iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: XII Month of publication: December 2024

DOI: https://doi.org/10.22214/ijraset.2024.66196

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XII Dec 2024- Available at www.ijraset.com

An Overview of the Methodologies in Drought Forecasting

Angel R [1], Kokila Ramesh[2], Anita Chaturvedi[3]
[1]Research Scholar, Jain (Deemed-to-be-University), Bengaluru- 560078, India
[2]Associate Professor, Faculty of Engineering and Technology, Jain (Deemed-to-be University), Bengaluru 560078, India
[3] Professor, Faculty of Engineering and Technology, Jain (Deemed-to-be University), Bengaluru- 560078, India

Abstract: Drought has resulted in the massive destruction of lives, livestock and economy, locally and globally. Under the current circumstances, the dry weather conditions are prevalent leading to water stress resulting in imbalance of the ecosystem. The persistent dry weather eventually results in drought conditions, which indirectly affects the economy of the country. As a result, it has become imperative to assess and mitigate the effects of the droughts. With the advancements in the machine learning techniques, there has been a substantial improvement in the level of accuracy of the prediction of drought events. The motivation of this paper is to review those machine learning techniques such as Extreme Gradient Boost, ARIMA, Long Short-Term Memory (LSTM), Random Forest used in forecasting drought events. Considering the extensively used drought indices, that is, the Standard Precipitation Index (SPI) and Standardized Precipitation Evaporation Index (SPEI) together with the various meteorological parameters, the influence of the climate variables on drought forecasts are given more importance in the present study. This study also includes review on single and hybrid models used for forecasting droughts. Further the gaps in the reviews have been identified by critically observing the studies in the research articles which will enhance the efficiency of the predictive models.
Keywords: Drought indices, Review, Analysis, Modelling, Forecasting, Droughts, SPEI, Machine Learning Models, LSTM, ARIMA.

I. INTRODUCTION

Droughts are the most expensive natural event. With the continuous change in the climate, the global average surface temperature is on the rise. “There is a 66% chance that the annual average temperature of the near surface global temperature is 1.5^0 C above the pre-industrial level for at least one year between 2023 and 2027” [1]. This temperature increase leads to dryness of the weather, consequently increases the evapotranspiration and this results in the water stress. The incidence of drought events have increased in the past decades. The World Health Organization reports that around 55 million people have been affected by drought globally each year [2]. It affects the livelihood and livestock. Water shortages have impacted about 40% of the global population. There is a possibility that by 2030, 700 million people will be displaced due to drought [2].

A. Definition of Drought

According to Whilhite and Glantz [3], “there cannot be a universal definition of drought” [3]. The absence of a single definition for drought has handicapped the scientists to identify the

onset of drought. There are various definitions accepted by the different institutions. National Integrated Drought Information Systems (NIDIS), has defined drought as “A deficiency of precipitation over an extended period of time, resulting in a water shortage” [4]. The consequence of droughts, causes an imbalance in the demand and the supply of water, affecting water needs of the ecosystem [25]. Hence quantification of drought becomes pivotal to manage droughts.

B. Classification of Drought

Mishra and Singh [5] have classified different types of Droughts.

Meteorological drought- It occurs due to the effect of the mereological variables like precipitation, rainfall, temperature, evapotranspiration, cloud cover etc. It occurs when the dry weather pattern exists.

Agricultural drought – This occurs when there is insufficient supply of water to the crops, due to lack of water in the surface and the subsurface or ground level[5]. This results in the reduction in soil moisture, where the water supply does not meet the demand of the water requirement of the crops. The soil moisture during this period is very less or nil.

Hydrological drought – This is the result of decrease of water levels in the natural water resources such as rivers, ponds, lakes etc. and in the other storage tanks and the irrigation canals, dams, reservoirs, groundwater, etc. [5].
Socio -economic Drought- The non-availability of water for an extended period of time will lead to malnutrition in the population, lack of fodder for cattle which results in the death of the same, mass migration, and other economic activities such as tourism, hospitality industry will be affected.

