DROUGHT AND PRECIPITATION MODEL FOR WATER RESOURCE MANAGEMENT ON THE ZAMBEZI RIVER

Shamiso Paskwavaviri

Zambezi River Authority

shamiso.paskwavaviri@zambezira.org

Mobile No: +263715793238/+260776591937

Shamiso Paskwavaviri holds a Bachelor of Engineering in Civil and Water Engineering, (2022) and is currently a graduate trainee at the Zambezi River Authority. She contributes to the Batoka Gorge Hydroelectric Scheme and Kariba Dam Rehabilitation Projects, with specific involvement in project management.

Lucy Mwape

Zambezi River Authority

lucy.mwape@zambezira.org

Mobile No: +260765609745

Lucy Mwape is a Civil Engineer with experience on major projects like the Kariba Dam and Batoka Gorge Hydro Scheme, focusing on planning, safety, and stakeholder coordination. She is passionate about sustainable infrastructure and aims to specialize in renewable energy.

CO - AUTHORS

Patrick Sipatela

Zambezi River Authority

patrick.sipatela@zambezira.org

Patrick Sipatela, holds a MEng and BEng from The University of Zambia, including project management certification from the University of Cape Town. He is a distinguished Engineer with over 13 years of experience and has excelled in roles such as Designs Manager, Technical Manager, Project Manager including provision of technical advisory as an independent consultant leading complex construction projects with high-level technical expertise and project management skills.

‘Renowned for driving organizational efficiency and ensuring compliance with Health, Safety, and Environmental regulations, Patrick adeptly manages cross-functional teams throughout the project lifecycle. He builds robust stakeholder relationships to achieve project goals and has extensive experience in construction project management, risk management, environmental compliance, and quality assurance. He is a Professional member of the Engineering Institution of Zambia and passionately leads the ambitious 2400MW Batoka Gorge Hydro-electric Scheme for the Zambezi River Authority.

Kwanga Miyanza

Zambezi River Authority

kwanga.miyanza@zambezira.org

Kwanga Miyanza is an Environmental Engineer with 15 years' experience in mining, manufacturing, and environmental consultancy. Currently working for ZRA as EHS Officer on the BGHES project.

Munyaradzi Munodawafa

Zambezi River Authority

munyaradzi.munodawafa@zambezira.org

Munyaradzi C. Munodawafa is an Electrical engineer with more than 40 years of experience in the Energy Sector, ranging from power engineering, fossil fuels, renewable energy, water resources and large dam management. Currently implementing the US$ 300 million Kariba Dam Rehabilitation Project and concluding the pre-development activities for the development of the 2400 MW Batoka Gorge Hydro-Electric Scheme. He possesses diverse experience in water resource engineering and environmental management. He is a Fellow of the Zimbabwe Institute of Engineers (ZIE), Corporate Member of the Engineering Institute of Zambia (EIZ); and a Contributor to the Engineers Without Borders programmes, jointly organized by the Zimbabwe Institute of Engineers and the Engineering Institute of Zambia.

1.0 ABSTRACT

The Zambezi River Basin, particularly the Kariba Lower Catchment (KLC), is increasingly vulnerable to climate-induced droughts and precipitation variability, exemplified by the severe 2023/2024 El Niño event that reduced Lake Kariba's usable storage to 13.8% in 2024, causing an over 85% decline in hydropower generation. To address these challenges, this study develops a localized drought prediction model for the KLC using Long Short-Term Memory (LSTM) neural networks, which leverage long-term meteorological and hydrological data from the Zambezi River Authority alongside statistical tools such as the Standardized Precipitation Index (SPI) and Gamma distribution fitting to enhance forecasting accuracy. The methodology includes data cleaning, stationarity testing, LSTM model training, and validation through goodness-of-fit tests and observed drought event comparisons, while exploring machine learning techniques to improve scalability. The resulting model supports climate-resilient hydropower planning, informs regional drought preparedness policies, and aligns with Zambia's Nationally Determined Contributions under the Paris Agreement, promoting sustainable water resource allocation and enhancing adaptive capacity in the basin. The study used 20 years of precipitation data from gauge stations in the Kariba Lower Catchment (KLC). The prediction model performed best at Gwayi and Ume River stations, likely due to fewer zero-precipitation values and better spatial coverage, which improved the Gamma distribution fit. The limited dataset may have affected model accuracy at other stations by not fully capturing long-term variability.

