowledge of local water officers, field scientists, and other experts. Thus, if high usage is predicted in areas without adequate borehole coverage and/or functionality — known through prior or external sources — then resources should be devoted to new installations, maintenance, and other water infrastructure.

The models had a higher specificity than sensitivity, meaning there were fewer false positive results and the risk of allocating resources to increase water availability unnecessarily is lower. Conversely, this interpretation accepts that that some instances of need will be missed. We chose a threshold to balance sensitivity and specificity for the highest overall accuracy, but the two parameters can be changed based on programmatic priorities.

Operational drought monitoring for Sub-Saharan Africa has been relatively nascent in the last decade, when satellite data and land surface modeling were leveraged to establish historical climatology and forecast parameters related to meteorological, agricultural, and hydrological drought, such as precipitation, soil moisture, vegetation, and streamflow. Yet, most monitors do not include a unique groundwater component.

FEWS NET provides several drought indicators (McNally et al., 2019), and additionally the water supply and demand product for crops, the Water Requirement Satisfaction Index (WRSI) (Senay et al., 2015), and surface water levels, the Waterpoint Viewer (Senay et al., 2013). Acute water stress monitoring related these to population needs (McNally et al., 2019). Although, given that these analyses are typically generated from renewable freshwater resources and that the volume of stored groundwater is estimated to be 100 times that of annual renewable sources and 20 times the volume of freshwater lakes in Africa, groundwater as a resource to meet domestic, pastoral, and agricultural needs is missing from water scarcity assessments (MacDonald et al., 2012). These datasets in large part form the evidence base of food security classifications and risk outlooks.

Thus, we propose the DRIP groundwater use and demand maps will become another reference dataset for drought indices and expand the knowledge base for decision making in Kenya and other future operational contexts. A challenge will be how to systematically integrate localized groundwater withdrawal estimates as a quantitative feature, but we have identified potential to apply them to drought monitoring through expert interpretation and guidance with several key stakeholders.

4.2. Case study applications

Stakeholder consultations were a critical component of this research. We conducted regular consultations with scientists at the eastern and southern Africa SERVIR hub – the Regional Centre for Mapping of Resources for Mapping (RCMRD) – and FEWS NET in order to align the groundwater use and demand products with user needs. We also organized two user engagement workshops. The first workshop was held in February 2021 and was attended by drought management officers from the NDMA. The second meeting was held in June 2021 and focused on the rollout of prototype products for Marsabit County in Kenya. Participants in the second meeting were drawn from NDMA, the Kenya Meteorological Department, drought humanitarian agencies including the Kenya Red Cross Society (KRCS), Mercy Corps, and Food for the Hungry, and the county departments of information technology and water services.

We received three important recommendations from these meetings: (i) the need for participatory planning and co-development of community early warning systems that leverage groundwater use and demand information from DRIP, (ii) continuous improvement of the groundwater use and demand products through expert feedback and user-based validations, and (iii) linking the products to existing drought information dissemination mechanisms through the NDMA national and county drought bulletins. These recommendations are part of future co-development processes. The development of the groundwater products is also synergistic with other rangeland vegetation monitoring services by RCMRD and national drought early warning early actions program led by the KRCS.

4.2.1. The Famine Early Warning System Network (FEWS NET)

FEWS NET scientists and domain experts (co-authors McNally and Slinski), reviewed the DRIP groundwater platform for its potential to inform the FEWS NET analysis of acute food insecurity risk. The risk of acute food insecurity is a function of a particular hazard or shock, the vulnerability of a specific population, and that population's ability to cope, or recover, from the shock. This framework informs the FEWS NET scenario development process that allows for the projection of food insecurity eight months in advance for humanitarian assistance planning.

The ASALs of Kenya are subject to shocks that include drought, animal pests and diseases, limited access to dry season grazing, and cattle raiding. Some communities additionally experience inadequate access to water for domestic use and watering livestock, especially during the dry season. This results in the occurrence of waterborne disease and poor animal health which are additional shocks to the system. Much of the population in northern Kenya relies on livestock for food and cash income (FEWS NET, 2011). Thus, given the nature of the hazards and vulnerability of the livestock sector, the ability of this region to cope tends to be low. In late 2021, this region was experiencing Phase III - Crisis food insecurity which is characterized by households that either have above usual acute malnutrition or are marginally able to meet minimum food needs by depleting essential assets or engaging in negative coping strategies.