C. Indicators & Indices

"The drought indicators are the variables which describe the drought conditions. The common indicators assumed are precipitation, temperature, soil moisture, ground and reservoir water, streamflow and snowpack" [6].
Drought Indices help to convert the indicator data with respect to a common scale, which can be used to compare the droughts across different geographical locations and across different time scales. There are more than 150 drought indices listed based on the drought indicators in the hand book of indicators and indices [6]. One of the earliest indices that was developed is the - "Palmer Drought Severity Index (PDSI) [7], China Z Index (CZI) [8], Rainfall Anomaly Index (RAI) [9], Rainfall deciles [10], Crop Moisture Index (CMI) [11], Aridity Anomaly Index (AAI) [12], Standardised Precipitation Index (SPI) [13], Standardised precipitation Evapotranspiration Index (SPEI)" and many more. Among all the indices, the World Meteorological Organisation (WMO) has recommended SPI. The SPI index was developed by Mackee et.al.[15], uses only precipitation as the input variable. It is a simple index compared to the other complex drought indices. The droughts can be better monitored and compared using SPEI rather than SPI, as temperature is a fundamental variable in drought occurrence together with precipitation [16].

D. Modelling of Droughts

It is observed in the existing literature that different modelling techniques are available to forecast the droughts. The models can be described as dynamical or data driven in nature. The dynamical models are based on physical laws involving hydrological cycle-land, ocean and atmosphere cycle. If the initial and the boundary conditions are accurately considered, then the prediction is bound to be accurate. Since the natural cycles are not yet completely understood, converting them into mathematical equations or differential equations is still a challenge. While the data driven models are constructed on the past data and its interaction with the other atmospheric variables. These interactions are based on the empirical relationships between the variables and are derived based on the measures of the relation. It is less time consuming and the ease to handle the complex data has made it more popular in predicting the weather data. [17] classified the drought modelling techniques and are as follows. Stochastic models -It is the time series model which includes Auto Regressive Integrated Moving Average (ARIMA) models and Seasonal Auto Regressive Integrated Moving Average (SARIMA) models. Probabilistic modelling that includes the Copula and Markov models, The Artificial Neural Network (ANN) models have performed extremely well when compared with the other models. Hybrid models- which is the combination of different models- such as copulas with ANN, ARIMA with ANN. so that it can be tailored according to the needs. So that the efficiency of each of the models can be used. This contributes to developing a model with a good performance.

II. REVIEW OF DROUGHT FORECASTING

In a study of [18], the training data used were from 1901 to 2010, the testing data from 2011 to 2018 were obtained from the Climate Research Unit for New South Wales, Australia. The indicator variables used were precipitation, mean temperature, vapour pressure, cloud cover and potential evapotranspiration [18]. The droughts were predicted using Long Short-Term Memory (LSTM) at two different time scales SPEI1, SPEI3. Receiver Operating Characteristic-based Area Under the Curve (ROC-AUC) approach was used to identify the drought classes. Here, the hydrometeorological variables alone were considered, and did not include the climate variables. The deep learning method was able to effectively predict the results even without the climate variables. The indicators such as vapour pressure and cloud cover are observed to be influencing factors of drought, like the temperature.
The drought forecasts [19] using SPEI and SPI for varying timescales was analysed. The data sets from 1959 to 2013 were used to train the data using LSTM and the year 2014 was to test the data under the used modelling technique. The input variables considered were the precipitation, minimum temperature and maximum temperature, windspeed, sunshine, humidity and evapotranspiration [19]. The correlation analysis was performed to identify the indicators affecting droughts index SPI and SPEI. The correlation of the indicators - humidity and temperature were considered with the indices. While the correlation existed between humidity and SPEI and not with SPI. For the sake of uniformity, the indicators humidity and temperature were considered. A combination of 4 models were developed- SPI with temperature, SPI with humidity, SPEI with temperature and SPEI with humidity for varying timescales-1,6,12, using ARIMA.

The performance of the ARIMA model for 1month-timescale using SPI and SPEI was higher, while for 6- and 12-month time scales were not very convincing. The accuracy of LSTM models for 1-month timescale is 99% for SPI AND SPEI. LSTM outperformed ARIMA model in the case of 6, 12-month timescale. The multivariate approach using LSTM has shown better results compared to the univariate method of ARIMA. In another study of [20], comparison between machine learning algorithms- XGB - Extreme Gradient Boost and RF- Random Forest and the Deep Learning models namely, LSTM - Long Short-Term Memory and CNN-Convolutional Neural Network. Seven scenarios with different choices of the input variables- “precipitation(P), Temperature average(T), temperature minimum, Temperature maximum, wind(W), relative humidity (RH), sunshine hours (SH) and solar radiation (SR) were analyzed for two different timescales SPEI 3 and SPEI-6 [20]. The study area was Tibetan plateau, China. The data from 1980 to 2019 were extracted. The input data was normalised for better understanding of the structure. Agnew’s approach was used to determine the drought classes. The various characteristics of drought, such as drought area, drought severity and drought duration were quantified. Principal component Analysis was used to locate drought patterns. The performance of machine learning models- XGB and RF were considered to be better compared to the deep learning models CNN and LSTM. It has been observed that there existed a significant correlation between wind speed and relative humidity.