Keywords:

Standardised Precipitation Index; Drought monitoring; Water Resource Management; Gamma distribution; LSTM model

2.0 INTRODUCTION

The Zambezi River Basin, situated in southern Africa, currently offers vast water resources for social and economic development for the eight riparian countries that constitute the watershed. (Angola, Botswana, Malawi, Tanzania, Namibia, Zambia, Mozambique Zimbabwe). One of its key features is Lake Kariba, a major reservoir that plays a central role in regional energy production. As the primary water source for hydropower stations in both Zambia and Zimbabwe, the reliable allocation of water at Kariba is critical for energy security, economic stability, and cross-border cooperation.

However, the basin is increasingly threatened by the impacts of climate change, population growth, and economic development, which are intensifying pressure on water resources. Southern Africa is particularly susceptible to climate variability, often experiencing cycles of droughts and floods that severely affect vulnerable populations and stress existing infrastructure.

Drought monitoring and flood prediction are essential tools for proactive water management in the Kariba sub-basin. Accurate drought forecasting enables better water allocation to hydropower stations, supports emergency preparedness, informs operational decisions, and facilitates long-term planning for water use across competing sectors. Despite advancements in modelling approaches, challenges remain due to data scarcity, short historical records, and model calibration limitations—especially in ungauged or under-monitored catchments. Traditional hydrological and stochastic models often fail to capture the complexity of extreme events, particularly under nonstationary and changing climate conditions.

This study aims to develop and evaluate a Gamma distribution-based drought prediction model for the Lower Kariba Catchment. The model is based on historical hydrological records from 2000 to 2020, utilizing inflow measurements from four rain gauge stations. The Gamma distribution is well-suited for hydrological applications, particularly in representing the skewed nature of rainfall data. Its mathematical flexibility and simplicity make it appropriate for estimating the probability of drought and wet periods, especially in data-constrained environments such as this one.

3.0 LITERATURE REVIEW

3.1 SPI as a drought index.

The Standardized Precipitation Index (SPI) is the most used indicator worldwide for detecting and characterizing meteorological droughts. The SPI indicator, which was developed by McKee et al. (1993), and described in detail by Edwards and McKee (1997), measures precipitation anomalies at a given location, based on a comparison of observed total precipitation amounts for an accumulation period of interest (e.g. 1, 3, 12, 48 months), with the long-term historic rainfall record for that period. The historic record is fitted to a probability distribution (the “gamma” distribution), which is then transformed into a normal distribution such that the mean SPI value for that location and period is zero Precedents in Southern Africa or similar catchments.

3.2 Long Short-Term Memory (LSTM) Networks

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network (RNN) designed to capture long-term dependencies in time-series data. They are particularly effective in hydrology for modeling non-linear and sequential data such as rainfall and streamflow (Hochreiter and Schmidhuber, 1997). Due to their ability to retain historical patterns, LSTMs have outperformed traditional models in rainfall prediction tasks, making them suitable for forecasting applications in drought-prone regions.

3.3 Applications in Africa and Data-Sparse Environments

Husak et al. (2007) advanced the application of the Gamma distribution across the African continent, fitting it to interpolated monthly rainfall grids derived from the Collaborative Historical African Rainfall Model (CHARM). Their findings showed that the Gamma distribution was suitable for representing monthly rainfall in over 98% of tested locations across Africa, as validated using the Kolmogorov–Smirnov (KS) goodness-of-fit test. The Gamma distribution is also widely used in calculating the Standardized Precipitation Index (SPI), a leading drought indicator. As reported by McKee et al. (1993), the SPI relies on fitting a Gamma distribution to historical precipitation data to compute standardized deviations from the mean. This allows for easy classification of drought severity levels over different timescales.