The predictions of groundwater use and demand can be helpful to the FEWS NET scenario development process. As previously mentioned, groundwater demand is indicative of inadequate access to water from other sources, like surface water and shallow water infrastructure due either to non-functionality or drought. In this way, groundwater demand forecasts could alert food security analysis of a shift in behavior toward using available groundwater. In many locations, however, the groundwater maps and overlaid in-situ sensor data show that households have inadequate access to groundwater during crucial periods (e.g., dry season) due to non-functionality of infrastructure at nearby boreholes.

This unmet need, which contributes to the risk of food insecurity, could be highlighted in FEWS NET reporting and in turn addressed by decision-makers (e.g. programmers of humanitarian assistance or national water ministries) in a number of different ways: 1) encouraging water point rehabilitation, including repair and maintenance of pumps, in locations where DRIP indicates that wells are not properly functioning; 2) alternative solutions like water trucking could be employed in locations where wells do not exist or are unable to be repaired; 3) institutionalization of reliable access to groundwater through installation of new boreholes and increased support of pump maintenance in longer term planning documents. Preventive action to ensure sustainable groundwater access may be particularly useful in locations that suffer from unreliable water trucking due to inaccessibility or political factors.

Prior to the 2021 dry season and before the current drought in Kenya became apparent, our forecasts would have indicated an increased demand in groundwater relative to 2020 (Table 2). An increased reliance on groundwater in times of drought may, in some cases, involve traveling long distances to Kenya's EDE strategic boreholes. These predictions were consistent with the observed situation in the September 2021 Key Message Update (FEWS NET, 2021): "an atypically high number of livestock are migrating to dry season grazing areas driven by the decline in rangeland and water resources. Between July and August, livestock trekking distances to watering points increased by 60-90 percent, likely driving the 13-55 percent decline in milk production compared to the three-year average. [...] Overall, the decline in livestock productivity and body conditions is constraining household access to food and income and maintaining area-level Crisis (IPC Phase 3) outcomes across pastoral areas."

Taken together, knowledge of groundwater demand and FEWS NET risk profiles could have initiated concrete early actions such as borehole rehabilitation and pump repair or alternative reliable access to water through, for example, water trucking to improve water and food security during this year's drought.

4.2.2. The Kenya National Drought Management Authority (NDMA)

NDMA exercises coordination across drought risk management and establishes mechanisms, either on its own or with stakeholders, that will end drought emergencies in Kenya (NDMA, 2015). NDMA has headquarter offices in Nairobi, Kenya and has established sub-offices in 23 arid and semi-arid counties considered vulnerable to drought. County Steering Groups (CSG) manage the coordination of drought and early warning information at the county level in Kenya (USAID, 2018). The CSG is co-chaired by the county governor and county commissioner while the NDMA is the secretariat. Based on NDMA's overall mandate and role in guiding the agenda and discussion at CSG meetings, they have been targeted as a critical user of data on groundwater.

From predicted groundwater demand, NDMA and other stakeholders can identify mitigation activities against the effects of drought, such as identifying areas with predicted high water demand whose borehole pumps would require preventive maintenance. The Ministry of Water at the county would conduct timely budgeting, procurement of spare parts, and plan visits by borehole technicians to specific sites. Further, by analyzing the predicted water demand in relation to production capacities of boreholes in an area, NDMA and stakeholders could advise communities on migration that is usually triggered by the search for water. Such timely advice has the potential to avert violence between communities over competition of resources that has been common during the drought period. Thus, knowledge of groundwater patterns and forecasts helps public agencies address multi-dimensional impacts of drought, including health, livelihoods, and conflict.

4.3. Limitations and Future Work

In northern Kenya, groundwater supply is not a limiting factor. Recharge and fossil quantities in aquifers are higher than abstraction rates and are capable of providing groundwater (Mumma et al., 2011). This assumption would need to be revisited when applying this framework in new settings. Instead, being able to extract groundwater from functional infrastructure is a limitation that may contribute to unmet demand for water and under or overestimate use in our maps. We know from sensor reports and the motivation behind the DRIP theory of change (Thomas et al., 2020) that mechanical boreholes in this region are, in fact, often non-functional. In this study, our data filters attempted to remove the most persistent periods of non-functionality, and we have begun to identify and correct for functional status (Thomas et al., 2021). Our maps represent areas where dependence on groundwater at strategically located boreholes is high under the assumption that EDE boreholes do exist and are functioning in that area; if neither of these assumptions prove to be valid, that in and of itself may be the justification to initiate a response.