The case study [21] was conducted in Jiangsu province in China, with three different methods- ARIMA, PROPHET- (is a Python based open-source software - effective in the case of missing data, with outliers and reckoning patterns) and LSTM drought prediction. Remote sensing data such as NDVI, LST- Land Surface Temperature, Climate data such as SPEI, Evapotranspiration-ETP, humidity, precipitation, wind speed and pressure and Biophysical data, soil moisture were considered [21]. Droughts were forecasted using SPEI values for the three models. Maps were created. For LSTM, the efficiency using Root Mean Square Error (RMSE) metric was 0.001 and concluded that LSTM was the best model as compared to ARIMA and PROPHET using the metrics, R2, RMSE, Mean Absolute Error (MAE).

Due to the complex nature of the weather data, the traditional methods [22] of forecasting the drought, is more suitable for the linear trend, and fails to cope with the non-linear trend existing in the data. With the development of the Artificial neural network, the drawbacks of the traditional methods have been overcome to a certain extent. This paper utilizes 6 models forecasting the drought using datasets from 613 stations from 1980 to 2019 with the multivariate drought index, SPEI. ARIMA was used to predict SPEI 24 and was more accurate compared to SPEI1. The precipitation and temperature data sets from the stations were used. The deep learning model considered was LSTM and Support Vector Machine (SVR) for various timescales. Support vector Regression (SVR) model was observed to be better suited for the prediction of the long-term drought prediction. Hybrid models with different combinations of standalone models such as ARIMA-LSTM, ARIMA-SVR and LS-SVR were developed [22]. The prediction using the ARIMA-LSTM model was far more efficient compared to the individual models and other hybrid models. In the ARIMA-LSTM model, the linear patterns were analysed by the ARIMA model. The residue between predicted and actual observations are then fed into the LSTM model. The nonlinear features were extracted and forecasted. Finally, the predicted linear feature and the non-linear features were integrated to get the complete prediction of the hybrid models. The aim of this paper is to improve the accuracy of short-term prediction. Of the considered models, the ARIMA-LSTM, performs better with long lead time.

Lagged climate variables and stacked LSTM [23] were used to forecast the long lead time drought The hydro-meteorological variables - precipitation, potential evapotranspiration, temperature, cloud cover along with “sea surface temperature (SST) indices such as Pacific Decadal oscillation, Southern oscillation Indices, Indian ocean Dipole Indices, Nino Indices and southern Annular mode” were used as the input variable, along with SPEI [23]. The lag periods were determined for SPEI and the variables, using Cross correlation with various climate indices. The optimum lag period obtained was as large as 8 months for Pacific Decadal Oscillation and zero for Indian Dipole Oscillation. The lagged variables used for prediction considerably, improves the forecasting for long lead time. It was observed that cloud cover and precipitation resulted in high correlation. Deep learning algorithm - Stacked LSTM, was used to forecast. The LSTM models perform better than the traditional models.

The meteorological drought was predicted using the Wavelet - ARIMA- LSTM model using China Z-score Index (CZI) [24]. The Grey prediction model is used for the drought characterization of the forecasted values. The precipitation data of 51 years is used for training and testing of data. This data is first normalized. This method involved the decomposition and reconstruction of data. First, the data is classified into high frequency, low frequency components using wavelet decomposition [24]. ARIMA estimates the low frequency coefficients, while LSTM estimates the high frequency coefficients which are non-linear. The predicted coefficients are then used to reconstruct the data. Comparison was made between ARIMA, LSTM, Wavelet-ARIM-LSTM. Then the forecasted values of ARIMA and LSTM are utilized to reconstruct the data. The CZI was used to characterize the drought. Finally, the Discrete Grey model was used to categorise the drought classes. Grey models work well when the sample size is small. The model predicted better in the humid regions.