4.0 STUDY AREA

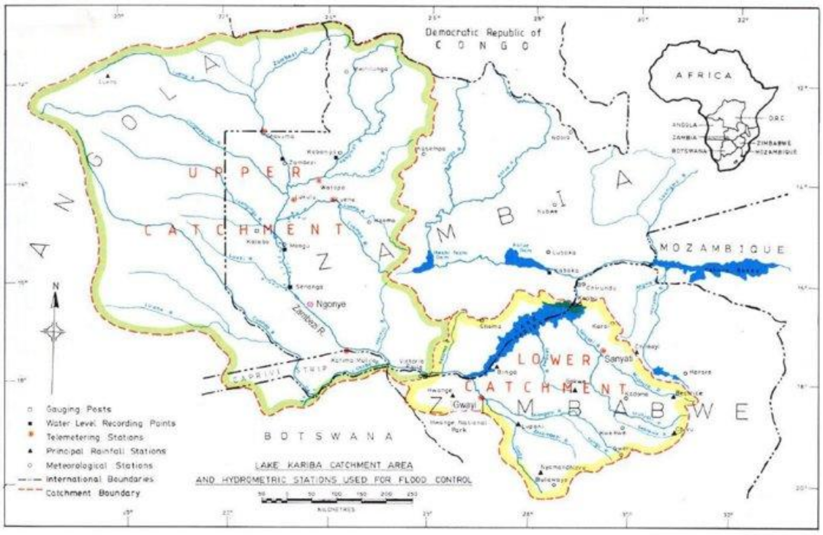


Figure 1: Upper and lower catchment of Zambezi River Basin

The lower catchment of the Kariba Dam shown in Figure 1 lies within the Zambezi River Basin, straddling parts of Zambia and Zimbabwe. Longitudes of 15°S and 21°S, and Latitudes of 25°E and 32° E in the Kariba Catchment. KLC lies between Victoria Falls and Lake Kariba. This region falls within a semi-arid climate zone, characterized by high rainfall variability and a distinct wet and dry season. Average annual rainfall is low and unevenly distributed, making the area particularly vulnerable to frequent droughts, some of which have significantly affected water availability. The lower Kariba catchment is of strategic importance due to its role in hydropower generation, irrigated agriculture, and as a vital ecological zone supporting diverse flora and fauna. Reliable rainfall modelling and drought forecasting in this region are critical for water resource planning and transboundary cooperation.

5.0 DATA AND METHODOLOGY

5.1 Data Sources

The database comprises of precipitation collected from more than 4 rain gauge stations in Kalomo, Ume River, Gwayi River and Livingstone. For more than half a century, the Kariba catchment has been gauged, and runoff records have been kept at various stations, as shown in Figure 2. The precipitation data was obtained from the Zambezi River Authority. The data obtained was pre-processed for quality and consistency by missing data filling and outliers' identifications.

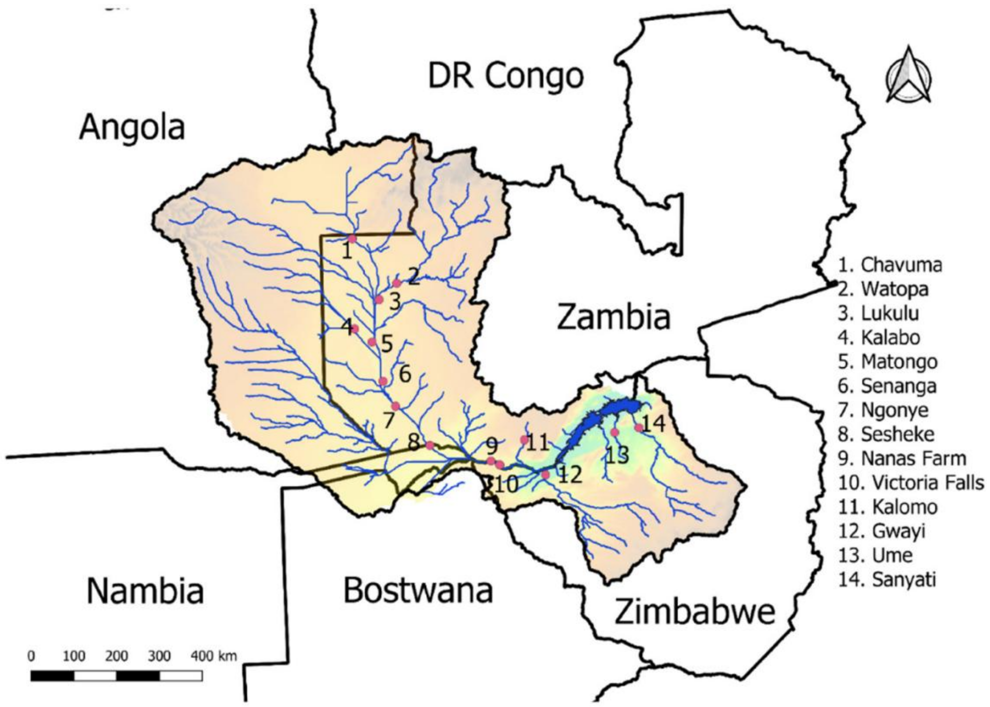


Figure 2 shows Location of gauging stations in the Kariba Catchment 10, 11, 12, 13

