Development of a Drought Indicator Using Machine Learning Algorithm

A Case Study of Zambia

05th December 2022

Colombo, Sri Lanka
Rim Sleimi, Surajit Ghosh, Giriraj Amarnath

CGIAR
INITIATIVE ON
Climate Resilience

Authors

Rim Sleimi[1], Surajit Ghosh[1], Giriraj Amarnath[1]

[1] International Water Management Institute (IWMI), Battaramulla, Sri Lanka

Suggested Citation

Sleimi R, Ghosh S, Amarnath G. 2022. *Development of a Drought Indicator Using Machine Learning Algorithm*. CGIAR Climate Resilience Initiative.

Acknowledgments

This work was carried out with support from the CGIAR Initiative on Climate Resilience, ClimBeR. We would like to thank all funders who supported this research through their contributions to the CGIAR Trust Fund.

Table of Contents

1. Introduction

Under the constant pressure of Spatio-temporal alterations of precipitation regimes coupled with the increasing demand for water, the frequency and intensity of drought events are constantly changing worldwide. This least understood weather-related phenomenon has severe environmental and socioeconomic impacts [1–11]. It is recognized as one of the leading environmental challenges to farmers worldwide as the availability of water is very crucial in farming [2, 5, 6, 8, 11–22]. In Africa for example, severe droughts followed by famine occurred in the 1910s, 1940s, 1960s, 1970s, and 1980s [23]. In the past 10 years alone, sub-Saharan Africa has endured three severe droughts (2015-2016, 2018-2019, and 2020-2021). Chikoore et al.(2021) [24], have shown that the 1983 and 1992 droughts have been exceeded by 2015, 2016, and 2019 in Southern Africa. Despite the expanding global adverse effects of droughts and their importance as a critical driver of ecological and evolutionary dynamics, current understanding of its physical causes and their potential consequences are still limited [25]. The bulk of the current research has concentrated on developing systems to monitor and forecast droughts, rather than focusing on the intricacies and the driving forces of drought. Regardless of the underlying mechanisms, a drought event occurs when the specified need for water exceeds water supplies. This complex phenomenon develops slowly and can be broadly classified into four types; meteorological, hydrological, agricultural, and socioeconomic droughts [1, 3, 6, 10, 11, 13, 14, 16, 17, 22, 26, 27]. As agricultural drought is directly linked to crop failure and famine, it has attracted widespread attention worldwide. Understanding the progression and the potential impacts of droughts in southern Africa is of utmost importance because agriculture is the basic economic activity for the majority of the population in these countries including Zambia. Furthermore, These countries' development is hindered by the reoccurrence of drought events as the allocated resources are redirected towards limiting the catastrophic repercussions and losses. Therefore, an early detection will play an integral role in controlling its occurrence, as well as reducing the complexity and consequences of drought [2, 6]. To mitigate the impacts of drought an effective and timely monitoring and forecasting system is required to aid in developing an early warning system that can provide valuable insights about the drought development and changes in drought-prone areas to support drought mitigation decision-making [1, 2, 6–8, 10–12, 14, 16, 17, 20, 22, 27–35].

So far, most of the efforts dedicated to drought investigation and research are focused on meteorological drought assessment and thus solely consider precipitation in their analysis [16, 28, 36–44]. This is because, drought occurrence is mostly determined by rainfall performance in a given location, with droughts occurring in areas of both low and high rainfall. The Standardized Precipitation Index (SPI) is one of the most commonly used precipitation-based indices and it is recommended by the World Meteorological Organization (WMO) to assess the impacts of drought worldwide regardless of the climate. However, drought conditions are associated with multiple factors, namely precipitation, temperature, evapotranspiration, soil moisture, and vegetation health which is especially true for agricultural drought. Thus, drought indicators based on a single factor are not sufficient to characterize complicated drought conditions and their wide impact [20]. A better way to approach this problem is by taking into consideration the interactions among the aforementioned influencing factors. Several integration methods have been reported in the consulted literature: 1) methods based on numerical models and on data assimilation; 2) gauge-only products, derived only from gauge data, 3) satellite-based products, based only or partially on satellite data. The latter products use station data or data from models for calibration or bias adjustment. The commonly used numerical models in drought-related research include: the water balance models (standardized precipitation evapotranspiration index (SPEI), Palmer Drought Severity Index (PDSI), Palmer Modified Drought Index (PMDI)), crop models [8], data mining models (Vegetation Drought Response Index (VegDRI), United States Drought Monitor (USDA) [45]).

Compared to the other data integration methods, these models are not widely used because they are limited by the inherent uncertainties of both the model's parameterization and the observations it uses. Gauge-only products are based on ground measurements of meteorological and agricultural variables that are directly related to drought. Such variables are daily measurements of rainfall, temperature, and soil moisture. Usually, these variables are aggregated to weekly or monthly values and used to compute drought indicators [46]. In situ measurements from gauges can be accurate. However, in data-sparse regions such as the case of African countries, drought monitoring using direct measurements is not possible given the high cost, maintenance, and upgrade requirements of the climate stations [6]. Especially since the national meteorological agencies are often underfunded. In fact, the number of weather stations is often under the minimum recommended by World Meteorological Organization (WMO) and has been decreasing. In addition, ground-based indices are local (point measurements) and limited in characterizing detailed spatial distributions of drought at regional scales [3, 6, 8, 17, 26]. The use of interpolation and extrapolation methods can lead to uncertainties due to the algorithms used and the complex topographic conditions. Remote sensing-based approaches offer a reliable data source to indirectly derive such information. With the recent advent of open-source satellite imagery, timeliness, medium to high temporal, spectral, and spatial resolution data can be easily acquired with global coverage. Remote measurements, using satellite imagery, are being considered a more effective and efficient method for such a task. Optical, thermal and microwave remote sensing are already a validated tool for drought monitoring in Europe [26], Africa [2, 10, 28], Australia [17], and Asia [1, 3, 5–8, 14, 29, 30]. In the consulted literature, numerous methods have been implemented to derive remote sensing-based drought indicators, namely linear combinations (Integrated Scaled Drought Index (ISDI) [47], Vegetation Temperature Precipitation Condition Index (VTPCI) [19], Scaled Drought Condition Index (SDCI) [48], he Microwave Integrated Drought Index (MIDI) [49], principal component analyses (Effective Meteorological Drought Index (EMDI) [6], agricultural drought indicator (ADI) [31], synthesized drought index (SDI) [13], the new drought index (PMDI) [50], The comprehensive drought index (CDI) [32], combined drought indicator for Marathwada (CDI_M) [34], joint distributions (multivariate standardized drought index (MSDI) [51]), and machine learning approaches (random forest [1, 22], boosted regression trees [22], and Cubist [22], deep feed forward neural network (DFNN) [5], Artificial neural network (ANN) [14]). Linear models are widely used for their simple theoretical basis and their straightforwardness. For example, $SDCI = (1/4)*scaled\ LST + (2/4)*scaled\ TRMM + (1/4)*scaled\ NDV\ I$ [48]. where *LST* is the land surface temperature, *TRMM* is the Tropical Rainfall Measuring Mission rainfall data and *NDV I* is the normalized difference vegetation index; Similarly, by assigning different weights to each unique remote sensing index ISDI was developed using the following formula $ISDI = \alpha*Scaled\ NDV\ I + \beta*Scaled\ LST + \gamma*Scaled\ PCP + \lambda*Scaled\ SM$[47].

Although their implementation is still in its infant stages, with the introduction of modeling based on machine learning and deep learning methods, there is more potential for early warning system development. Machine learning is a group of advanced statistical methods that can be used for prediction and classification tasks. These methods generalize from experience, by learning from a training dataset. Coupled with the recent development in computing power, ML is able to process large amounts of different data in a reasonable time. They are widely used in many fields of science and engineering and currently provide unprecedented opportunities for a variety of geosciences applications as well [52]. More recently, the emergence of deep learning has drawn growing interest in the computer vision community. Deep learning is a set of processing layers that learn representations of data with multiple levels of abstraction [53]. Unlike machine learning methods, they are able to extract intricate structures in large datasets instead of handcrafting features that are designed based mainly on domain-specific knowledge. Deep learning applied to remote sensing data is an attractive application that can provide more consistent results over time and space. This combination can

help overcome common challenges of traditional methods such as subjective judgment, time constraints, cost, and generalization. Early drought monitoring can also benefit from such applications that can provide reliable maps to help guide decision-making. However, demarcating areas with a potentially high risk of droughts using space-borne data and deep learning methods is still in the developing stages. The current implementation of machine learning focuses on the temporal forecast of drought indices and drought patterns. For example, Adede et al. (2019) [10], used 10 precipitation and vegetation condition indices lagged over 1, 2, and 3-month time steps to predict future values of vegetation condition index aggregated over a 3-month time period (VCI3M) using ANN. VCI3M was used here as a proxy for drought monitoring. Shen et al.(2019) [5] used Deep Forward Neural Network (DFNN) to predict the relative soil moisture in Henan province, China. In a similar study conducted by Liu et al. (2020) [14], ANN was used to predict the available water content (AWC) over the North China Plain. Prodhan et al.(2021) '[3] conducted a assessed the performance of distributed Random Forest (DRF), Gradient Boosting Machine (GBM), and a Deep Forward Neural Network (DFNN) in predicting the soil moisture deficit index (SMDI) in South Asia. Their results showed that the DFNN model outperformed the other two models. Achite et al. (2022) [16] also evaluated the hydrological drought forecasting potential of machine learning models (i.e., Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Support Vector Machine (SVM), and Decision Trees(DT)) in the Wadi Ouahrane basin in the northern part of Algeria. Their results showed that all the ML models accurately predicted hydrological drought, while the SVM model outperformed the other ML models. Feng et al.(2019) [17] compared the performance of three machine learning methods, i.e., bias-corrected random forest (BRF), multi-layer perceptron neural network (MLP), and support vector machine (SVM), in reproducing and predicting ground-based SPEI in dry and wet areas. Their results showed that BRF performed the best and had a better agreement with station-based drought maps and wheat yields than the station-based SPEI. Despite the good performances of the reported models in the aforementioned studies, it is worth mentioning that they suffer from several limitations: 1) Most of these studies rely on lengthy records of gauge datasets, 2) focus on one forecasting a single variable as a drought proxy 3) The machine learning models used take into consideration the temporal dimension of the dataset and discard the spatial dimension. 4) There is no consensus about the computational power needed to reproduce such results, 5) The conducted research is applicable on a regional scale and tailored to specific climate conditions. These limitations hinder the reproducibility of such methods in African countries such as the case of Zambia.

As indicators of climate change and unsustainable use of water continue to surface in Zambia, and rightfully raise sustainability concerns, the need for innovative methods of monitoring and managing droughts, within the country's capabilities, is becoming more prevalent. The construction of the drought monitoring system requires a paradigm shift. It should adapt to the practical needs of specific regions in terms of the diverse climate conditions over space, the different physical causes of droughts, the cost, and the potentially impacted sectors.

We define an efficient drought early warning system as a system that:

- Can be used for both monitoring and forecasting;
- Integrates the different factors that can explain drought;
- Simple and straightforward. It doesn't consider the underlying stochastic mechanisms of drought;
- Able to apply readily available open-source satellite data, to detect or predict the occurrence, duration, and spatial patterns of droughts;

- Takes advantage of the open-source cloud-computing and AI technologies; • Developed for extensive application in different climate-hydrology regions;
- Easily deployed with near real-time response.

To answer these requirements, we propose an integrated drought monitoring and forecasting system on a shortterm scale through the integration of precipitation, temperature, soil moisture, evapotranspiration, and vegetation as explanatory variables to monitor drought in Zambia using Principal Component Analysis (PCA) and convolutional long short term memory (ConvLSTM). The aim of this study is to 1) Derive a composite drought index (CMDI) that encapsulates the effects of different drought indicators (precipitation, vegetation, evapotranspiration, soil, and temperature); using PCA. 2) Implement ConvLSTM for Spatio-temporal prediction of CMDI. 3) Demonstrate that online platforms such as Google Earth Engine [54, 55]- a cloud-based geospatial analysis, as well as Google Colab and Kaggle which offer free yet limited GPU usage, are effective tools for carrying out the analysis of global geospatial big data at scale. Therefore, leveraging them can help bypass the need for expensive and limited datasets and hardware. 4) Compare the spatial-temporal distribution of CMDI with multi-time scale SPI (3-month to 12-month) and Gross Primary production (GPP). The methodology is shown in the flowchart, Figure. 2.

2. Materials and Methods

Study Area

The study area is Zambia. An East African country that lies between longitudes 22°E and 34°E and latitudes 8°S and 18°S with an area of 752,617 km^2. Figure 1 shows the location of Zambia and the selected meteorological station under study.

Because of its geographical location, Zambia experiences a predominantly subtropical climate, characterized by three distinct seasons: a hot and dry season (mid-August to mid-November), a wet rainy season (midNovember to April), and a cool dry season (May to mid-August) [56]. The mean temperature is around 30°C in summer and can get as low as 5°C in the winter [36]. Rainfall is largely influenced by the Inter Tropical Convergence Zone (ITCZ) which is a low-pressure system caused by the convergence of trade winds as well as the El Niño/Southern Oscillation (ENSO) phenomenon [57].

The highest rainfall is observed in the north, especially in the northwest and the northeast with an annual average of 1,400 mm; then, it tends to decrease towards the south with an annual average of 700 mm [33, 56, 57]. Furthermore, the majority of rainfalls are concentrated in the six months from November to April (wet rainy season). Consequently, for most rain-fed crops the sowing and growing seasons correspond to the rainy season. Being an agrarian economy, the country's cultural, economic, and social life is heavily dependent on the onset and end of the rainy season as well as the amount of rain it brings. Over the years, The lack of adequate precipitation resulted in hunger.

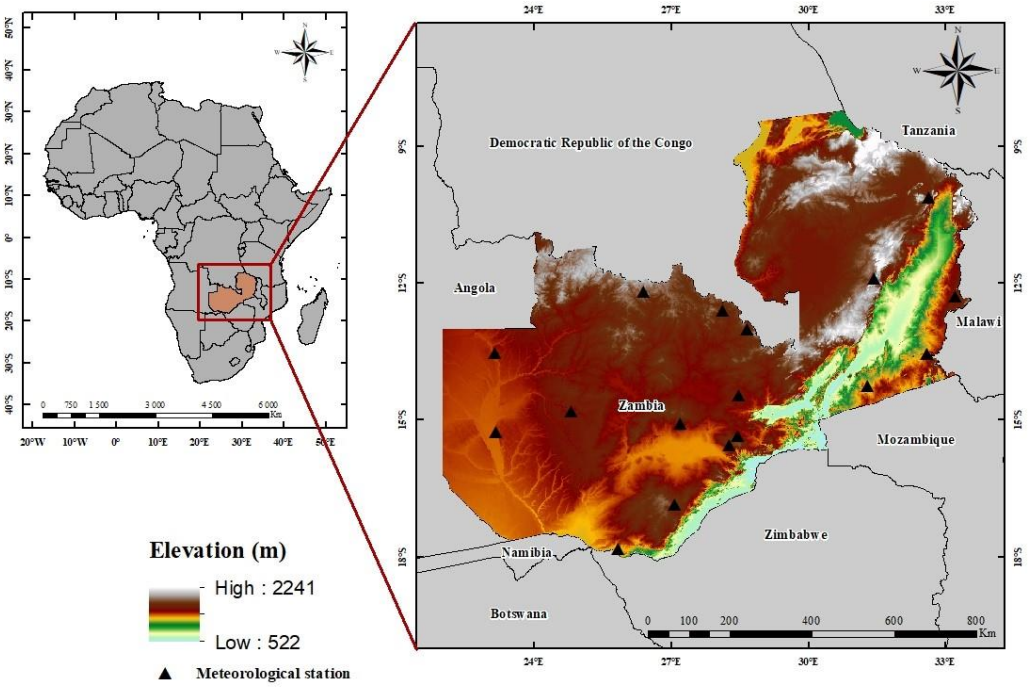


Figure 1: *Geographical location of the study area*

Dataset and preprocessing

Remote Sensing Data

In this study, datasets from remote sensing were used to calculate drought indices as well as to validate the combined drought index. The specifications of these datasets are presented in table 1.

The Terra and Aqua Moderate–Resolution Imaging Spectroradiometer (MODIS) evapotranspiration (ET), land surface temperature (LST), and Normalized Difference Vegetation Index (NDVI) products were used to calculate the ETCI, TCI, and VCI remotely sensed drought factors respectively. We Combine MODIS Terra and Aqua products to optimize the data quality. Precipitation data from Integrated Multi-satellitE Retrievals for GPM (IMERG) was used for the calculation of the Precipitation Condition Index (PCI), due to its consistency and sufficient length which make the data more reliable for climate variability analysis [58]. For the calculation of SMCI, we used soil moisture data at depths 1040cm. This data was derived from GLDAS, NOAH land surface model of NASA.

For validation, we used the MODIS Terra and Aqua Gross Primary Productivity (GPP) data, as well as the Standardized Precipitation Index (SPI) calculated based on the IMERG precipitation dataset.

Table 1 *The specification of remote sensing data used in this study*

Data Sources	Variables	Temporal Resolution	Spatial Resolution
MODIS	SR	8 days	500 m
	ET	8 days	500 m
	LST	8 days	1 km
	NDVI	16 days	250 m
	GPP	8 days	500 m
IMERG	Precipitation	Half hourly	0.1° * 0.1°
GLDAS	Soil Moisture	3-hourly	0.25° * 0.25°

Insitu data

Considerable studies have highlighted the effectiveness of rainfall data as an indicator for monitoring meteorological drought. Therefore, daily precipitation measured at the site was used as the reference index of meteorological drought to evaluate the reliability of the proposed CMDI. The 2016 to 2020 rainfall data of 17 sites in Zambia were measured. Subsequently, monthly precipitation data were obtained by summation.

Preprocessings

As seen in the data summary Table 1, the raw data is available in varying resolutions and quality and therefore needs to be filtered and homogenized before use. As for the experimental setup we used Google Earth Engine (GEE). A cloud computing platform that was launched by Google, in 2010. GEE provides free access to numerous remotely sensed datasets as well as computing power, facilitating big geo data processing and analysis [54]. In addition, geemap -a Python package for interactive mapping with Google Earth Engine and Google Colaboratory were also used for the construction of the drought index. For computational reasons, we restricted the study period to the growing season (November-April) [59] from 2012 until 2022. All datasets are aggregated as monthly datasets: summation in the case of precipitation and evapotranspiration, and the average for the case of GPP, LST, SM, and NDVI. Monthly minimum and maximum over the entire study period were also computed for all datasets except GPP. Eventually, all datasets are rescaled to 250m spatial resolution using the bicubic interpolation method.

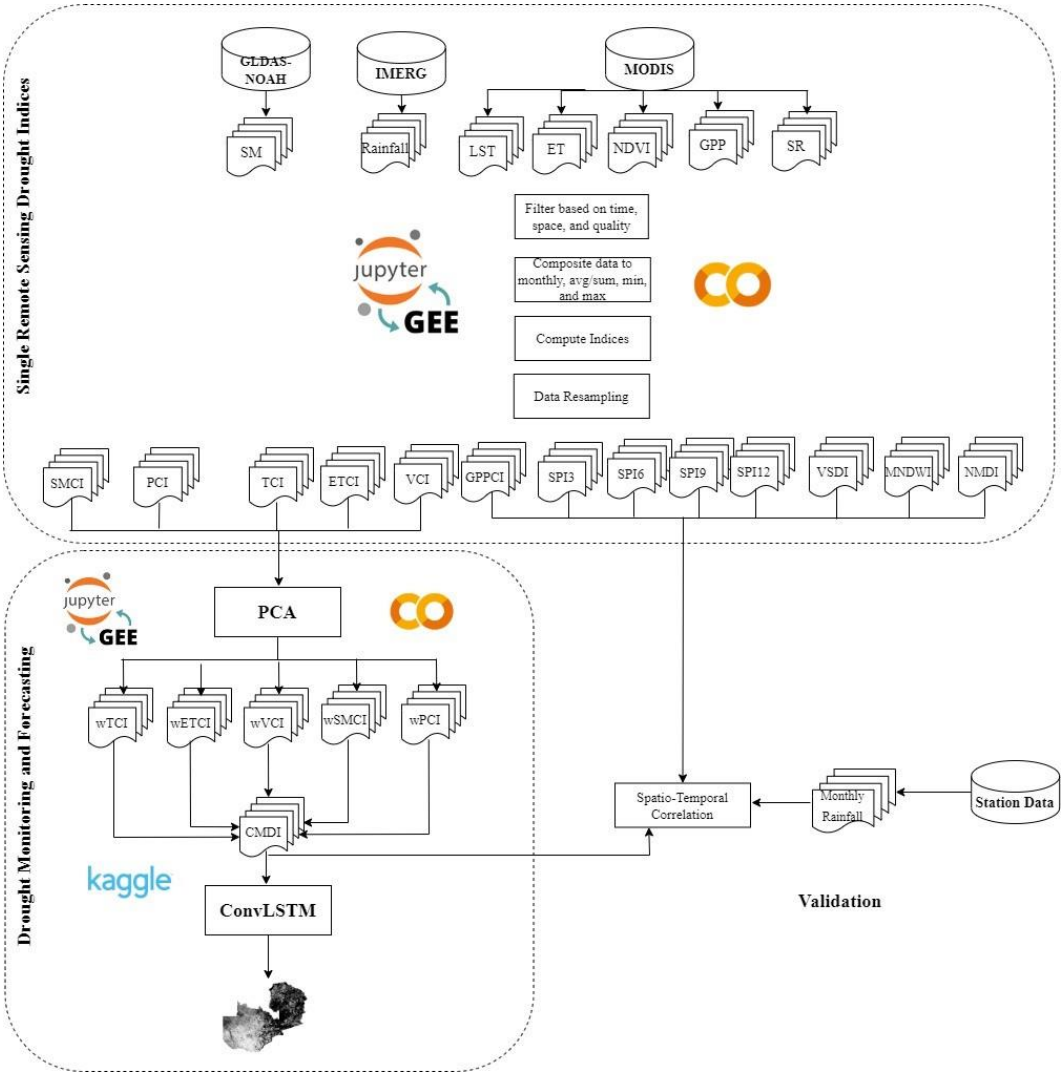


Figure 2 The general flow of methodology followed for this study

Methods

Remote Sensing Idices

As input parameters for the construction of the drought index in the study areas, we selected the variables related to agricultural and meteorological drought conditions. The precipitation factor greatly influences meteorological drought which can be calculated using the precipitation condition index (PCI). Crop growth is hindered by high surface temperature, which can be measured in this study using the temperature condition index (TCI). Vegetation shows its response to a state of low precipitation and deficit soil moisture. This situation can be measured using the VCI, SMCI, and ETCI drought indices. Thus, the five variables used for model input parameters are presented in Table 3 with detailed descriptions.

- Precipitation condition index (PCI):

The Precipitation Condition Index (PCI), reflects meteorological drought. It was defined for the detection of the precipitation deficits from climate signal and it has the following expression:

$$PCI = \frac{P_i - P_{min}}{P_{max} - P_{min}} \quad (1)$$

Where P_i is the smoothed monthly precipitation. P_{max} and P_{min} are maximum and minimum precipitation, respectively for the corresponding month during the entire study period. PCI ranges from 0 to 1. In case of a meteorological drought which has an extremely low precipitation, the PCI is close or equal to 0, and at flooding conditions, the PCI is close to 1.

- Vegetation condition index (VCI):

The Vegetation Condition Index (VCI)is a pixel based normalization of the NDVI. VCI was developed by Kogan et al. (1995) [60] to control for local differences in ecosystem productivity by reducing the influences of geographical and ecological factors to the spatial variability of NDVI [61]. VCI is recognized as a better indicator of drought stress on vegetation than NDVI [61]. The VCI is computed as:

$$VCI = \frac{NDVI_i - NDVI_{min}}{NDVI_{max} - NDVI_{min}} \quad (2)$$

Where $NDVI_i$ is the smoothed monthly *NDV I* value , $NDVI_{min}$ and $NDVI_{max}$ are the absolute minimum and maximum *NDV I* values respectively in the corresponding month during the entire study period. VCI changes from 0 to 1, corresponding to the changes in vegetation condition from extremely unfavorable to optimal. In case of an extremely dry month, the vegetation condition is poor and the VCI is close or equal to 0. Whilst at optimal conditions of vegetation, the VCI is close to 1

- Temperature Condition Index (TCI):

The temperature condition index (TCI), a remote sensing based thermal stress indicator is proposed by Kogan et al.(1995) [62] to determine temperature-related drought phenomenon. TCI assumes that drought event will decrease soil moisture and cause land surface thermal stress, and there are a higher LST in the drought year than the same month of normal years. The TCI is computed as:

$$TCI = \frac{LST_{max} - LST_i}{LST_{max} - LST_{min}} \quad (3)$$

Where LST_i is the smoothed monthly *LST* values, LST_{min} and LST_{max} are the minimum and maximum values respectively of *LST* in the corresponding month during the study period. Under drought conditions, the TCI is close to or equal to 0, and in wet conditions the TCI is close to 1.

- Evapotranspiration Condition Index (ETCI):

$$ETCI = \frac{ET_i - ET_{min}}{ET_{max} - ET_{min}} \quad (4)$$

Where ET_i is the smoothed monthly *ET* values, ET_{min} and ET_{max} are the minimum and maximum values respectively, of *ET*in the corresponding month during the study period.

Agricultural crops are sensitive to soil moisture. During drought conditions, the soil moisture deficit in the root zone during various stages of the crop growth cycle will have a profound impact on crop yield. SMCI and ETCI are able to reflect developing short-term dry conditions, thus responding to agricultural drought.

- Soil Moisture Condition Index (SMCI):

$$SMCI = \frac{SM_i - SM_{min}}{SM_{max} - SM_{min}} \quad (5)$$

Where SM_i is the smoothed monthly *SM* values, SM_{min} and SM_{max} are the minimum and maximum values respectively, of *SM*in the corresponding month during the study period.

- Standardized Precipitation Index (SPI):

The Standardized Precipitation Index (SPI) [63] was calculated using the climate indices python package based on the IMERG monthly precipitation data. To calculate SPI, the rainfall time series is first fitted to a gamma probability distribution function which is then transformed into a normal distribution so that the mean SPI is zero. The SPI is recommended by the World Meteorological Organization [64]. It has been effectively used as a validation dataset in several drought monitoring and forecasting scenarios [1, 12–14, 16, 22, 30, 31, 35, 48, 50, 65]. SPI was designed to quantify the precipitation deficit for multiple time scales. These time scales reflect the impact of drought on the availability of different types of water resources. In this study SPI was calculated on 3-, 6- and 9-month time scales, which correspond to the past 3, 6, and 9 months of precipitation totals respectively as these scales have been found to be useful for monitoring various drought types [63]. Positive SPI values indicate greater than mean precipitation and negative values indicate less than mean precipitation. Based on the SPI values, a classification system was developed to define the intensity of the drought events (Table 2).

Table 2 *SPI Drought categories [66]*

Category	SPI Values
Extremely Wet	SPI>=2
Very Wet	1.5=<SPI<2
Moderately Wet	1.0=<SPI<1.5
Near Nromal	-0.99=<SPI<1.0
Moderately Dry	-1.0=<SPI<-1.5
Severely Dry	-1.5=<SPI<-2
Extremely Dry	SPI=<-2

- Modified Normalized Difference Water Index (MNDWI):

A main limitation of the NDWI is that it cannot suppress the signal noise coming from the land cover features of built-up areas efficiently. Moreover, the water body has a stronger absorbability in the SWIR than that in the NIR and the built-up class has stronger radiation in the SWIR than in the NIR band. Based on this finding the MNDWI is defined as follows:

$$MNDWI = \frac{Green - SWIR}{Green + SWIR} \quad (6)$$

Compared to NDWI, water bodies have greater positive values in MNDWI, because water bodies generally absorb more SWIR light than NIR light; soil, vegetation, and built-up classes have smaller negative values because they reflect more SWIR light than green light.

- Normalized Multi-Band Drought Index (NMDI):

NMDI [67] combines information from multiple near-infrared, and short wave infrared channels as shown in the equation below. Which enhances the sensitivity to drought severity. It uses the difference between two liquid water absorption channels centered at 1640 nm and 2130 nm as the soil and vegetation moisture-sensitive band and therefore it is suitable for monitoring soil and vegetation moisture.

$$NMDI = \frac{R_{860nm} - (R_{1640nm} - R_{2130nm})}{R_{860nm} + (R_{1640nm} - R_{2130nm})} \quad (7)$$

- Visible and Shortwave Drought Index (VSDI):

Similar to NMDI, VSDI [68] is also used for monitoring soil and vegetation moisture. Thus, efficient for agricultural drought monitoring over different land-cover types during the plant-growing season. As shown in the equation below, it is based on the reflectance of shortwave infrared (SWIR) red and blue channels.

$$VSDI = 1 - [(\rho SWIR - \rho blue) + (\rho red - \rho blue)] \quad (8)$$

Table 3 *Description of the different indices*

Factors	Index	Source	Reference
Agricultural	Vegetation Condition Index (VCI)	MODIS (NDVI)	Kogan et al. (1990)[69]
	Soil Moisture Condition Index (SMCI)	GLDAS, NOAH	Narasimhan et al.(2005)[12]
	Gross Primary Production Condition Index (GPPCI)	MODIS (GPP)	
Meteorological	Temperature Condition Index (TCI)	MODIS (LST)	Kogan et al. (1995) [62]
	Evapotranspiration Condition Index (ETCI)	MODIS (ET)	Narasimhan et al.(2005) [12]
	Precipitation Condition Index (PCI)	IMERG	Zhang et al. (2013)[49]
	Standardized Precipitation Index (SPI)		Mckee et al. (1993) [70]
Other Remote Sensing indices	Visible and Shortwave Drought Index (VSDI)	MODIS (SR)	Zhang et al. (2013) [68]
	Modified Normalized Difference Water Index (MNDWI)		Xu et al. (2006) [71]
	Normalized Multi-Band Drought Index (NMDI)		Wang et al. (2007) [67]

Drought Model Development

Figure shows the process flow diagram of this study. The proposed approach is divided into two steps: Step 1 uses PCA to derive the combined drought monitoring index based on the drought factors (PCI, VCI, ETCI, TCI, and SMCI). and step 2 uses ConvLSTM uses the outputs from step 1 as training data to forecast CMDI. There are two reasons why the final drought-forecasting model combines two machine learning approaches. First, the model structure becomes complex when all independent variables are used as input variables for

ConvLSTM. This requires much more memory and processing time compared to using PCA. Therefore, the temporal patterns of each drought index and spatial information were used in PCA and ConvLSTM, respectively, to enable drought forecasting even in a memory-limited environment. The second reason is that there is information redundancy in the drought factors. Therefore, retaining relevant information using PCA can reduce the complexity of the problem.

- PCA:

Principal Component Analysis (PCA) [72], is a dimensionality reduction technique, that has been frequently used to reduce the dimension of a large dataset of predictors by transforming it into a smaller subset that retains the maximum of information from the initial dataset. Because of its simple theoretical basis and its straightforwardness, it has been widely and successfully applied to drought research globally [6, 13, 31, 32, 34, 50]. Although the initial selection of drought factors (VCI, TCI, PCI, SMCI, ETCI) are well established in the literature to monitor drought, they are correlated. Furthermore, it is difficult to use these variables to retrieve drought directly due to the non-linear relationship between each other in different seasons and the effects of outliers, especially in a complex environment. Therefore, PCA is performed here to determine the weights for the drought factors. This produces a different weight value per year and per month for each input variable (Figure 4). When the CMDI is calculated for February, the weights will be different from when the CMDI is calculated for April, etc. The main steps of this technique include (1) Standardization of the data (2) Calculation of the covariance matrix. (3) Computing the eigenvectors and eigenvalues of the covariance matrix (4) Solving the principal components, generally only considering the principal components with eigenvalues exceeding 1. (5) Calculation of the contribution rate.
In this last step, the aim is to determine the weights of the input variables. These weights are derived from the loadings of the eigenvector that corresponds to the principal component that holds most of the variance in the dataset. In this study we considered a cumulative contribution rate exceeding 85% that corresponds with an eigenvalue exceeding 1. Next, a linear model of the combined drought index can be constructed. The resulting

comprehensive index will have the following formula:

$$CMDIy,m = WV\,CI,y,m * V\,CIy,m + WTCI,y,m * TCIy,m + WPCI,y,m * PCIy,m + WETCI,y,m * ETCIy,m + WSMCI,y,m * SMCIy,m \quad (9)$$

Where y is the year (2012 to 2022), and m is the month of the year (November, December, January, February, March, April).
W is the optimal weighting for the input variables based on the PCA results.
VCI is the vegetation condition index, TCI is the temperature condition index, ETCI is the evapotranspiration condition index, PCI is the precipitation condition index, SMCI is the soil moisture condition index, CMDI is the combined drought monitoring index.

Table 4 CMDI Drought Categories

Category	CMDI
Exceptional drought	0.0<=CMDI<0.1
Extreme drought	0.1<=CMDI<0.2
Severe drought	0.2<=CMDI<0.3
Moderate drought	0.3<=CMDI<0.4
Abnormally dry	0.4<=CMDI<0.5
No drought	CMDI=>0.5

In order to evaluate the performance of PCA, we compared the CMDI with other agricultural and meteorological drought indices (Table.3), using: Pearson's correlation coefficient (R):

$$R = \frac{n(\sum(y\hat{y}) - \sum(y)\sum(\hat{y})}{\sqrt{[n\sum(y^2) - (\sum(y))^2][n\sum(\hat{y}^2) - (\sum(\hat{y}))^2]}} \quad (10)$$

n is the number of samples and y and $\hat{y}$ are the values of CMDI and the validation drought indices, respectively. A good agreement between the two indices corresponds to a value of R close to 1.

• ConvLSTM

Spatiotemporal prediction of drought is challenging due to the constant alterations of precipitation regimes coupled with the recent trends of climate change, which require dependencies on both temporal and spatial domains. Long Short-Term Memory (LSTM) [73] architectures can handle temporal correlations very effectively, but they fail to capture spatial information well, as the input, hidden states, cells, and gates are all vectors. To circumvent this issue, ConvLSTM was introduced by Shi et al.(2015) [74]. In addition to the long-short term modeling, ConvLSTM takes 3D data as input, such as our CMDI dataset, and it replaces matrix multiplication with convolution operations at each gate in the LSTM cell. By doing so, it captures underlying spatial features by convolution operations in multiple-dimensional data. A standard ConvLSTM cell is illustrated in Figure 3, and the key equations of are as follows:

$$it = \sigma(Wxi * Xt + Whi * Ht{-}1 + bi) \quad (11)$$

$$ft = \sigma(Wxf * Xt + Whf * Ht{-}1 + bf) \quad (12)$$

$$ot = \sigma(Wxo * Xt + Who * Ht{-}1 + bo) \quad (13)$$

$$gt = tanh(Wxc * Xt + Whc * Ht{-}1 + bc) \quad (14)$$

$$Ct = ft \circ Ct{-}1 + it \circ gt \quad (15)$$

$$H_t = o_t \circ tanh(C_t) \quad (16)$$

where i_t, f_t, and o_t are input, forget, and output gate. W is the weight matrix, X_t is the current input data, H_{t-1} is previous hidden state, and C_t is the cell state. For a more detailed explanation of ConvLSTM, the reader is refered to Shi et al.(2015) [74].

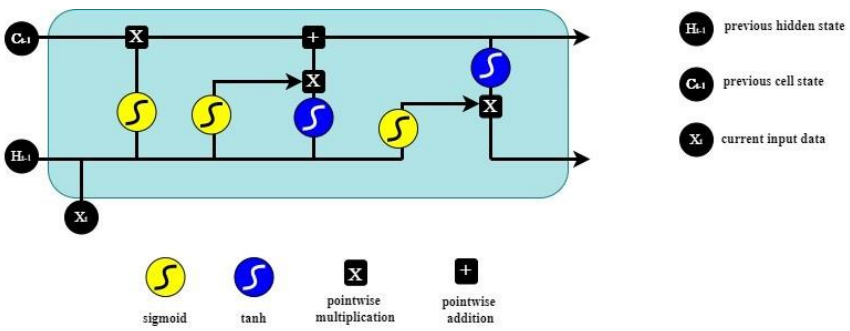


Figure 3: *Illustration of the standard ConvLSTM cell*

In this study, 8 temporally consecutive monthly CMDI were applied to forecast drought conditions in the next month through real-time learning from 2016 to 2022. For example, to forecast drought conditions on January 2018, the ConvLSTMmodel is updated through striding the 8 consecutive monthly data from January 2016 to January 2017. The stride period of 8 months was determined considering the computational efficiency and the impact of annual phenology. The network was is shown in Figure 4 where the model's configuration is many-to-one. The input data *X* of the model receives 8 items of data according to the time interval, and the input data of each node is $256 * 256$. The output are the generated images of the output gate of the last cell, which is the expected CMDI for the input data.

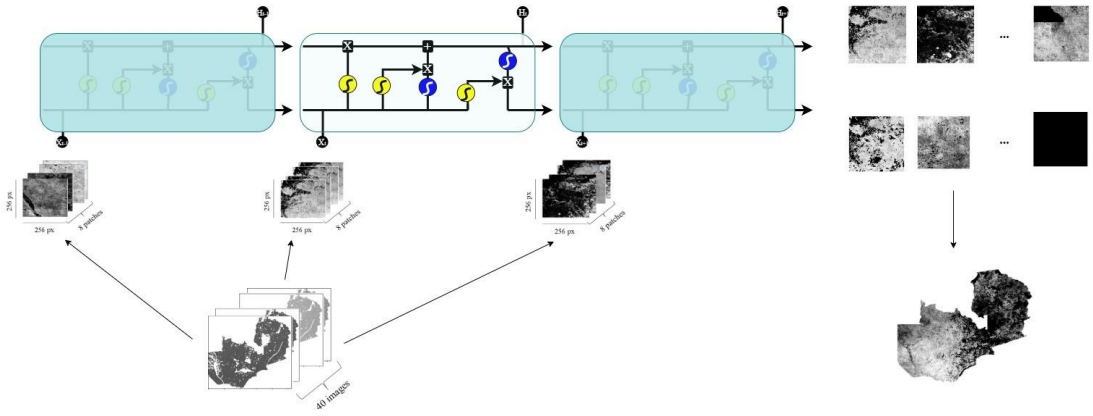


Figure 4: *ConvLSTM architecture*

Training Procedure

We patchify the 40 CMDI timestamps that cover all rainy seasons (November-April) from 2016 to 2022 into $256 * 256$ patches. We split the resultant 340 data samples into 306 training volumes and 34 validation volumes. We then use a batch of 8 by randomly choosing 8 patches from the 340 volumes. After testing various combinations of parameters, the ConvLSTM structure (Figure 4) was determined to have 3 layers. We replaced the standard convolutions with depth-wise separable convolutions to reduce computational overhead. These convolutions have shown their efficiency in terms of memory usage [75]. We finally train our networks for just 50 epochs using ADAM optimizer [76] and with a constant learning rate $lr = 1.10^4$, and it took about $1h7min$ to learn (one-stack ConvLSTM). The Root Mean Square Error (RMSE) was used to measure the prediction accuracy.

(17)

n is the number of samples and y and $y\hat{}$ are the values of reference and predicted drought indices, respectively. The forecasting skills are useful when RMSE is closer to 0. All experiments were run on Kaggle, and the testbed environment configuration was as follows:

- CPU: Intel(R) Xeon(R) 16GB @ 2.30GHz;
- GPU: NVIDIA TESLA P100 13GB;
- RAM: 20 GB;
- Framework: Pytorch 1.11.0, Python 3.7.12

3. Results and Discussion

Construction of the Cloud-based Multisource Drought Index (CMDI)

As mentioned before, the main factors involved in the occurrence, development, and evolution of drought were selected to construct the combined drought indicator. These factors are the standardized forms of the NDVI, LST, ET, SM, and precipitation, denoted as VCI, TCI, ETCI, SMCI, and PCI, respectively. PCA is applied to integrate all the input parameters for the development of the CMDI. PCA takes care of the monthly variations in the input variables and their changing percentage contribution with respect to space and time. Which makes it effective for drought monitoring. The first principal component was chosen because it holds most of the variance in the dataset. The importance of each factor was then obtained based on the eigenvector loadings corresponding to the chosen PC. Next, the importance of each factor was squared, and a linear model of the combined drought index was constructed. Table 6 indicates the specific weights for each input parameter (VCI, TCI, PCI, ETCI, and SMCI) for the rainy season of 2019-20. Using April 2020 as an example, fig.5 represents the spatial distribution of the contribution of each input parameter and the corresponding CMDI model is as

follows:

$$CMDI2020,4 = 0.17 * VCI2020,4 + 0.05 * TCI2020,4 + 0.09 * PCI2020,4 + 0.46 * ETCI2020,4 + 0.26 * SMCI2020,4 \quad (18)$$

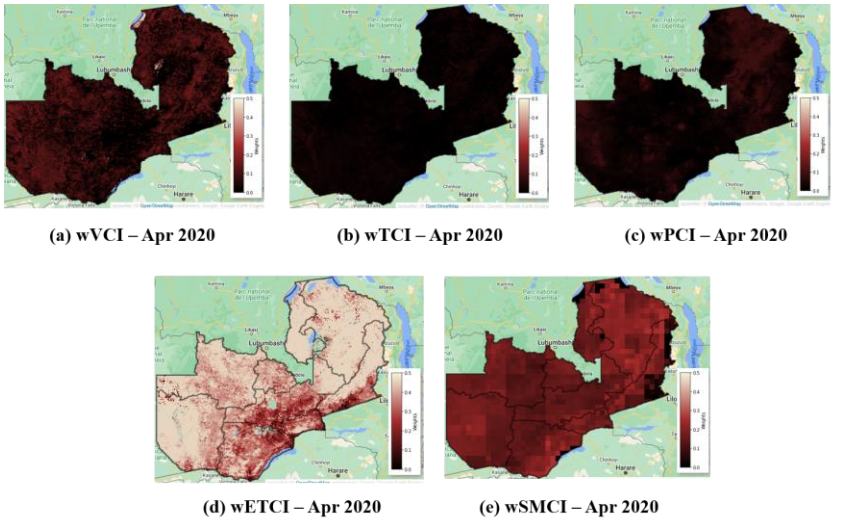


Figure 5: *Spatial pattern of percent contribution of vegetation condition index (VCI), temperature condition index (TCI), precipitation condition index (PCI), evapotranspiration condition index (ETCI), and soil moisture condition index (SMCI) over Zambia for April 2020*

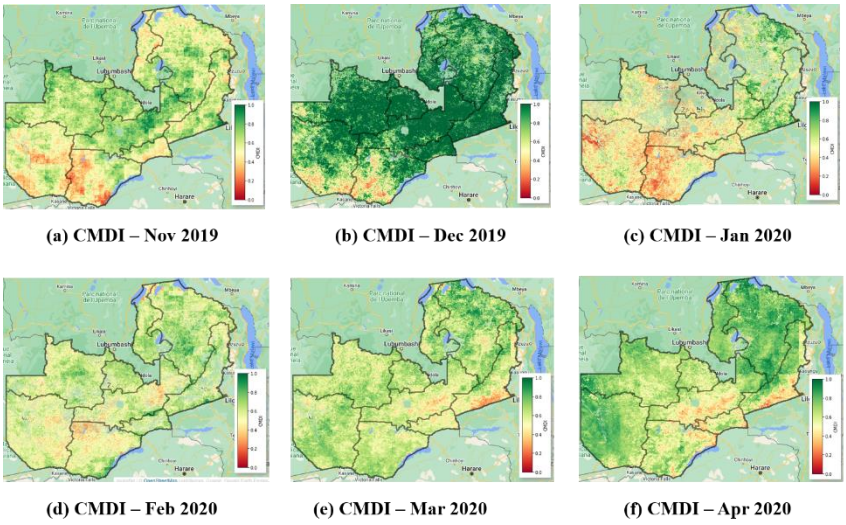
On an annual scale, during the rainy season of 2019-20, SMCI is the highest contributing factor (34%), in the CMDI followed by VCI (20%), ETCI (18.5%), TCI(18%), and PCI (13.5%). The time series of CMDI maps (Figure 6) have highlighted the rainy season of 2019-2020, in particular, the months of January, February, March, and April as one of the driest months over the study period. During these months, Zambia experienced moderate-to-extreme drought conditions, and the average CMDI values for most of the regions were less than 0.5. This drought severity continued for the next 2 years During this period, the average spatial drought spread varied from 25% to 90% with respect to each month.

In the past 8 years (2015–2022), the PCA-based method detected the rainy seasons of 2015-16, 2019-20, 202021, and 2021-22 as severe drought years, with conditions ranging from moderate to extreme drought across the season. Appendix A shows the spatial distribution of dry spells in these years. These results align with the findings of Hulsman et al.(2021) [77], Alemaw et al. (2022) [78], and the Famine Early Warning Systems Network [79]. For the agriculture sector, in particular, these observed months are crucial for the overall cropping cycle. The increasing drought frequencies over these months heavily impact the sowing activities and it may also directly impact the following harvest season in Zambia. For the study area as a whole, the Southern, NorthWestern, Western, Muchinga, and Luapula regions were frequently under rigorous dry spells during the rainy season from 2015-2022. According to Kulkarni et al.(2020) [34] 3-4 occurrences of severe to extreme drought events over any region in the span of 18 years is alarming. Higher elevation and less availability of surface water resources are the common characteristics for almost all of the regions mentioned

above. These characteristics might have acted as additional parameters for the aggravation of drought occurrence and severity and thus the reduction in agricultural productivity over the 2019–2020 rainy season.

Table 5: PCA-based spatially averaged monthly weight values of the input parameter for the rainy season 20212022

2019-20	VCI	TCI	PCI	ETCI	SMCI
Nov	0.12	0.1	0.27	0.13	0.4
Dec	0.16	0.19	0.19	0.15	0.34
Jan	0.21	0.19	0.06	0.1	0.47
Feb	0.23	0.26	0.14	0.09	0.3
Mar	0.27	0.25	0.06	0.18	0.27
Apr	0.17	0.05	0.09	0.46	0.26



(a) CMDI – Nov 2019 (b) CMDI – Dec 2019 (c) CMDI – Jan 2020

(d) CMDI – Feb 2020 (e) CMDI – Mar 2020 (f) CMDI – Apr 2020

Figure 6: *Spatio-Temporal representation of PCA-based CMDI maps for the rainy season 2019–20*

Table 6: PCA-based spatially averaged monthly weight values of the input parameter for the rainy season 2015-
2016

2015-16	VCI	TCI	PCI	ETCI	SMCI
Nov	0.19	0.14	0.23	0.2	0.27
Dec	0.23	0.15	0.17	0.22	0.25
Jan	0.24	0.03	0.28	0.1	0.36
Feb	0.27	0.06	0.14	0.19	0.37
Mar	0.44	0.06	0.21	0.07	0.25
Apr	0.24	0.1	0.04	0.36	0.3

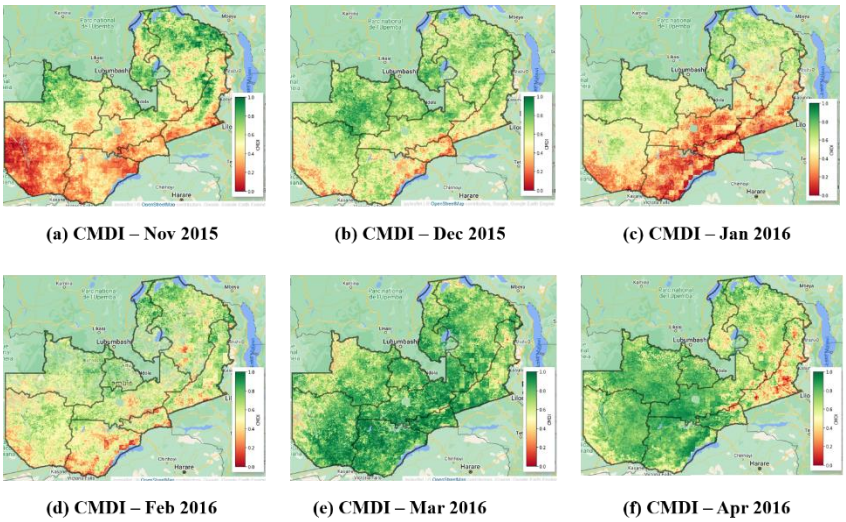


Figure 7: Spatio-Temporal representation of PCA-based CMDI maps for the rainy season 2015–16

CMDI and its relation to the Gross Primary Production (GPP):

Several studies have noted the significant impacts of severe to exceptional drought conditions on crop yield [80]. Zambia in particular has shown the highest reduction risks in maize crops during the drought years [24, 81]. Intense dry spells in Zambia always have a dire influence on the community and environment. Therefore, in this study, we have tried to understand the Spatio-temporal association between GPP and the monthly CMDI intensities using the statistical correlation method. Scatter plots and the corresponding trends were plotted, for a selection of data points (n=2500) that cover the agricultural areas of Zambia during April 2020. To reduce the influence of other non-agro-climatological parameters on crop production, all the GPP datasets were detrended as shown in figure 8. The analysis results show that the correlation between CMDI and GPP is significant ($P < 0.01$) and positive. The correlation coefficient r is 0.59. Although not shown here, the r values of the remaining years are less different. Therefore, CMDI can reflect the changes in the GPP in different months of each year between 2012 and 2022, and so CMDI can reflect temporal changes in the GPP.

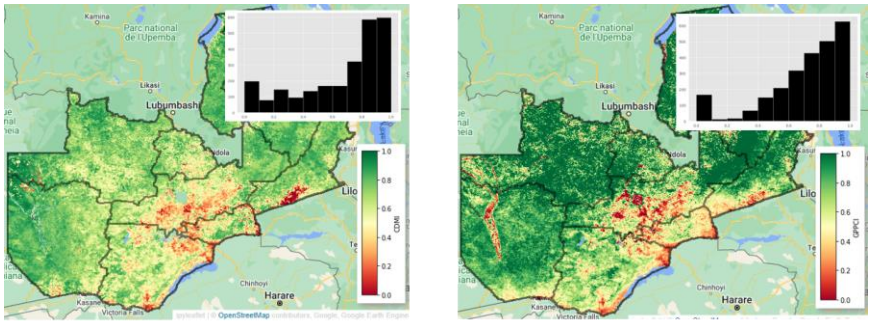


Figure 8 Correlation of the CMDI and GPPCI for April 2020

Comparison of CMDI with the standardized precipitation index (SPI) and other remote sensing indices:

To further illustrate the applicability of the proposed CMDI in drought monitoring, the CMDI was compared with other drought indicators. Because most drought factors used to construct the CMDI are common monitoring indicators of meteorological and agricultural drought. The SPI is a widely used and well-established meteorological drought index that has been used to construct other comprehensive drought monitoring indices or drought monitoring and prediction systems as well as to validate newly proposed drought monitoring indices. Therefore, the SPI calculated based on the IMERG precipitation data, as a representative index of meteorological drought, and was utilized to evaluate the suitability of CMDI as well as the other indicators for meteorological drought monitoring. The correlation analyses were conducted between the SPI at 3, 6, 9, and 12 months time scale, with the following indices: CMDI, VCI, TCI, PCI, SMCI, ETCI, VSDI, MNDWI, and NMDI for February 2020. As shown in Table 7, the correlations between SPI-1, SPI-3, SPI-6, and SPI-9 and CMDI are generally higher than those of any of the single remote sensing drought indices. The range of correlation coefficients R is 0.407 0.532. This may be because of the influence of factors other than precipitation meteorological drought is quite significant. The correlation coefficients between the proposed drought index and SPI-3 and SPI-6 are the largest among all of the remote sensing indices. This finding indicates that the CMDI is superior to other drought indices in reflecting short- and medium-term meteorological droughts.

Although the correlations between other remote sensing indices and the SPI at different time scales are not as large as the correlations of the CMDI, meaningful observations can still be made: (1) VSDI, MNDWI, NDMI are not related to any of the SPIs, indicating that these indices are not suitable for monitoring meteorological droughts in the study area. This is understandable because they are based on the spectral reflectance of vegetation, and thus more suitable for monitoring agricultural droughts than other types of droughts. Furthermore, we can see that MNDWI and VSDI are negatively correlated to GPPCI. (2) Although insignificant, the correlations between the SPIs and the different components of CMDI (VCI, TCI, PCI, ETCI, and SMCI) show, that PCI has the highest correlation among those indices with the SPIs. Which further emphasizes that precipitation is one of the main factors for drought. Followed by SMCI, and this highlights the relationship between meteorological and agricultural drought. The correlations between VCI and GPPCI, and ETCI and GPPCI are

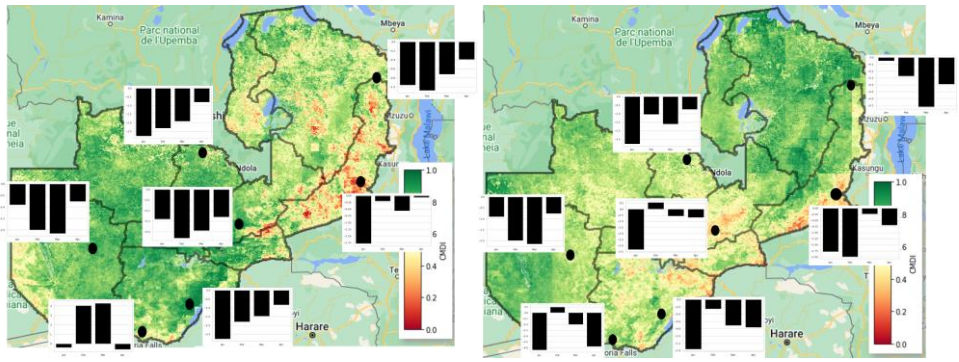
higher compared to the other factors. This highlights the influence that ET exerts on agricultural drought is larger than those on meteorological drought.
Based on these findings, the proposed drought index can reflect both meteorological drought and agricultural drought. We think that the introduction of the ET and SM variables further improved the performance of CMDI.

Table 7 *Correlation coefficients (r) for remote sensing indices with the SPI (Notes: P<.01 for all the experiments, and the bold values are the maximum value of each column)*

Remote Sensing Index	SPI3	SPI6	SPI9	SPI12	GPPCI
MNDWI	-0.069	-0.058	-0.06	0.003	-0.398
NMDI	0.002	-0.004	-0.003	0.005	-0.083
VSDI	-0.117	-0.094	-0.096	-0.016	-0.368
VCI	-0.105	-0.080	-0.085	-0.055	0.319
TCI	0.060	0.014	0.017	-0.001	-0.042
PCI	0.353	0.261	0.392	0.229	-0.075
ETCI	0.021	0.007	0.012	0.006	0.365
SMCI	0.074	0.254	0.225	0.408	-0.170
CMDI	0.632	0.513	0.407	0.445	0.697

Comparison of CMDI with station-based monthly rainfall

We also compare the Rainfall anomaly with the spatial distribution of drought monitored by the CMDI during April 2016 and April 2020 (Figure ??). It should be noted that in Figure ??, data are missing from some stations on certain months, and certain days. Despite these limitations, there is an observed consistency and agreement in the spatial distributions of precipitation and the CMDI. Areas with large CMDI values have large monthly precipitation values; similarly, areas with small CMDI values have small monthly precipitation values. This finding indicates that the CMDI can accurately determine the spatial distribution of drought; thus, it can be used to monitor the Spatio-temporal distribution of drought and its evolution. As shown in the maps in Figure ??, the Spatio-temporal evolution of drought in Zambia from November 2019 to January 2020 shows the formation of drought in nearly the whole country (CMDI < 0.5). Over this rainy season, the drought intensity gradually increased, and the differences in regional drought began to appear.



(a) (b)

Figure 9 *Spatial distributions of the CMDI and monthly rainfall anomaly during April 2016 (left) and April 2020 (right)*

CMDI Forecasting using ConvLSTM

We trained the proposed model (Figure 4) with the Adam optimizer at a learning rate of $1e^{-4}$ and weight decay of $1e^{-5}$. We used random shuffling mini-batches for learning; the mini-batch size was set to 12. It seems that the depth-wise separable convolution operation efficiently extracted the underlying features from the data and enabled quick training. ConvLSTM took about 2h 34min to learn (one-stack ConvLSTM) and has required 50 epochs to reach a loss of 0.14 on the validation dataset. The two-stacked ConvLSTM model showed more stable performance than the one-stacked ConvLSTM. However, it took longer to run.

Our experiments show that using 6 consecutive CMDI images is a good compromise in terms of results and computational time. Figure 10 shows the CMDI and its prediction using ConvLSTM for April 2020. We used CMDI images of April 2019, November 2020, December 2020, Jan 2020, and March 2020 as input data. We can see that both images correlate well. But ConvLSTM seems to overestimate the drought severity in some regions.

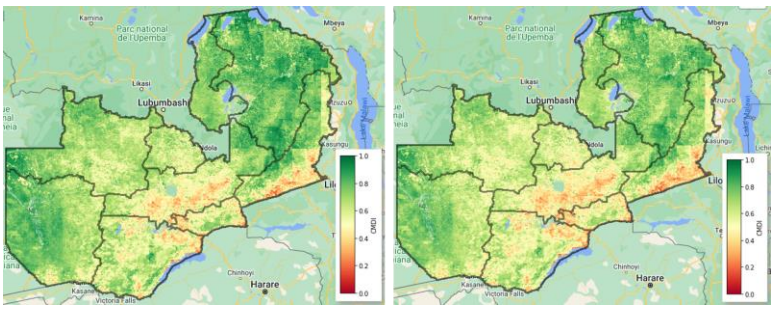


Figure 10 *Spatio-Temporal prediction of CMDI for April 2022: CMDI(left), prediction (right)*

To generate the CMDI prediction for November 2022 (Figure 11), we used the previous 6 CMDI images (i.e, November 2020, December 2020, January 2021, February 2021, March 2021, and April 2021).

Based on this prediction, we can see that the North-western, Eastern, and Muchinga regions are at high risk of drought during the rainy season of 2022-2023.

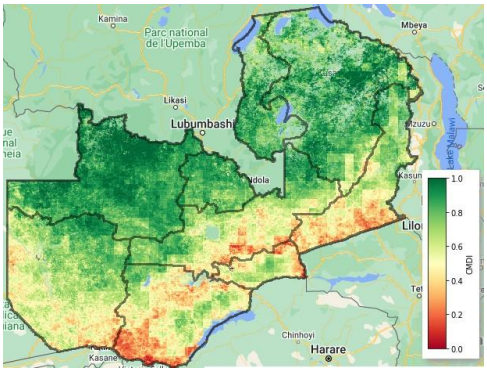


Figure 11 *Spatio-Temporal prediction of CMDI for November 2022*

4. Conclusion

Spatiotemporal monitoring and forecasting of drought within the country's capabilities are crucial to reducing the socio-economic damage, through informed decision-making. Taking Zambia as a pilot site, in this paper we provided such a system, PCA-ConvLSTM, by deriving and forecasting a new drought index and leveraging open-source platforms such as Google Earth Engine, Google Colab, and Kaggle to carry out the analysis of global geospatial big data at scale and training our model. For the accurate assessment of agro-climatological drought severities, CMDI integrates the different explanatory variables of drought (temperature, evapotranspiration, vegetation, soil moisture, and precipitation) using PCA. Considering the rainy season from 2012-2022, CMDI was capable of identifying the spatial extent of historical and current drought events over the study area. Moreover, the spatial and temporal distribution of drought determined by the CMDI agrees with that of SPI-3, SPI-6, SPI-9, and SPI-12 month scale data, monthly station rainfall data, and GPP, showing a high correlation. In the later stages of this study, ConvLSTM was used to forecast drought conditions for the next rainy season of 2022-23.

Our findings show that developing drought early warning systems using only satellite remote sensing data when the measured data are incomplete can be realized through the CMDI. This system is also scalable since it doesn't take into consideration the mechanism of drought-induced interaction between drought-causing factors. However, our system has room for improvement.

All the input parameters used for this study were initially available in different spatial resolutions and at a coarse spatial resolution. Future work should experiment with high-resolution input data. We can further improve the forecasting performance of ConvLSTM with mechanisms like attention.

References

[1] H. Han, J. Bai, J. Yan, H. Yang, and G. Ma, "A combined drought monitoring index based on multi-sensor remote sensing data and machine learning," *Geocarto International*, vol. 36, no. 10, pp. 1161–1177, 2021.

[2] F. Mou, T.-A. Akwasi, M. Li, M. Zhu, Y. He, Z. He, Y. Xiao, J. Ren, J. Xia, X. Zhang, *et al.*, "Drought monitoring in sub-sahara africa," in *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*, pp. 6902–6905, IEEE, 2020.

[3] F. A. Prodhan, J. Zhang, F. Yao, L. Shi, T. P. Pangali Sharma, D. Zhang, D. Cao, M. Zheng, N. Ahmed, and H. P. Mohana, "Deep learning for monitoring agricultural drought in south asia using remote sensing data," *Remote Sensing*, vol. 13, no. 9, p. 1715, 2021.

[4] R. Inoubli, A. B. Abbes, I. R. Farah, V. Singh, T. Tadesse, and M. T. Sattari, "A review of drought monitoring using remote sensing and data mining methods," in *2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 1–6, IEEE, 2020.

[5] R. Shen, A. Huang, B. Li, and J. Guo, "Construction of a drought monitoring model using deep learning based on multi-source remote sensing data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 79, pp. 48–57, 2019.

[6] N. Neeti, C. A. Murali, V. Chowdary, N. Rao, and M. Kesarwani, "Integrated meteorological drought monitoring framework using multi-sensor and multi-temporal earth observation datasets and machine learning algorithms: A case study of central india," *Journal of Hydrology*, vol. 601, p. 126638, 2021.

[7] J. Rhee and J. Im, "Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data," *Agricultural and Forest Meteorology*, vol. 237, pp. 105–122, 2017.

[8] F. A. Prodhan, J. Zhang, T. P. P. Sharma, L. Nanzad, D. Zhang, A. M. Seka, N. Ahmed, S. S. Hasan, M. Z. Hoque, and H. P. Mohana, "Projection of future drought and its impact on simulated crop yield over south asia using ensemble machine learning approach," *Science of The Total Environment*, vol. 807, p. 151029, 2022.

[9] H. West, N. Quinn, and M. Horswell, "Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities," *Remote Sensing of Environment*, vol. 232, p. 111291, 2019.

[10] C. Adede, R. Oboko, P. W. Wagacha, and C. Atzberger, "A mixed model approach to vegetation condition prediction using artificial neural networks (ann): case of kenya's operational drought monitoring," *Remote Sensing*, vol. 11, no. 9, p. 1099, 2019.

[11] B. Ayugi, E. O. Eresanya, A. O. Onyango, F. K. Ogou, E. C. Okoro, C. O. Okoye, C. M. Anoruo, V. N. Dike, O. R. Ashiru, M. T. Daramola, *et al.*, "Review of meteorological drought in africa: historical trends, impacts, mitigation measures, and prospects," *Pure and Applied Geophysics*, pp. 1–22, 2022.

[12] B. Narasimhan and R. Srinivasan, “Development and evaluation of soil moisture deficit index (smdi) and evapotranspiration deficit index (etdi) for agricultural drought monitoring,” *Agricultural and forest meteorology*, vol. 133, no. 1-4, pp. 69–88, 2005.

[13] L. Du, Q. Tian, T. Yu, Q. Meng, T. Jancso, P. Udvardy, and Y. Huang, “A comprehensive drought monitoring method integrating modis and trmm data,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 23, pp. 245–253, 2013.

[14] X. Liu, X. Zhu, Q. Zhang, T. Yang, Y. Pan, and P. Sun, “A remote sensing and artificial neural networkbased integrated agricultural drought index: Index development and applications,” *Catena*, vol. 186, p. 104394, 2020.

[15] X. Zhou, P. Wang, K. Tansey, S. Zhang, H. Li, and L. Wang, “Developing a fused vegetation temperature condition index for drought monitoring at field scales using sentinel-2 and modis imagery,” *Computers and electronics in agriculture*, vol. 168, p. 105144, 2020.

[16] M. Achite, M. Jehanzaib, N. Elshaboury, and T.-W. Kim, “Evaluation of machine learning techniques for hydrological drought modeling: A case study of the wadi ouahrane basin in algeria,” *Water*, vol. 14, no. 3, p. 431, 2022.

[17] P. Feng, B. Wang, D. Li Liu, and Q. Yu, “Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in south-eastern australia,” *Agricultural Systems*, vol. 173, pp. 303–316, 2019.

[18] R. Maity, M. I. Khan, S. Sarkar, R. Dutta, S. S. Maity, M. Pal, and K. Chanda, “Potential of deep learning in drought assessment by extracting information from hydrometeorological precursors,” *Journal of Water and Climate Change*, vol. 12, no. 6, pp. 2774–2796, 2021.

[19] E. Ghoneim, A. Dorofeeva, M. Benedetti, D. Gamble, L. Leonard, and M. AbuBakr, “Vegetation drought analysis in tunisia: A geospatial investigation,” *Journal of Atmospheric Earth and Science*, vol. 1, no. 002, 2017.

[20] Y. Han, Z. Li, C. Huang, Y. Zhou, S. Zong, T. Hao, H. Niu, and H. Yao, “Monitoring droughts in the greater changbai mountains using multiple remote sensing-based drought indices,” *Remote Sensing*, vol. 12, no. 3, p. 530, 2020.

[21] M. H. Bazrkar, J. Zhang, and X. Chu, “Hydroclimatic aggregate drought index (hadi): a new approach for identification and categorization of drought in cold climate regions,” *Stochastic Environmental Research and Risk Assessment*, vol. 34, no. 11, pp. 1847–1870, 2020.

[22] S. Park, J. Im, E. Jang, and J. Rhee, “Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions,” *Agricultural and forest meteorology*, vol. 216, pp. 157–169, 2016.

[23] M. Hassanein, A. Khalil, and Y. Essa, “Assessment of drought impact in africa using standard precipitation evapotranspiration index,” *Journal of Climate*, vol. 11, pp. 75–81, 2013.

[24] H. Chikoore and M. R. Jury, “South african drought, deconstructed,” *Weather and Climate Extremes*, vol. 33, p. 100334, 2021.

[25] H. Balti, A. B. Abbes, N. Mellouli, I. R. Farah, Y. Sang, and M. Lamolle, “A review of drought monitoring with big data: Issues, methods, challenges and research directions,” *Ecological Informatics*, vol. 60, p. 101136, 2020.

[26] L. Crocetti, M. Forkel, M. Fischer, F. Jurečka, A. Grlj, A. Salentinig, M. Trnka, M. Anderson, W.-T. Ng, Ž. Kokalj, *et al.*, “Earth observation for agricultural drought monitoring in the pannonian basin (southeastern europe): current state and future directions,” *Regional environmental change*, vol. 20, no. 4, pp. 1–17, 2020.

[27] H. Park, K. Kim, and D. K. Lee, “Prediction of severe drought area based on random forest: Using satellite image and topography data,” *Water*, vol. 11, no. 4, p. 705, 2019.

[28] A. Belayneh and J. Adamowski, “Drought forecasting using new machine learning methods,” *Journal of Water and Land Development*, 2013.

[29] S. Park, E. Seo, D. Kang, J. Im, and M.-I. Lee, “Prediction of drought on pentad scale using remote sensing data and mjo index through random forest over east asia,” *Remote Sensing*, vol. 10, no. 11, p. 1811, 2018.

[30] H.-J. Chu, R. F. Wijayanti, L. M. Jaelani, and H.-P. Tsai, “Time varying spatial downscaling of satellitebased drought index,” *Remote Sensing*, vol. 13, no. 18, p. 3693, 2021.

[31] K. Hazaymeh and Q. K. Hassan, “A remote sensing-based agricultural drought indicator and its implementation over a semi-arid region, jordan,” *Journal of Arid Land*, vol. 9, no. 3, pp. 319–330, 2017.

[32] J.-l. Wang and Y.-h. Yu, “Comprehensive drought monitoring in yunnan province, china using multisource remote sensing data,” *Journal of Mountain Science*, vol. 18, no. 6, pp. 1537–1549, 2021.

[33] S. Tirivarombo, D. Osupile, and P. Eliasson, “Drought monitoring and analysis: standardised precipitation evapotranspiration index (spei) and standardised precipitation index (spi),” *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 106, pp. 1–10, 2018.

[34] S. S. Kulkarni, B. D. Wardlow, Y. A. Bayissa, T. Tadesse, M. D. Svoboda, and S. S. Gedam, “Developing a remote sensing-based combined drought indicator approach for agricultural drought monitoring over marathwada, india,” *Remote Sensing*, vol. 12, no. 13, p. 2091, 2020.

[35] S. Park, J. Im, D. Han, and J. Rhee, “Short-term forecasting of satellite-based drought indices using their temporal patterns and numerical model output,” *Remote Sensing*, vol. 12, no. 21, p. 3499, 2020.

[36] B. Musonda, Y. Jing, V. Iyakaremye, and M. Ojara, “Analysis of long-term variations of drought characteristics using standardized precipitation index over zambia,” *Atmosphere*, vol. 11, no. 12, p. 1268, 2020.

[37] P. Mahmoudi, A. Rigi, and M. Miri Kamak, “A comparative study of precipitation-based drought indices with the aim of selecting the best index for drought monitoring in iran,” *Theoretical and Applied Climatology*, vol. 137, no. 3, pp. 3123–3138, 2019.

[38] S. Shukla, A. C. Steinemann, and D. P. Lettenmaier, “Drought monitoring for washington state: Indicators and applications,” *Journal of Hydrometeorology*, vol. 12, no. 1, pp. 66–83, 2011.

[39] C. Lai, R. Zhong, Z. Wang, X. Wu, X. Chen, P. Wang, and Y. Lian, “Monitoring hydrological drought using long-term satellite-based precipitation data,” *Science of the total environment*, vol. 649, pp. 1198– 1208, 2019.

[40] F. Li, H. Li, W. Lu, G. Zhang, and J.-C. Kim, “Meteorological drought monitoring in northeastern china using multiple indices,” *Water*, vol. 11, no. 1, p. 72, 2019.

[41] Z. T. Abdulrazzaq, R. H. Hasan, and N. A. Aziz, “Integrated trmm data and standardized precipitation index to monitor the meteorological drought,” *Civil Engineering Journal*, vol. 5, no. 7, pp. 1590–1598, 2019.

[42] X. Zhang, Z. Su, J. Lv, W. Liu, M. Ma, J. Peng, and G. Leng, “A set of satellite-based near real-time meteorological drought monitoring data over china,” *Remote Sensing*, vol. 11, no. 4, p. 453, 2019.

[43] A. Shalishe, A. Bhowmick, and K. Elias, “Meteorological drought monitoring based on satellite chirps product over gamo zone, southern ethiopia,” *Advances in Meteorology*, vol. 2022, 2022.

[44] A. K. Samantaray, M. Ramadas, and R. K. Panda, “Changes in drought characteristics based on rainfall pattern drought index and the cmip6 multi-model ensemble,” *Agricultural Water Management*, vol. 266, p. 107568, 2022.

[45] J. F. Brown, B. D. Wardlow, T. Tadesse, M. J. Hayes, and B. C. Reed, “The vegetation drought response index (vegdri): A new integrated approach for monitoring drought stress in vegetation,” *GIScience & Remote Sensing*, vol. 45, no. 1, pp. 16–46, 2008.

[46] S. Mukherjee, A. Mishra, and K. E. Trenberth, “Climate change and drought: a perspective on drought indices,” *Current Climate Change Reports*, vol. 4, no. 2, pp. 145–163, 2018.

[47] J. Lu, G. J. Carbone, and P. Gao, “Mapping the agricultural drought based on the long-term avhrr ndvi and north american regional reanalysis (narr) in the united states, 1981–2013,” *Applied geography*, vol. 104, pp. 10–20, 2019.

[48] J. Rhee, J. Im, and G. J. Carbone, “Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data,” *Remote Sensing of environment*, vol. 114, no. 12, pp. 2875–2887, 2010.

[49] A. Zhang and G. Jia, “Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data,” *Remote sensing of Environment*, vol. 134, pp. 12–23, 2013.

[50] Y. Zhang, X. Liu, W. Jiao, X. Zeng, X. Xing, L. Zhang, J. Yan, and Y. Hong, “Drought monitoring based on a new combined remote sensing index across the transitional area between humid and arid regions in china,” *Atmospheric Research*, vol. 264, p. 105850, 2021.

[51] Z. Hao and A. AghaKouchak, "Multivariate standardized drought index: a parametric multi-index model," *Advances in Water Resources*, vol. 57, pp. 12–18, 2013.

[52] D. J. Lary, A. H. Alavi, A. H. Gandomi, and A. L. Walker, "Machine learning in geosciences and remote sensing," *Geoscience Frontiers*, vol. 7, no. 1, pp. 3–10, 2016.

[53] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[54] M. Amani, A. Ghorbanian, S. A. Ahmadi, M. Kakooei, A. Moghimi, S. M. Mirmazloumi, S. H. A. Moghaddam, S. Mahdavi, M. Ghahremanloo, S. Parsian, *et al.*, "Google earth engine cloud computing platform for remote sensing big data applications: A comprehensive review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5326–5350, 2020.

[55] Q. Zhao, L. Yu, X. Li, D. Peng, Y. Zhang, and P. Gong, "Progress and trends in the application of google earth and google earth engine," *Remote Sensing*, vol. 13, no. 18, p. 3778, 2021.

[56] J. Thurlow, T. Zhu, and X. Diao, "The impact of climate variability and change on economic growth and poverty in zambia," 2009.

[57] climateknowledgeportal, "climate-data historical zambia." https://climateknowledgeportal.worldbank.org/country/zamb Last accessed on 2022-07-06.

[58] S.-h. Jiang, L.-y. Wei, L.-l. Ren, L.-q. Zhang, M.-h. Wang, and H. Cui, "Evaluation of imerg, tmpa, era5, and cpc precipitation products over mainland china: Spatiotemporal patterns and extremes," *Water Science and Engineering*, 2022.

[59] Food and Agricultural Organization of the United Nations, "Giews - global information and early warning system." https://www.fao.org/giews/countrybrief/country.jsp?code=ZMBlang=ar, Last accessed on 202207-06.

[60] F. N. Kogan, "Droughts of the late 1980s in the united states as derived from noaa polar-orbiting satellite data," *Bulletin of the American Meteorological Society*, vol. 76, no. 5, pp. 655–668, 1995.

[61] S. M. Quiring and S. Ganesh, "Evaluating the utility of the vegetation condition index (vci) for monitoring meteorological drought in texas," *Agricultural and Forest Meteorology*, vol. 150, no. 3, pp. 330–339, 2010.

[62] F. N. Kogan, "Application of vegetation index and brightness temperature for drought detection," *Advances in space research*, vol. 15, no. 11, pp. 91–100, 1995.

[63] T. B. McKee, "Drought monitoring with multiple time scales," in *Proceedings of 9th Conference on Applied Climatology, Boston, 1995*, 1995.

[64] M. Hayes, M. Svoboda, N. Wall, and M. Widhalm, "The lincoln declaration on drought indices: universal meteorological drought index recommended," *Bulletin of the American Meteorological Society*, vol. 92, no. 4, pp. 485–488, 2011.

[65] A. Hadri, M. E. M. Saidi, and A. Boudhar, “Multiscale drought monitoring and comparison using remote sensing in a mediterranean arid region: a case study from west-central morocco,” *Arabian Journal of Geosciences*, vol. 14, no. 2, pp. 1–18, 2021.

[66] M. R. Mansouri Daneshvar, A. Bagherzadeh, and M. Khosravi, “Assessment of drought hazard impact on wheat cultivation using standardized precipitation index in iran,” *Arabian Journal of Geosciences*, vol. 6, no. 11, pp. 4463–4473, 2013.

[67] L. Wang and J. J. Qu, “Nmdi: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing,” *Geophysical Research Letters*, vol. 34, no. 20, 2007.

[68] N. Zhang, Y. Hong, Q. Qin, and L. Liu, “Vsdi: a visible and shortwave infrared drought index for monitoring soil and vegetation moisture based on optical remote sensing,” *International journal of remote sensing*, vol. 34, no. 13, pp. 4585–4609, 2013.

[69] F. N. Kogan, “Remote sensing of weather impacts on vegetation in non-homogeneous areas,” *International Journal of remote sensing*, vol. 11, no. 8, pp. 1405–1419, 1990.

[70] T. B. McKee, N. J. Doesken, J. Kleist, *et al.*, “The relationship of drought frequency and duration to time scales,” in *Proceedings of the 8th Conference on Applied Climatology*, vol. 17, pp. 179–183, California, 1993.

[71] H. Xu, “Modification of normalised difference water index (ndwi) to enhance open water features in remotely sensed imagery,” *International journal of remote sensing*, vol. 27, no. 14, pp. 3025–3033, 2006.

[72] K. Pearson, “Liii. on lines and planes of closest fit to systems of points in space,” *The London, Edinburgh, and Dublin philosophical magazine and journal of science*, vol. 2, no. 11, pp. 559–572, 1901.

[73] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735– 1780, 1997.

[74] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, “Convolutional lstm network: A machine learning approach for precipitation nowcasting,” *Advances in neural information processing systems*, vol. 28, 2015.

[75] F. Chollet, “Xception: Deep learning with depthwise separable convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251–1258, 2017.

[76] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.

[77] P. Hulsman, H. H. Savenije, and M. Hrachowitz, “Satellite-based drought analysis in the zambezi river basin: Was the 2019 drought the most extreme in several decades as locally perceived?,” *Journal of Hydrology: Regional Studies*, vol. 34, p. 100789, 2021.

[78] B. F. Alemaw, “The recent droughts of 2019/20 in southern africa and its teleconnection with enso events,” *Atmospheric and Climate Sciences*, vol. 12, no. 2, pp. 246–263, 2022.

[79] FEWS NET:Famine Early Warning Systems Network , “High food assistance needs persist due to consecutive droughts and erratic 2019/20 rainfall.” https://fews.net/southern-africa/alert/april-2-2020-0, Last accessed on 2022-07-12.

[80] I. Cohen, S. I. Zandalinas, C. Huck, F. B. Fritschi, and R. Mittler, “Meta-analysis of drought and heat stress combination impact on crop yield and yield components,” *Physiologia Plantarum*, vol. 171, no. 1, pp. 66–76, 2021.

[81] G. C. Matchaya, G. Tadesse, and A. N. Kuteya, “Rainfall shocks and crop productivity in zambia: Implication for agricultural water risk management,” *Agricultural Water Management*, vol. 269, p. 107648, 2022.