III. DISCUSSION

Droughts are one of the least understood natural calamities of nature resulting in devastating effects on the ecosystem. The frequency of droughts in the current times is alarming. Hence there is an urgent need for a reliable forecasting model. The efficiency of the forecasting models with long lead time is the pressing priority. The fundamental requirements of a forecasting model are historical data, a suitable index and a forecasting model. The historical data can be of gridded data or the remote sensing data of the required parameters. Secondly there are several drought indices available and the recommended meteorological drought indices by WMO are SPI - which considers only precipitation as the parameter and SPEI - considers temperature and precipitation as the parameter. The performance of SPI and SPEI have been significantly higher, as compared to the other drought indices. Nevertheless, the prediction of droughts can be further enhanced by the use of climate indices such as Indian Ocean Dipole, Southern Oscillation Index, ENSO, etc. The meteorological index along with the climate indices have resulted in a better accuracy. Another element in the prediction of the drought, is the type of models used. Among the data driven models, the hybrid models such as ARIMA-LSTM, ARIMA-SVR, Wavelet-ARIMA-LSTM etc, have had higher performance accuracy than the single models such as LSTM, ARIMA, XGB, Random Forest etc. Thus, the essential components of the modelling should be chosen carefully for the accurate prediction of droughts.

The RMSE for SPI -1 with humidity and that of SPEI-1 with humidity is 0.011 and 0.2 respectively, which gives a good accuracy for the region of Hyderabad [19]. It may be observed that in both the cases the RMSE is low and it may not be the only variable included for predicting the drought events, but this may be one of the steps. This accuracy was achieved using the deep learning LSTM which performed better than the ARIMA model. [18] used SPEI with LSTM to predict the droughts for New South Wales of Australia and RMSE was found to be 0.018 for SPEI-1 and 0.027 for SPEI-3. Here the only indices used are SPEI-1 and SPEI-3 and are not connected to other atmospheric indices such as Southern Oscillation index (SOI), Indian Dipole Index, snow cover and so on. Since LSTM is used to predict drought events, it may perform during the training period, but the actual challenge might arise during forecasting. The machine learning models, namely Extreme Gradient Boost, Random Forest, Convolution neural Network and the LSTM models were used in the prediction in Tibet plateau [20]. Nash-Sutcliffe model efficiency for SPEI -3 was more than 0.75 for Extreme Gradient Boost models and RF models. However, the inclusion of atmospheric variables might improve the efficiency. [23] has included the atmospheric parameters by considering Southern Oscillation Index, Indian Ocean Dipole, along with the meteorological parameters considering the stacked LSTM model. It was able to consider the nonlinear relationship between the variables along with ENSO, but the forecasting efficiency needs to be estimated as per the model, which is not considered in this article. [24] used the hybrid model of WAVELET-ARIMA-LSTM over north east China with the CZI index. The correlation coefficient was found to be as close as 0.880 between the model and the precipitation estimates. To effectively model, predict and forecast the extreme events in any country, one has to concentrate on the efficiency of the model through the training period. If in the training period efficiency crosses 95%, then there is a possibility of predicting and forecasting with better accuracy and can be relied upon. With the availability of various machine learning models, and considering all possible correlated variables of the drought, drought forecast can be improvised. This will provide the researchers to proceed in a correct path to come out with better accuracy models to forecast drought events with better confidence levels.

IV. CONCLUSION

Drought has a massive impact on the ecosystem and has inevitably compromised human existence. Considering the climate changes, the frequency of occurrence of droughts is inclined to increase in the future. This paper reviews the various drought indices, atmospheric variables and the models with different methodology used to predict droughts. It highlights the different drought indices used in the models. The effectiveness of various parameters such as the meteorological parameters and inclusion of the climatic parameters in the prediction are examined through the articles reviewed. The comparison of different models used such as the individual models and the hybrid models are included. This review provides sufficient information to the researchers for using a proper methodology to overcome the gaps in the existing models.

REFERENCES

[1] World Meteorological Organisation., (2023, March 17), Global Temperatures set to reach new records in next five years. https://wmo.int/news/media-centre/global-temperatures-set-reach-new-records-next-five-years.
[2] World Health Organisation, (n.d.). Drought, https://www.who.int/health-topics/drought/#tab=tab_1
[3] Wilhite, Donald A. and Glantz, Michael H., "Understanding the Drought Phenomenon: The Role of Definitions" (1985). Drought Mitigation Center Faculty Publications. 20. http://digitalcommons.unl.edu/droughtfacpub/20
[4] National integrated Drought Information System, (n.d.), Drought Basics, https://www.drought.gov

[5] Mishra, A. K., & Singh, V. P. (2011). Drought modeling – A review, Journal of Hydrology.