5.2 SPI Computation in Gamma Distribution

The Standardized Precipitation Index (SPI) was computed to quantify meteorological drought conditions using rainfall data fitted to a Gamma distribution. The Gamma distribution was selected due to its suitability for modelling positively skewed, non-negative precipitation data (Thom, 1958; Wilks, 2011). The cumulative distribution function (CDF) was then transformed into a standard normal distribution to derive SPI values (Guttman, 1999). A 3-month time scale was used to capture short-term drought patterns (World Meteorological Organization, 2012). The SPI values were interpreted based on widely accepted thresholds as shown in Table 1. The mean SPI trends were classified per year and month, enabling a standardized comparison of rainfall anomalies across time and location.

Table 1 Classification of drought based on SPI

SPI Values	Class
>2	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
−0.99 to 0.99	Near normal
−1.0 to −1.49	Moderately dry
−1.5 to −1.99	Severely dry
<−2	Extremely dry

5.3 LSTM Model Validation

The performance of the precipitation prediction model was evaluated using a Long Short-Term Memory (LSTM) neural network. A 3-month rolling time series was prepared for each of the four rain gauge stations. The dataset from each station, covering the period 2000 to 2024, was split into training and testing sets using an 80:20 ratio. The first 80% of the data was used to train the LSTM model, while the remaining 20% was used for testing and validation. Model validation was conducted by comparing the LSTM-predicted rainfall values against the actual observed precipitation values for each station. The comparison focused on both temporal consistency and magnitude accuracy. To quantitatively assess the model's predictive performance, three standard statistical metrics were used:

- Root Mean Square Error (RMSE) – to measure the average magnitude of prediction errors, where.
- Mean Absolute Error (MAE) – to assess the average absolute deviation between predicted and observed values, where lower MAE indicates more accurate predictions.
- Coefficient of Determination (R^2) – to evaluate how well the model captured the variability in observed data.

6.0 RESULTS

6.1 SPI Drought Index

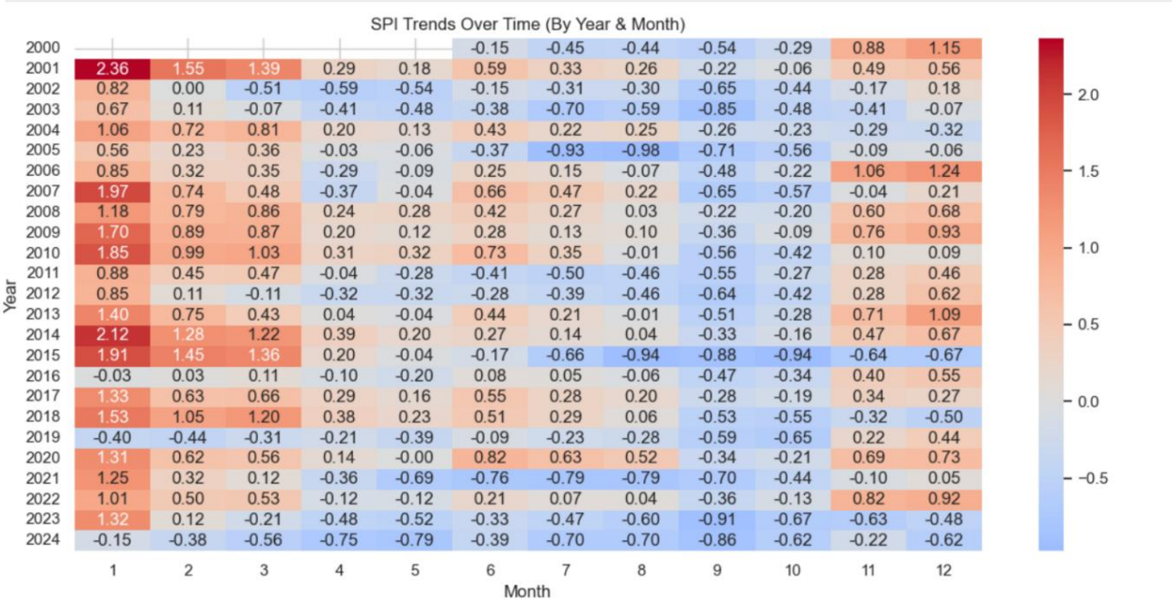
The outcome of the Kolmogorov–Smirnov (KS) goodness-of-fit test indicated that the inflow data follow a Gamma distribution at all stations. Consequently, SPI values were calculated for Livingstone, Gwayi River, Ume River, and Kalomo stations. These SPI values, computed across multiple years, are presented in Annex A.