The inclusion of forecasted climatology as model features, to supplement or supplant the current lagged observations, should be explored. Predicting the dry season from conditions at a one to four month lead was demonstrated here because hydrological and agricultural drought, and subsequent reliance on groundwater, should be strongly correlated to conditions during the long rains season, March through May

(Shukla et al., 2021). However, incorporation of other forecasted drought indicators, which are often built from decades of historical climatology, could improve our predictive ability during years outside the range of values observed since in-situ borehole monitoring.

Further development of the gridded groundwater maps and DRIP service will focus on application and amplifying the wider context of groundwater health and potential for agricultural and rangeland productivity. Installation of groundwater level sensors and modeling of trends in groundwater level will address aquifer health, local overdraft, and sustainability of groundwater use. As groundwater is promoted and leveraged as a drought mitigation strategy, appropriate development and management supported by monitoring will be critical.

5. Conclusion

Groundwater represents an opportunity to increase reliable water supplies in Africa and provide a buffer against drought. Improving accessibility to groundwater through better maintenance of water systems and responsible development of new water schemes are mitigation strategies that reduce drought risk and improve resilience. Sustainable, effective, and local management of groundwater resources and infrastructure is a necessary precondition to achieving water security, as is accurate monitoring of changing water needs in a changing climate. The Drought Resilience Impact Platform fine resolution gridded dry season groundwater use and demand maps are a novel data source that can be used directly to allocate resources as well as provide a groundwater component to other water and food insecurity indices as part of multi-sectoral, multi-dimensional drought early warning and early action.

6. Funding

The work described in this publication was made possible through support provided by the National Aeronautics and Space Administration, under the terms of Grant No. 80NSSC20K0150; the National Science Foundation under the terms of Award No. 1738321; and the U.S. Agency for International Development, under the terms Cooperative Agreement No. AID-615-A-15-00008. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Agency for International Development.

Appendix A. Characteristics of Data Features

Features, or covariates, provided to groundwater prediction models. Theme relates to grouped features held out to test variable importance; see Figure 5. Res is the data product's native resolution. [1]Population density was used as an independent variable in pump runtime models, but for per capita volume predictions it was only used to scale the outcome and not as a model feature.

Coded Name	**Theme**	**Description**	**Units**	**Res**
amenities	admin	presence of major or primary roads, markets, or towns within 10 km	numeric (0 - 4)	–
county	admin	county	discrete (Turkana, Marsabit, Isiolo, Wajir, Garissa)	–
dem	admin	elevation from Digital Elevation Model	m	30 m
x	admin	longitude, east/west location	decimal degrees in WGS84	–
y	admin	latitude, north/south location	decimal degrees in WGS85	–
corrneighbor_*	behavior	average daily pump runtime or volume of interpolated 5 most linearly correlated sensored boreholes	hours or L/per capita	5 km
geoneighbor_*	behavior	average pump runtime or volume of 5 closest by Euclidean distance sensored boreholes	hours or L/per capita	–
geoneighbor_dist	behavior	distance between site or grid centroid and 5 closest sensored boreholes	km	–
proxy_*	behavior	modeled average daily pump runtime or volume based on static site characteristics	hours or L/per capita	–
boreholes_krapid	demandproxy	number of boreholes (inventory identified by Kenya RAPID) within 10 km	numeric	–
dams_or_pans	demandproxy	number of dams and sand pans (identified by Kenya RAPID) within 10 km	numeric	–
livestock_h2o	demandproxy	estimated amount of water required for livestock in 2015	discrete (>0.26, >0.50, >1.0, >1.5, >3.5 $m^3/km^2/day$)	subcounty
livestock_route	demandproxy	presence of major livestock migration route within 10 km	binary (0, 1)	–
month_factor	demandproxy	categorical month	discrete (1 - 12)	–

Coded Name	Theme	Description	Units	Res
people_h2o	demandproxy	estimated domestic water demand in 2015	discrete (>1030, >2000, >4000, >6000, >8000 m^3/day)	subcounty
popldens[1]	demandproxy	estimated population density, number of people per square kilometer	numeric	1 km
baseflow	landsurface	total baseflow, streamflow that is sustained between precipitation events	mm	3 km
et	landsurface	total actual evapotranspiration	mm	3 km
gvf	landsurface	greenness vegetation fraction	proportion (0 - 1)	4 km
maxtemp	landsurface	average daily maximum temperature	degrees Celsius	50 km
ndvi	landsurface	average 10-day maximum Normalized Difference Vegetation Index value	proportion (0 - 1)	250 m
runoff	landsurface	total amount of flow of water across the ground surface after it no longer infiltrates the soil	mm	3 km
soilm1	landsurface	average daily soil moisture in top layer at depths of 0 - 10 cm	m^3/m^3	3 km
soilm2	landsurface	average daily soil moisture at depths of 10 - 40 cm	m^3/m^3	3 km
soilm3	landsurface	average daily soil moisture at depths of 40 - 100 cm	m^3/m^3	3 km
soilm4	landsurface	average daily soil moisture at depths of 100 - 200 cm	m^3/m^3	3 km
chirps	surfacewater	total precipitation from Climate Hazards Group InfraRed Precipitation with Station (CHIRPS) v2.0	mm	5 km
naturalh2o	surfacewater	presence of floodplains, lakes, springs, or river basins with 10 km	numeric (0 - 4)	–
tahmo_chirps	surfacewater	total localized precipitation, CHIRPS v2.0 scaled and bias-corrected with Trans-African Hydro-Meteorological Observatory (TAHMO) weather stations	mm	5 km
waterpoint_depth	surfacewater	average daily interpolated relative surface water depth	percentage (0 - 100)	5 km

References

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

Cabot Venton, C., 2018. Economics of Resilience to Drought: Kenya Analysis. Technical Report. USAID Center for Resilience.

Case, J.L., LaFontaine, F.J., Bell, J.R., Jedlovec, G.J., Kumar, S.V., Peters-Lidard, C.D., 2014. A Real-Time MODIS Vegetation Product for Land Surface and Numerical Weather Prediction Models. IEEE Transactions on Geoscience and Remote Sensing 52, 1772–1786. doi:10.1109/TGRS.2013.2255059.

Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Association for Computing Machinery, New York, NY, USA. pp. 785–794. doi:10.1145/2939672.2939785.

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H., Ceccato, P., 2018. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Quarterly Journal of the Royal Meteorological Society 144, 292–312. doi:10.1002/qj.3244.

Dinku, T., Hailemariam, K., Maidment, R., Tarnavsky, E., Connor, S., 2014. Combined use of satellite estimates and rain gauge observations to generate high-quality historical rainfall time series over Ethiopia. International Journal of Climatology 34, 2489–2504. doi:10.1002/joc.3855.

FAO, 2021. AQUASTAT. URL: https://www.fao.org/aquastat/en/.

FEWS NET, 2011. Livelihoods zoning "plus" activity in Kenya. Technical Report. URL: https://fews.net/sites/default/files/documents/reports/KE_livelihood_profiles.pdf.

FEWS NET, 2021. Consecutive below-average agricultural seasons increase household food insecurity. URL: https://fews.net/east-africa/kenya/key-message-update/september-2021.

Friedman, J.H., Hastie, T., Tibshirani, R., 2010. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software 33, 1–22. doi:10.18637/jss.v033.i01.

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. doi:10.1038/sdata.2015.66.

Funk, C., Shukla, S., 2020. Drought Early Warning and Forecasting: Theory and Practice. Elsevier Inc.

Gelman, A., Jakulin, A., Pittau, M.G., Su, Y.S., 2008. A weakly informative default prior distribution for logistic and other regression models. The Annals of Applied Statistics 2. doi:10.1214/08-AOAS191.

Giesen, N.v.d., Hut, R., Selker, J., 2014. The Trans-African Hydro-Meteorological Observatory (TAHMO). WIREs Water 1, 341–348. doi:10.1002/wat2.1034.

Hofste, R.W., Kuzma, S., Walker, S., Sutanudjaja, E.H., Bierkens, M.F.P., Kuijper, M.J.M., Sanchez, M.F., Beek, R.V., Wada, Y., Rodríguez, S.G., Reig, P., 2019. Aqueduct 3.0: Updated Decision-Relevant Global Water Risk Indicators URL: https://www.wri.org/research/aqueduct-30-updated-decision-relevant-global-water-risk-indicators.

KNBS, 2019. Volume I: Population by County and Sub-County. Technical Report. Kenya National Bureau of Statistics.

van der Laan, M.J., Polley, E.C., Hubbard, A.E., 2007. Super learner. Statistical Applications in Genetics and Molecular Biology 6. doi:10.2202/1544-6115.1309.

Landerer, F.W., Swenson, S.C., 2012. Accuracy of scaled GRACE terrestrial water storage estimates. Water Resources Research 48. doi:10.1029/2011WR011453.

Liebmann, B., Bladé, I., Funk, C., Allured, D., Quan, X.W., Hoerling, M., Hoell, A., Peterson, P., Thiaw, W.M., 2017. Climatology and Interannual Variability of Boreal Spring Wet Season Precipitation in the Eastern Horn of Africa and Implications for Its Recent Decline. Journal of Climate 30, 3867–3886. doi:10.1175/JCLI-D-16-0452.1.

Lloyd, C.T., Chamberlain, H., Kerr, D., Yetman, G., Pistolesi, L., Stevens, F.R., Gaughan, A.E., Nieves, J.J., Hornby, G., MacManus, K., Sinha, P., Bondarenko, M., Sorichetta, A., Tatem, A.J., 2019. Global spatio-temporally harmonised datasets for producing high-resolution gridded population distribution datasets. Big Earth Data 3, 108–139. doi:10.1080/20964471.2019.1625151.

MacDonald, A.M., Bonsor, H.C., Dochartaigh, B., Taylor, R.G., 2012. Quantitative maps of groundwater resources in Africa 7. doi:10.1088/1748-9326/7/2/024009.

McNally, A., Verdin, K., Harrison, L., Getirana, A., Jacob, J., Shukla, S., Arsenault, K., Peters-Lidard, C., Verdin, J.P., 2019. Acute Water-Scarcity Monitoring for Africa. Water 11. doi:10.3390/w11101968.

Mouselimis, L., 2021. KernelKnn: Kernel k Nearest Neighbors. URL: https://CRAN.R-project.org/package=KernelKnn.

Mumma, A., Lane, M., Kairu, E., Tuinhof, A., Hirji, R., 2011. Kenya Groundwater Governance Case Study. Technical Report. World Bank. Washington, DC. URL: https://openknowledge.worldbank.org/handle/10986/17227.

MUS Group, 2013. Briefing Note: Matching Water Services with Water Needs. Technical Report. Multiple Use Water Services Group. URL: https://www.musgroup.net/node/6.

Nagel, C., Beach, J., Iribagiza, C., Thomas, E.A., 2015. Evaluating Cellular Instrumentation on Rural Handpumps to Improve Service Delivery—A Longitudinal Study in Rural Rwanda. Environmental Science & Technology doi:10.1021/acs.est.5b04077.

NDMA, 2015. Common Programme Framework for Ending Drought Emergencies. Technical Report. National Drought Management Authority. Nairobi, Kenya. URL: https://www.ndma.go.ke/index.php/resource-center/send/43-ending-drought-emergencies/4251-common-programme-framework.

Nicholson, S.E., Funk, C., Fink, A.H., 2018. Rainfall over the African continent from the 19th through the 21st century. Global and Planetary Change 165, 114–127. doi:10.1016/j.gloplacha.2017.12.014.

Niu, G.Y., Yang, Z.L., Mitchell, K.E., Chen, F., Ek, M.B., Barlage, M., Kumar, A., Manning, K., Niyogi, D., Rosero, E., Tewari, M., Xia, Y., 2011. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. Journal of Geophysical Research: Atmospheres 116. doi:10.1029/2010JD015139.

OCHA, 2021a. Kenya: Drought - 2014-2021. URL: https://reliefweb.int/disaster/dr-2014-000131-ken.

OCHA, 2021b. Kenyan ASAL counties continue to face the brunt of the Drought. URL: https://reliefweb.int/report/kenya/kenyan-asal-counties-continue-face-brunt-drought.

Physical Sciences Laboratory, . CPC Global Daily Temperature. URL: https://psl.noaa.gov/data/gridded/data.cpc.globaltemp.html.

Polley, E., Laan, M.v.d., 2010. Super Learner In Prediction. U.C. Berkeley Division of Biostatistics Working Paper Series URL: https://biostats.bepress.com/ucbbiostat/paper266.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

RCMRD Geoportal, 2015. Kenya SRTM 30 Meters. URL: http://geoportal.rcmrd.org/layers/servir%3Akenya_srtm30meters.

Senay, G., Velpuri, N., Bohms, S., Budde, M., Young, C., Rowland, J., Verdin, J., 2015. Drought Monitoring and Assessment, in: Hydro-Meteorological Hazards, Risks and Disasters. Elsevier, pp. 233–262. URL: https://linkinghub.elsevier.com/retrieve/pii/B9780123948465000096.

Senay, G.B., Velpuri, N., Alemu, H., Md Pervez, S., Asante, K.O., Kariuki, G., Taa, A., Angerer, J., 2013. Establishing an operational waterhole monitoring system using satellite data and hydrologic modelling: Application in the pastoral regions of East Africa. Pastoralism: Research, Policy and Practice 3. doi:10.1186/2041-7136-3-20.

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. doi:10.1371/journal.pone.0242883.

Stevens, F.R., Gaughan, A.E., Linard, C., Tatem, A.J., 2015. Disaggregating Census Data for Population Mapping Using Random Forests with Remotely-Sensed and Ancillary Data. PLOS ONE 10. doi:10.1371/journal.pone.0107042.

The World Bank, 2021. Population growth (annual %) - Kenya. URL: https://data.worldbank.org/indicator/SP.POP.GROW?locations=KE.

Thomas, E., Wilson, D., Kathuni, S., Libey, A., Chintalapati, P., Coyle, J., 2021. A contribution to drought resilience in East Africa through groundwater pump monitoring informed by in-situ instrumentation, remote sensing and ensemble machine learning. Science of The Total Environment 780. doi:10.1016/j.scitotenv.2021.146486.

Thomas, E.A., Kathuni, S., Wilson, D., Muragijimana, C., Sharpe, T., Kaberia, D., Macharia, D., Kebede, A., Birhane, P., 2020. The Drought Resilience Impact Platform (DRIP): Improving Water Security Through Actionable Water Management Insights. Frontiers in Climate 2. doi:10.3389/fclim.2020.00006.

Thomas, E.A., Needoba, J., Kaberia, D., Butterworth, J., Adams, E.C., Oduor, P., Macharia, D., Mitheu, F., Mugo, R., Nagel, C., 2019. Quantifying increased groundwater demand from prolonged drought in the East African Rift Valley. Science of The Total Environment 666, 1265–1272. doi:10.1016/j.scitotenv.2019.02.206.

Thomson, P., Bradley, D., Katilu, A., Katuva, J., Lanzoni, M., Koehler, J., Hope, R., 2019. Rainfall and groundwater use in rural Kenya. Science of The Total Environment 649, 722–730. doi:10.1016/j.scitotenv.2018.08.330.

Tolk, L., de Vries, A., de Wildt, S., 2016. 3R/MUS Analysis: Data Inventory, Kenya RAPID. Technical Report. Acacia Water. URL: https://kenyarapid.acaciadata.com/.

Turman-Bryant, N., Nagel, C., Stover, L., Muragijimana, C., Thomas, E.A., 2019. Improved Drought Resilience Through Continuous Water Service Monitoring and Specialized Institutions—A Longitudinal Analysis of Water Service Delivery Across Motorized Boreholes in Northern Kenya. Sustainability 11, 3046. doi:10.3390/su11113046.

UNICEF, 2017. Kenya Humanitarian Situation Report. Technical Report. United Nations Children's Emergency Fund. URL: https://reliefweb.int/report/kenya/unicef-kenya-humanitarian-situation-report-20-february-2017.

USGS FEWS NET Data Portal, 2017. eMODIS AQUA Normalized Difference Vegetation Index (NDVI). URL: https://earlywarning.usgs.gov/fews/product/448.

Vargas, M., Jiang, Z., Ding, H., NOAA JPSS Program Office, 2015. NOAA JPSS Visible Infrared Imaging Radiometer Suite (VIIRS) Green Vegetation Fraction (GVF) from NDE.

Wilhite, D.A., Glantz, M.H., 1985. Understanding the Drought Phenomenon: The Role of Definitions. Water International 10, 111–120.

Wilson, D.L., Coyle, J.R., Thomas, E.A., 2017. Ensemble machine learning and forecasting can achieve 99% uptime for rural handpumps. PLOS ONE 12. doi:10.1371/journal.pone.0188808.