[6] Handbook of Drought Indicators and Indices (n.d.), pg. 3, https://www.droughtmanagement.info/literature/GWP_Handbook_of_Drought_Indicators_and_Indices_2016.pdf.

[7] Palmer, W.C., 1965: Meteorological Drought. Research Paper No. 45, US Weather Bureau, Washington, DC.

[8] Wu, H., M.J. Hayes, A. Weiss and Q. Hu, 2001: An evaluation of the Standardized Precipitation Index, the China-Z Index and the Statistical Z-score. International Journal of Climatology, 21: 745–758.

[9] van Rooy, M.P., 1965: A Rainfall Anomaly Index independent of time and space. Notos, 14: 43–48.

[10] Gibbs, W.J. and J.V. Maher, 1967: Rainfall Deciles as Drought Indicators. Bureau of Meteorology Bulletin No. 48, Melbourne, Australia.

[11] Palmer, W.C., 1968: Keeping track of crop moisture conditions, nationwide: the new Crop Moisture Index. Weatherwise, 21: 156–161. DOI: 10.1080/00431672.1968.9932814

[12] World Meteorological Organization, 2012: Chapter 6, Agrometeorological Forecasting, 6–15

[13] Guttman, N.B., 1998: Comparing the Palmer Drought Index and the Standardized Precipitation Index. Journal of the American Water Resources Association, 34: 113–121. DOI: 10.1111/j.1752-1688.1998.tb 05964.x.

[14] Vicente-Serrano, S.M., S. Begueria and J.I. Lopez-Moreno, 2010: A multi-scalar drought index sensitive to global warming: the Standardized Precipitation Evapotranspiration Index. Journal of Climate, 23: 1696–1718. DOI: 10.1175/2009JCLI2909.1.

[15] McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration of time scales. Eighth Conference on Applied Climatology, American Meteorological Society, Jan17-23, 1993, Anaheim CA, pp.179-186.

[16] Ashok K. Mishra, Vijay P. Singh, Drought modeling – A review, Journal of Hydrology, Volume 403, Issues 1–2, 2011, Pages 157-175, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2011.03.049.

[17] Ashok K. Mishra, Vijay P. Singh, A review of drought concepts, Journal of Hydrology, Volume 391, Issues 1–2, 2010, Pages 202-216, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2010.07.012.

[18] Dikshit A, Pradhan B, Huete A. An improved SPEI drought forecasting approach using the long short-term memory neural network. J Environ Manage. 2021 Apr 1;283:111979. doi: 10.1016/j.jenvman.2021.111979. Epub 2021 Jan 19. PMID:33482453.

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

[20] A. Mokhtar et al., "Estimation of SPEI Meteorological Drought Using Machine Learning Algorithms," in IEEE Access, vol. 9, pp. 65503-65523, 2021, doi: 10.1109/ACCESS.2021.3074305.