The Ume and Gwayi River stations recorded no drought classification, with average SPI indices above -0.5. Livingstone experienced 62.9% of the time in the no drought category, while Kalomo recorded 71.5%. Overall, the SPI values across all stations, Livingstone, Gwayi River, Ume River, and Kalomo stations show increased inflow indicative of a wet season—between November to March. In contrast, the months of April to October generally fall within the drought category, as shown in Table 2.

Severe drought conditions were observed during the 2002/2003 season, with extreme SPI indices above 2 recorded at Livingstone, Ume, and Kalomo stations. In 2001/2002 Zambia experienced a severe drought. Livingstone, Gwayi River, Ume River, and Kalomo stations shows drought clusters in 2023/2024 with a drought index of less than 0 indicating drought classification. This can be reflected with the 2023/2024 Elnino effects. This significantly led to lower lake levels and reduced power generation in Zimbabwe and Zambia. Drought clusters are presented in 2015/2016 and 2019/2020 were drought was experienced in Zambia and Zimbabwe. The droughts affected agriculture and electricity production due to lower levels in lake Kariba because of the delayed rainfall season and poor rainfall.

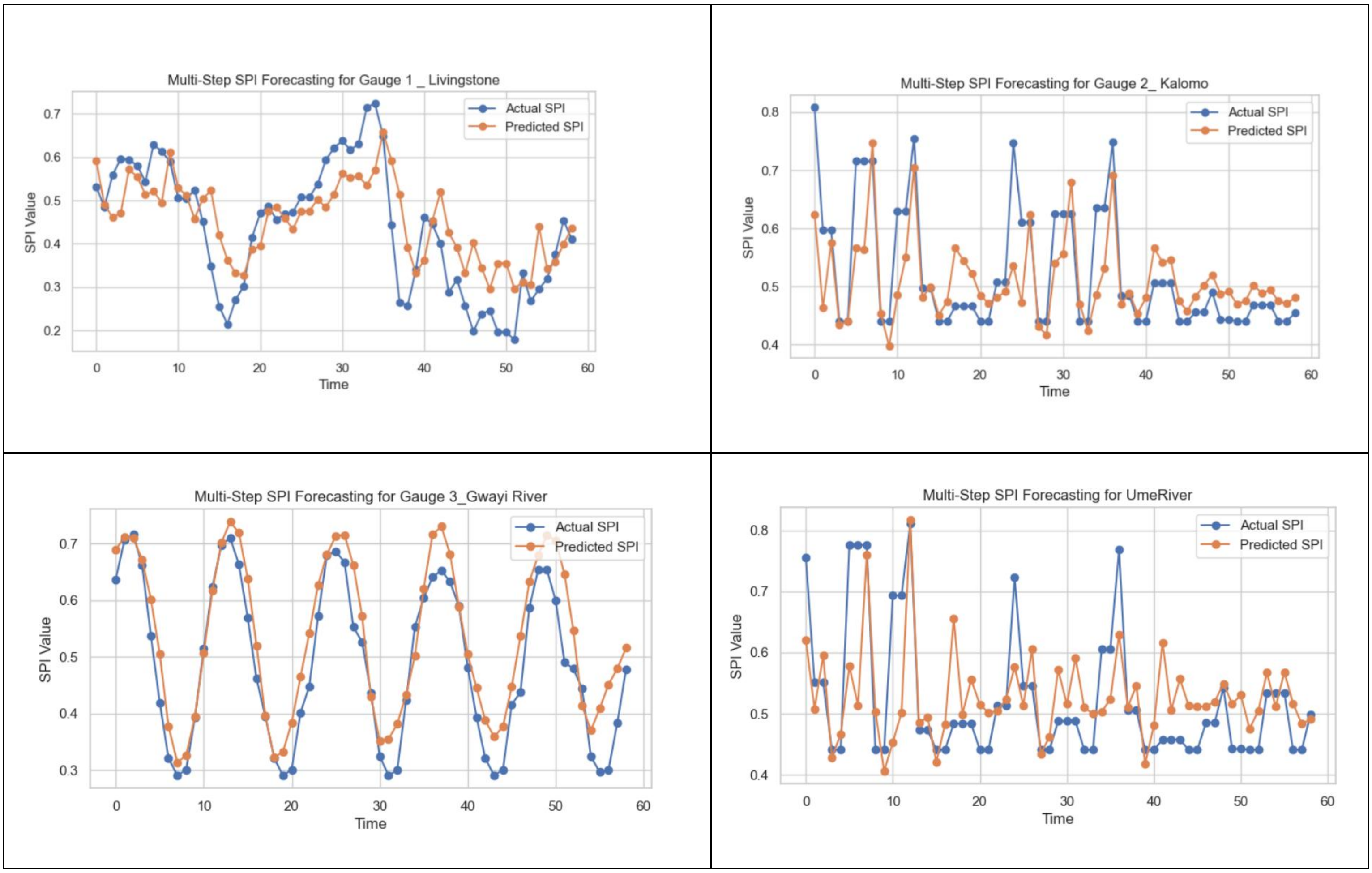
This drought monitoring is critical for informing water allocation decisions for hydropower generation at the Kariba North and South Bank Power Stations in Zimbabwe and Zambia.

Table 2 shows heat map for the SPI trends over time



Year	1	2	3	4	5	6	7	8	9	10	11	12
2000						-0.15	-0.45	-0.44	-0.54	-0.29	0.88	1.15
2001	2.36	1.55	1.39	0.29	0.18	0.59	0.33	0.26	-0.22	-0.06	0.49	0.56
2002	0.82	0.00	-0.51	-0.59	-0.54	-0.15	-0.31	-0.30	-0.65	-0.44	-0.17	0.18
2003	0.67	0.11	-0.07	-0.41	-0.48	-0.38	-0.70	-0.59	-0.85	-0.48	-0.41	-0.07
2004	1.06	0.72	0.81	0.20	0.13	0.43	0.22	0.25	-0.26	-0.23	-0.29	-0.32
2005	0.56	0.23	0.36	-0.03	-0.06	-0.37	-0.93	-0.98	-0.71	-0.56	-0.09	-0.06
2006	0.85	0.32	0.35	-0.29	-0.09	0.25	0.15	-0.07	-0.48	-0.22	1.06	1.24
2007	1.97	0.74	0.48	-0.37	-0.04	0.66	0.47	0.22	-0.65	-0.57	-0.04	0.21
2008	1.18	0.79	0.86	0.24	0.28	0.42	0.27	0.03	-0.22	-0.20	0.60	0.68
2009	1.70	0.89	0.87	0.20	0.12	0.28	0.13	0.10	-0.36	-0.09	0.76	0.93
2010	1.85	0.99	1.03	0.31	0.32	0.73	0.35	-0.01	-0.56	-0.42	0.10	0.09
2011	0.88	0.45	0.47	-0.04	-0.28	-0.41	-0.50	-0.46	-0.55	-0.27	0.28	0.46
2012	0.85	0.11	-0.11	-0.32	-0.32	-0.28	-0.39	-0.46	-0.64	-0.42	0.28	0.62
2013	1.40	0.75	0.43	0.04	-0.04	0.44	0.21	-0.01	-0.51	-0.28	0.71	1.09
2014	2.12	1.28	1.22	0.39	0.20	0.27	0.14	0.04	-0.33	-0.16	0.47	0.67
2015	1.91	1.45	1.36	0.20	-0.04	-0.17	-0.66	-0.94	-0.88	-0.94	-0.64	-0.67
2016	-0.03	0.03	0.11	-0.10	-0.20	0.08	0.05	-0.06	-0.47	-0.34	0.40	0.55
2017	1.33	0.63	0.66	0.29	0.16	0.55	0.28	0.20	-0.28	-0.19	0.34	0.27
2018	1.53	1.05	1.20	0.38	0.23	0.51	0.29	0.06	-0.53	-0.55	-0.32	-0.50
2019	-0.40	-0.44	-0.31	-0.21	-0.39	-0.09	-0.23	-0.28	-0.59	-0.65	0.22	0.44
2020	1.31	0.62	0.56	0.14	-0.00	0.82	0.63	0.52	-0.34	-0.21	0.69	0.73
2021	1.25	0.32	0.12	-0.36	-0.69	-0.76	-0.79	-0.79	-0.70	-0.44	-0.10	0.05
2022	1.01	0.50	0.53	-0.12	-0.12	0.21	0.07	0.04	-0.36	-0.13	0.82	0.92
2023	1.32	0.12	-0.21	-0.48	-0.52	-0.33	-0.47	-0.60	-0.91	-0.67	-0.63	-0.48
2024	-0.15	-0.38	-0.56	-0.75	-0.79	-0.39	-0.70	-0.70	-0.86	-0.62	-0.22	-0.62

Table 3 shows observed data vs predicted data for the gauging stations.



6.2 LSTM Performance Metrics

For each station, the predicted precipitation values were compared to the observed values over the test period. The graphs in Annex A illustrate time series plots of observed versus predicted precipitation.

The LSTM achieved the statistical metrics as follows.

1. Gwayi recorded the best model performance with the lowest MSE (0.0053) and RMSE (0.0725), and the highest R^2 value (0.7362), indicating a strong fit between predicted and observed values.
2. Livingstone also showed reasonable performance with an MSE of 0.0094, RMSE of 0.0969, and R^2 of 0.5707, suggesting moderate predictive accuracy. RMSE and MAE are low, indicating relatively small errors.
3. Ume and Kalomo performed less favourably. Ume had an MSE of 0.0076, RMSE of 0.0874, and R^2 of 0.3412, while Kalomo showed the weakest performance with an MSE of 0.0090, RMSE of 0.0948, and R^2 of just 0.1732, reflecting limited model reliability at these stations.

Table 4 shows the SPI forecast for Ume River

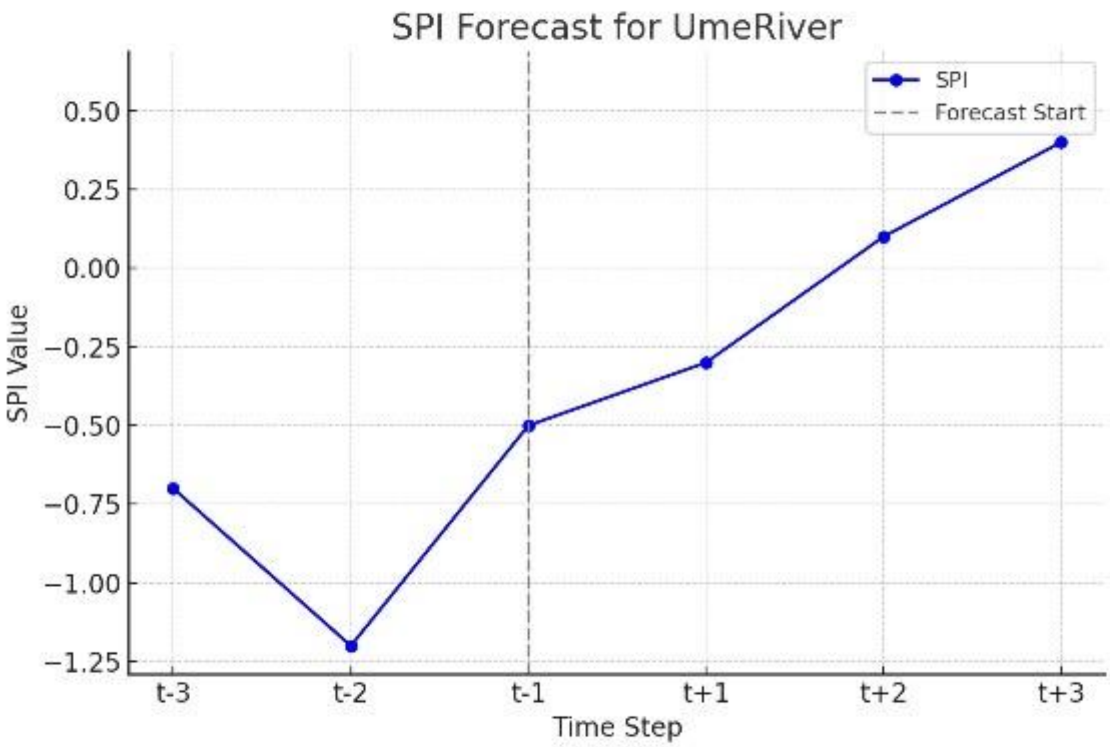


Table 4 presents the forecasted Standardized Precipitation Index (SPI) values for January, February, and March 2025 at the Ume River gauge station. The SPI values are –0.30 in January, 0.05 in February, and –0.40 in March. These values suggest generally dry conditions, with January and March falling into the mild drought category, while February indicates near-normal conditions.

SPI values between 0 and –0.99 typically signal mild to moderate drought, and values below –1.0 indicate more severe drought conditions (McKee et al., 1993). The forecasted values—although not extreme—highlight the persistence of below-average rainfall in two of the three months. This trend is consistent with historical observations that show the Ume River catchment receives relatively low rainfall compared to other stations in the Lower Kariba catchment.

These drought predictions are particularly important for water resource planning and management. For instance, reduced rainfall implies lower inflows into Lake Kariba, which affects water allocation to hydropower utilities, potentially leading to reduced electricity generation capacity. Additionally, drought forecasts inform agricultural planning, irrigation scheduling, and early warning systems for drought preparedness and mitigation across sectors that depend on seasonal rainfall.

7.0 CONCLUSION AND RECOMMENDATIONS

This study developed a localized drought prediction model for the Kariba Lower Catchment using Long Short-Term Memory (LSTM) networks, with input data processed through Gamma distribution-based SPI analysis. The model demonstrated strong performance in capturing seasonal rainfall variability, with R^2 values above 0.50 at well-instrumented stations. Despite limitations such as short data records and zero rainfall months, the Gamma distribution effectively modelled precipitation characteristics across the catchment. The tool provides reliable, short-term drought forecasts to support hydropower planning, water allocation, and climate resilience in the Zambezi Basin. Future work should focus on full-basin integration, inclusion of additional variables like soil moisture, evapotranspiration, and climate scenario modelling for long-term planning.

Despite its widespread use, fitting parametric distributions such as Gamma is not without challenges. In data-limited contexts, especially when records are short or spatial coverage is sparse as is the case in many African catchments — parameter estimation may be unstable. This is problematic when there are months with zero rainfall, which can skew the distribution and reduce model reliability. Methods such as maximum likelihood estimation (MLE) are often employed for parameter fitting, although they may be biased with small sample sizes. Some authors have addressed this by incorporating mixed models or using regionalisation techniques to improve parameter robustness in ungauged basins.

REFERENCES

Edwards, D.C. & McKee, T.B., 1997. Characteristics of 20th century drought in the United States at multiple time scales. *Climatology Report No. 97-2*. Fort Collins, CO: Colorado State University. gmao.gsfc.nasa.govdrought.emergency.copernicus.eu

Husak, G.J., Michaelsen, J.C. & Funk, C.C., 2007. Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. *International Journal of Climatology*, 27(7), pp. 935–944

Guttman, N.B., 1999. Accepting the Standardized Precipitation Index: a calculation algorithm. *Journal of the American Water Resources Association*, 35(2), pp.311–322.

Hasan, Balqees, 2011. Rainfall Statistical Analysis in Sudan: 1965-2005 using Time Series and regression methods. PhD. Dissertation, Islamic University of Um-Durman, Sudan. In Arabic.

Husak, G.J., Michaelsen, J.C. & Funk, C.C., 2007. Use of the gamma distribution to represent monthly rainfall in Africa for drought monitoring applications. *International Journal of Climatology*, 27(7), pp. 935–944

McKee, T.B., Doesken, N.J. & Kleist, J., 1993. The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, Anaheim, CA, pp.179–184.

Thom, H.C.S., 1958. A note on the Gamma distribution. *Monthly Weather Review*, 86(4), pp.117–122.

Wilks, D.S., 2011. *Statistical Methods in the Atmospheric Sciences*. 3rd ed. San Diego: Academic Press.

World Meteorological Organization (WMO), 2012. *Standardized Precipitation Index User Guide*. Geneva: WMO-No. 1090.

Hochreiter, S. and Schmidhuber, J. (1997) ‘Long Short-Term Memory’, *Neural Computation*, 9(8), pp. 1735–1780.

McKee, T.B., Doesken, N.J. and Kleist, J. (1993) ‘The relationship of drought frequency and duration to time scales’, *Proceedings of the 8th Conference on Applied Climatology*, Anaheim, CA, American Meteorological Society, pp. 179–184

ANNEX A: SPI-3 Time series tables