Drought Prediction by using Normalized Differential Vegetation Index and Long Short-Term Memory Alogrithm

Thiri Maung
Emergency Operations Centre
Ministry of Social Welfare, Relief and Resettlement
Nay Pyi Taw, Myanmar
thirimaung@ucsy.edu.mm

Thin Lai Lai Thein
Data Analytics Lab
University of Compurter Studies, Yangon
Yangon, Myanmar
tllthein@ucsy.edu.mm

***Abstract*— Drought is a natural phenomenon and it can affect the economy and environment of the country. Monitoring drought alone is not sufficient for the drought preparedness for the response. In order to ensure the mitigation of the impacts of the drought, the ability to predict the drought is necessary. Drought due to the unusually low rainfall amount can be directly significant to the vegetation and normalized differential vegetation index (NDVI) is a prominent indicator to detect drought. The trained model to predict NDVI for the future by using a long short-term memory (LSTM) algorithm will assist to forecast the potential drought for the upcoming period which can help the decision makers for the drought response.**

Keywords— drought, NDVI, LSTM

I. Introduction

A prolonged and severe deficiency in the availability of water phenomenon can be called drought and it can lead many challenges in human health, economy, agriculture and environment in the world today. Weather patterns are changing because of climate change and extreme and abnormal weather events including drought are happening more frequently.

As drought is a silent type in nature and compared to other disasters like floods and storms, it is more difficult to deliver the early warning. Drought is difficult to identify when it will start and end and only accessing and monitoring the impacts of drought can be mostly performed. Drought can be considered as the combination of meteorological, hydrological, and societal factors.

Not only in the agricultural sector but also in the livestock sectors can be threatened by drought and it can affect the food security of the country. Myanmar is a country where prone to the weather-related disasters because of its diverse agro-climatic conditions [1]. Myanmar faced drought situations in its history. The fresh event happened in 2019 and it caused in the early monsoon season in Magway Region located in the arid zone of Myanmar.

Agricultural drought is mainly related to various conditions like precipitation shortage, evapotranspiration and agricultural impact. Remotely sensed data from earth observation satellites like Landsat 8 can be utilized to detect vegetation conditions. The Normalized Differential Vegetation Index (NDVI) is an index of drought by detecting the healthy condition of the vegetation derived from the remote sensing data, particularly satellite imagery and broadly functional to express drought location and desertification.

Groundwater levels can also be simulated via short-term memory (LSTM) to predict the severity of drought using data-driven models [2]. Issuing early warning for agricultural drought based on NDVI is required to know the NDVI values in the upcoming period. Sequence prediction for NDVI values is suitable with LSTM because its ability to fetch long-term dependencies in time series data [3]. Having four gates in LSTM which can take the previous state as input to the present state, LSTM is relevant to use for predicting data in time sequence. [4].

Predicting drought is important for predicting weather and climate to prevent harmful effects on living organisms. The drought hazard conditions that happened in the past in the study area were identified by using Landsat 8 NDVI in this study. The historical NDVI values were extracted to feed in the LSTM neutral network for predicting the NDVI values for the drought in the future.

The necessary time-series NDVI values data to perform analysis of historical droughts were taken from the Landsat 8 satellite imageries by using the Google Earth Engine (GEE) platform. The aim of this study is to provide automatic drought detection for future. The rest of the paper is formed as follow.

Study region and data pre-processing to apply in this research will be described in section 2. The drought prediction algorithm intended to use will be discussed in section 3. Section 4 will explain the results of this research. Conclusions of the research will be provided in section 5.

II. Study Area and Data Pre-Processing with the Available Data

A. Study Region

The central Myanmar was chosen to perform this research because it is located in an arid zone with at least 40 degree Celsius average temperature in the summer. It covers an area of more than 54,000 km, there are 54 townships from 3 regions: namely Magway, Mandalay and lower part of Sagaing, according to the Dry Zone Greening Department (Myanmar's Natural Disasters, 2009) as shown in Figure 1. The study region received less rainfall than other parts of the country because it is located under the Rakhine mountainous regions which can weaken the southwest monsoon and may reduce rainfall. The temperature in the study area is very high, especially in the summer of Mar, Apr and May often reaching over 40 degrees Celsius. Land use and land cover changes in the region have resulted in insufficient rainfall and high temperatures resulting in dryness, and communities in the region are more vulnerable to droughts that threaten food security of the country.

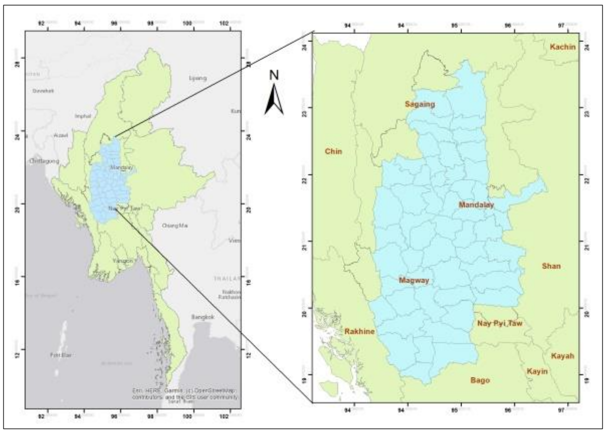

Figure 1. Study Area

B. *Data pre-processing with the available data*

Drought is characterized by three characteristics namely duration, spatial coverage and severity and it can be measured by using drought indices [5]. The Vegetation Health Index (VHI) is based on the Normalized Difference Vegetation Index (NDVI), which is a combination of products extracted from vegetation signals [6]. Satellite-based remote sensing has been widely utilized in the past several years from national level to global scales to monitor and observe many environmental events, including drought. Drought has been observed with different types of satellite sensors. Among the different satellites, Landsat series of satellites is famous because of its continuous data availability since 1972.

The plant can be released different colors of visible and near-infrared sunlight and these can be observed from the sensors equipped in the Landsat satellites to determine the green density of the vegetation on the land. The pigment in plant leaves called Chlorophyll can strongly absorb visible light (from 0.4 to 0.7 µm) to have photosynthesis. Almost all satellite vegetation indices use this difference formula to calculate plant growth density on Earth. The calculation for the NDVI is near-infrared radiation minus visible radiation divided by near-infrared radiation plus visible radiation. The mathematical formula is as below:

$$NDVI = (NIR - VIS)/ (NIR + VIS)$$

The calculation result can be ranged from minus one (-1) to plus one (+1) and close to plus one means the highest density of vegetation. The NDVI values of the observed area are the closer to minus one, the more potential to have the agricultural drought in that area. The historical NDVI values for the study region can be taken out from the Landsat 8 datasets by using the Google Earth Engine (GEE) which is a web-based platform.

The GEE online-platform is useful, for the countries that lack data from the past and have limited data processing facilities in high-performance to monitor silent and long-term behaviour disasters like drought. The GEE online-platform is also effective to classify multiple satellite imageries at the country level and with the possible to use the online-platform for higher scales sensors such as Landsat-8.

Moreover, the GEE can enable its users not only for Earth observation data acquisition, processing, analysis or visualization in speedily but also for monitoring or assessing any kind of natural hazards like flood and drought happened in a specific place in any time in the past or the present.

In this research, the data acquisition for NDVI values for monthly for each township in the study region in Myanmar was conducted on the GEE online-platform for the period 2015-2019. Long short-term algorithm (LSTM) is famous for time-series data modelling was applied in this study to predict the NDVI values for future to forecast the agricultural drought by using the past NDVI values in the past as the training dataset.

The flowchart for the study is as shown in Figure 2.

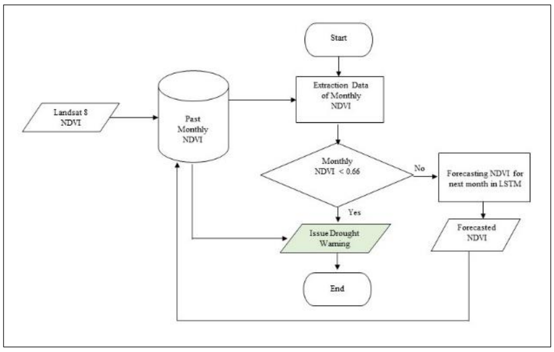

Figure 2. Flow chart

III. DROUGHT PREDICTION ALGORITHM

After obtaining the historical monthly NDVI values of the study area from the satellite imageries, data splitting was performed in the trained dataset, validation dataset, and test dataset. The trained dataset can be apply for training the LSTM model, the validation dataset can be applied for hyperparameter tuning, and the test set is intended to validate the NDVI forecasting the performance of model. LSTM is useful for predicting future values by keeping the data from the previous time series.

Once the predicted NDVI value from the LSTM model is obtained, that NDVI value will be checked by referencing the threshold value of greenness (> 0.66) in healthy vegetation. Then this predicted NDVI value is stored in the NDVI database to utilize as the input for next month prediction.

The algorithm to predict the NDVI as follow:

Algorithm:

Initialize monthly NDVI values,

while until end do

If NDVI value < 0.66, record as Drought

Otherwise Normal, then keep in NDVI and use that NDVI to predict for next month

end while

IV. RESULT AND DISCUSSION

Annual NDVI values for the research area for five years were examined, and the NDVI values in the study area was found out ranging from 0.33 to 0.66 below that of moderate healthy vegetation. During the 2015-2019 study period, particularly during 2015, which occurred as an El Niño event and vegetation conditions in unhealthy stage were occurred in 54 towns, 75% of the study area. The NDVI status of 25 out of 41 townships was unhealthy throughout the research period as shown in the bar chart of Figure 3.

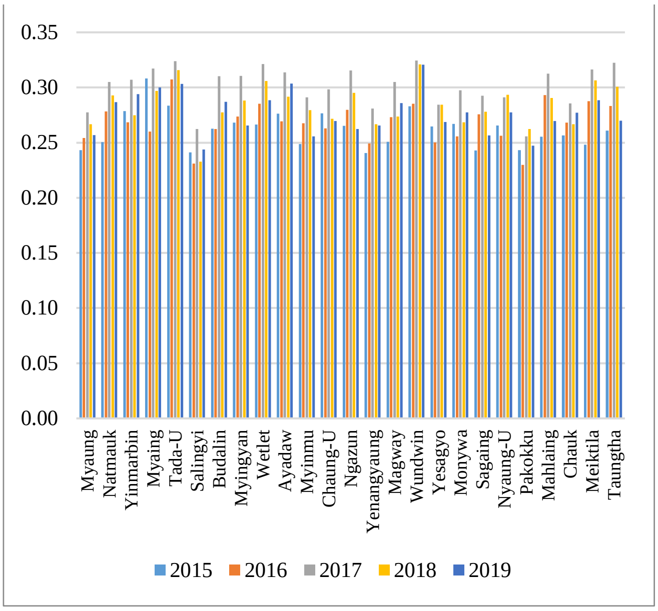

Figure 3. Chart of Towns with under NDVI value for vegetation in unhealthy stage

The research aims to apply a long short-term memory neural network that can study long-term data from the previous NDVI data obtained from Landsat 8 for predicting NDVI and future months of drought. Three years data from 2015 to 2017 were used as training dataset and 2018 and 2019 as validation dataset respectively.

The values estimated for NDVI by LSTM are normally satisfactory according to the weather condition trend in Myanmar. Estimated NDVI values for Oct, Nov and Dec are good and the NDVI is poor due to the impact of the hot season on the condition of the vegetation, but the NDVI values in other months are also practical.

V. CONCLUSION

The results from this study shown that the arid area of Myanmar is experiencing drought during the study period. Data derived from satellite observations such as Landsat 8 NDVI have contributed greatly to the spatial and temporal coverage for agricultural drought detection.

As NDVI is an indicator to detect drought condition, the trained model to predict NDVI for the future by using the LSTM algorithm can assist to forecast the potential drought for the future and it can help the stakeholders for the drought preparedness.

The potential drought for future can be predicted when the NDVI value is below 0.66, which is the reference value for unhealthy vegetation. It is found that the more trained dataset used in the LSTM, the more accurate for the predicted data.

ACKNOWLEDGMENT

First of all, I would like to express my special thanks to the University of Computer Studies (Yangon) and my Supervisor for her devotion to my research work and guidance.

REFERENCES

[1] Oo, A.T.; Van Huylenbroeck, G.; Speelman, S. Measuring the Economic Impact of Climate Change on Crop Production in the Dry Zone of Myanmar: A Ricardian Approach. Climate 2020, 8, 9.

[2] Lim, J.D.; Yang, J.S. Possibility analysis of future droughts using long short term memory and standardized groundwater level index. J. Korea Water Resour. Assoc. 2020, 53, 131–140.

[3] I. Park, H. S. Kim, J. Lee, J. H. Kim, C. H. Song and H. K. Kim, "Temperature prediction using the missing data refinement model based on a long short-term memory neural network", Atmosphere 2019, 10, 718.

[4] S. Poornima and M. Pushpalatha, "Drought prediction based on SPI and SPEI with varying timescales using LSTM recurrent neural network" Soft Computing 23, 8399–8412 (2019).

[5] Wilhite, D.A.; Michael, H.G. Understanding: The drought phenomenon: The role of definitions. Water Int. 1985, 10, 111–120. [Google Scholar]

[6] Tucker, C.J.; Vanpraet, C.; Boerwinkel, E.; Gaston, E.A. Satellite remote sensing of total dry matter production in the Senegalese Sahel. Remote Sens. Environ. 1983, 13, 461–474. [Google Scholar]