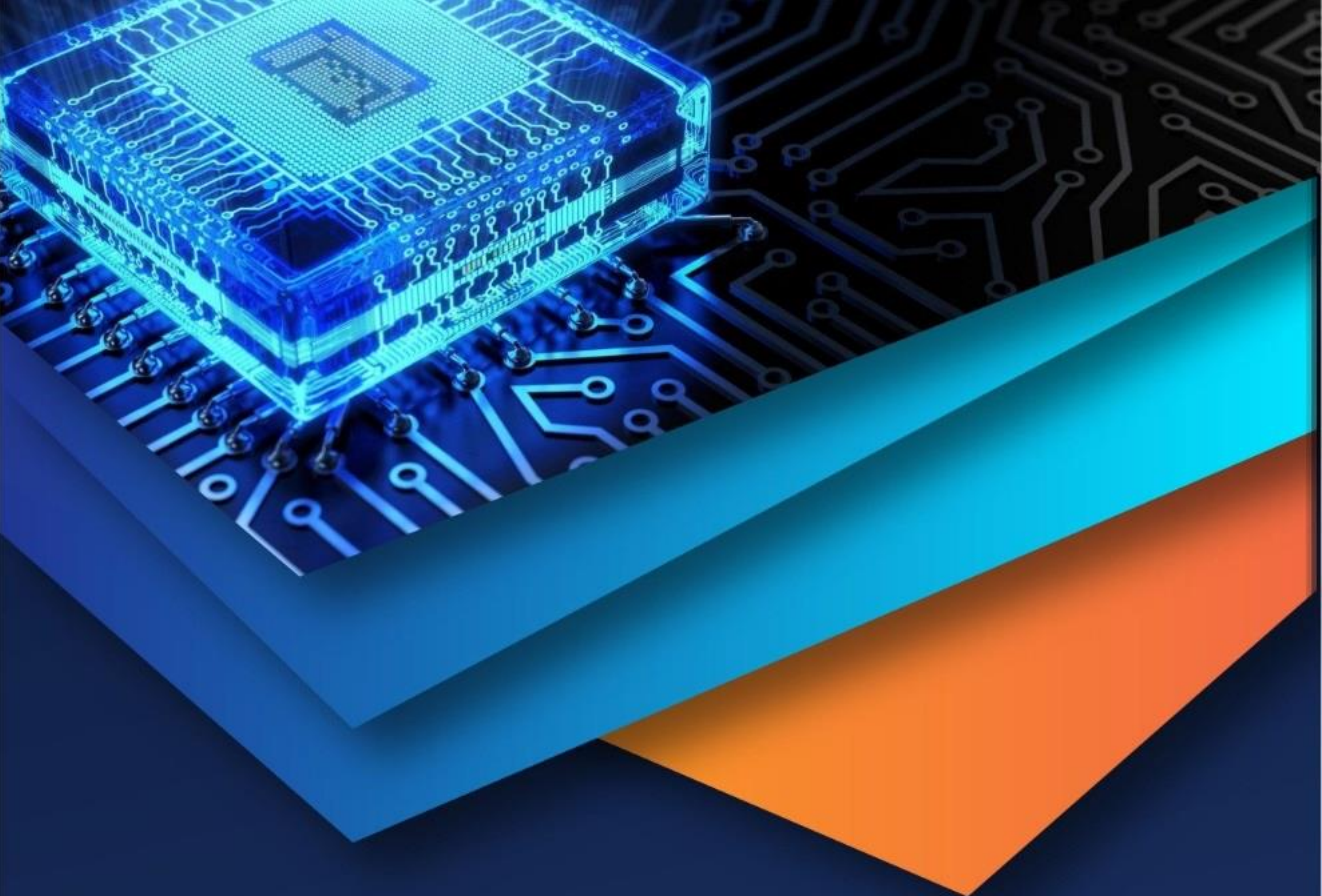
[21] H. Balti et al., "Big data based architecture for drought forecasting using LSTM, ARIMA, and Prophet: Case study of the Jiangsu Province, China," 2021 International Congress of Advanced Technology and Engineering (ICOTEN), Taiz, Yemen, 2021,pp. 1-8, doi: 10.1109/ICOTEN52080.2021.9493513.

[22] Xu D, Zhang Q, Ding Y, Zhang D. Application of a hybrid ARIMA-LSTM model based on the SPEI for drought forecasting. Environ Sci Pollut Res Int. 2022 Jan;29(3):4128-4144. doi: 10.1007/s11356-021-15325-z. Epub 2021 Aug 17. PMID: 34403057.

[23] Abhirup Dikshit, Biswajeet Pradhan, Abdullah M. Alamri, Long lead time drought forecasting using lagged climate variables and a stacked long short-term memory model, Science of The Total Environment, Volume 755, Part 2, 2021, 142638, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2020.142638.

[24] Wu, X., Zhou, J., Yu, H., Liu, D., Xie, K., Chen, Y., … Xing, F. (2021). The development of a hybrid wavelet-Arima-lstm model for precipitation amounts and drought analysis. Atmosphere, 12(1). https://doi.org/10.3390/ATMOS12010074

[25] Iban Ortuzar, Ana Serrano, Àngels Xabadia, Macroeconomic impacts of water allocation under droughts. Accounting for global supply chains in a multiregional context, Ecological Economics, Volume 211, 2023, 107904, ISSN 0921-8009, https://doi.org/10.1016/j.ecolecon.2023.107904.



10.22214/IJRASET

INDEX COPERNICUS

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

TOGETHER WE REACH THE GOAL

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY