International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Integrating Indigenous Knowledge with Deep Learning for Meteorological Drought Prediction

Leelavathy S. R.[1], A. Mary Mekala[2*]

Submitted: 09/10/2023 **Revised**: 29/11/2023 **Accepted**: 10/12/2023

Abstract: Drought is a natural disaster creating huge impacts in three areas of economy, environment and social. It becomes hard to predict drought, its onset and duration due to complex interaction of multiple factors. Though various scales and forecasting methods have been proposed, recent climatic variations caused by Global warming makes most of the scales and forecasting methods inaccurate. Most of existing solutions are based on seasonal behavior and correlation to other influencing factors like temperature and humidity. They are at larger coverage level and not specialized to cover smaller regions. Also with factors like global warming are affecting the baseline periodicity assumptions, there is a need to improvise meteorological factors based drought prediction methods. This work proposes a solution to this problem by integrating indigenous knowledge (IK) with deep learning forecasting methods through attention mechanism referred as IK fused attention networks. The indigenous knowledge view over precipitation, temperature, wind speed and humidity are integrated with LSTM based forecasting through attention mechanism to improve the accuracy of drought prediction. The performance of the proposed solution was tested against meteorological data collected from Karnataka disaster monitoring center for Chitradurga district of Karnataka. The proposed IK fused attention network is able to provide at least 1.2% higher NSE and 33% lower MAE in prediction of SPI compared to existing works

Keywords: *Drought forecasting, indigenous knowledge, attention vector, SPI*

I. Introduction

Drought is a natural disaster characterized by intense and persistent precipitation shortage which can last for months or years [1]. Compared to other natural disasters, the hazardous footprint of drought is higher as it generally over lasts for longer duration. It impacts food production, reduces life expectancy and shrinks economy. Droughts are generally of three types: Meteorological, Hydrological and Agricultural. Prolonged precipitation, higher temperature, higher winds, lower relative humidity and greater sunshine results in meteorological drought [2]. The persistence of meteorological drought depletes the ground water levels and reduces soil water deficiency creating agricultural drought. The impact of agricultural drought reduces biomass and agricultural yield creating hydrological drought [3]. Drought creates impacts in all level of economic, environmental and social. It creates adverse effects on food production, increasing farmers' suicide rate, excess heat, lower power generation, reduced industrial production, and human and animal health deterioration. The frequency and intensity of short span droughts has increased during last two decades [31]. These droughts have adverse effects on agricultural economy like India. In India agriculture is mostly dependent on monsoon rains. Agriculture contributes almost 14% of total gross domestic product and drought can cause huge financial and food security problems. The drought risk management process is heavily dependent on effective drought forecasting methods.

There are more than hundred different drought indices and various machine learning techniques proposed for various purposes of drought severity prediction, start/end of drought and location of drought [4].Some of the well-known drought indicators are Palmer Drought Severity Index (PDSI)[32]. Standard Precipitation Index (SPI)[33], Standardized Nonstationary Precipitation Index (SnsPI)[34], Joint Deficit Index (JDI)[35] and Copula-based Joint Drought Index[36]. Metrics like PDSI, JDI and CJDI require multiple meteorological inputs. Compared to them SPI is simple metric which is based only on precipitation data. SPI is a more robust meteorological index and it is used in many recent works successfully [37]. Most of the methods are based on historical values of environmental observations. Global warming has created large temperature variations. The global temperature increase of 0.5 to 2 degree Celsius and large climatic variations makes most of scales & machine learning techniques inaccurate and needs adaptation. Many attempts have been made integrating multiple factors like climatic, oceanic, agricultural yield to increase the prediction accuracy. In recent years, there is resurgent interest in applying traditional indigenous knowledge in drought prediction. Many indigenous knowledge based systems have been proposed (discussed in section II). Though use of local knowledge has many advantages, it is not a valid system in its own right. The strength of Indigenous knowledge systems is its locality specific prediction. But meteorological forecasting methods predict for a larger area. Integrating indigenous knowledge

[1,2]*School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India*
Email id: [1]leelavathys.r2015@vit.ac.in, [2]amarymekala@vit.ac.in

systems with meteorological forecasting methods can increase the prediction accuracy for the local regions and reduce the risks for farmers.

This work proposes a drought prediction system integrating indigenous knowledge with meteorological forecasting methods referred as IK fused attention networks. Deep learning forecasting with meteorological factors is augmented with attention vectors generated from indigenous knowledge systems. By this way, the meteorological forecasting coverage can be brought to cover smaller regions with a improved accuracy. Following are the contributions of this work

(i) Attention vector model for representing the indigenous knowledge (IK).

(ii) A novel LSTM improved with attention vector of IK to provide forecasting for a smaller coverage area

(iii) Validation of the proposed model for south Indian districts.

The rest of the paper is organized as follows. Section II surveys the drought prediction systems and presents the case of integration of IK with meteorological forecasting systems. Section III details the proposed integration of IK with deep learning based meteorological forecasting. Section IV presents the results of the proposed integrated forecasting method. Section V presents the conclusion and future scope of work.

II. Related Work

The survey is conducted in two categories of meteorological parameter based and traditional parameter based.

A. *Meteorological parameters based prediction*

Jena et al [6] fitted a Coupled Model Intercomparison Phase 5 (CMIP5) model over the historical monsoon rainfall data in Indian subcontinent to predict droughts. Assessment of flood or drought is made based only on monsoon rain fall distribution. Zhang et al [7] experimented with three models of multiple linear regression (MLR), long short term memory (LSTM) and random forest (RF) to predict flash drought. Eight meteorological variables: precipitation, air temperature (average, minimum & maximum), air pressure, relative humidity, wind speed, sunshine duration were collected from 1979 to 2016. A derived variable potential evapotranspiration is calculated from air temperature, humidity and wind speed. Three different models of MLR, LSTM and RF was fit with the nine variables as input and drought intensity as output. The output variable of drought intensity was constructed based on depletion rate of soil moisture. But the method is for larger area and drought intensity measurement based on soil moisture alone is not sufficient. Zhang et al [8] analyzed drought from perspective of crop production. The entire wheat growing season of October to April was considered for the study. The correlation between different stages of wheat growth to different types of droughts was analyzed. Four droughts of Meteorological, Hydrological, Soil Moisture and Vegetation were considered in this study. Time series data for every month from 1981 to 2013 were used in the study. The study provided influence of each drought type, their spatial temporal distribution and variation on wheat growth. But the method did not propose any prediction model.

Zhang et al [15] used LSTM to predict soil moisture drought. LSTM model was trained with historical soil moisture data from 1980 to 2012. Prediction based on historical data alone cannot yield higher accuracy as the baselines are disturbed due to climate variability. Poornima et al [16]used LSTM to predict SPI and SPEI predictions in time scale of 1, 6 and 12 months. The dataset consisting of maximum temperature, maximum relative humidity, minimum relative humidity, precipitation, wind speed, sunshine and evaporation were collected from year 1958 to 2014. LSTM is found to provide better compared to ARIMA. The study is at macro level and does not consider giving more weightage to certain variables depending on the geographic conditions. Sumin et al [17] developed a short term drought forecasting model based on temporal patterns of satellite based drought indexes and numerical model outputs. Two machine learning classifiers of convolutional LSTM and random forest (RF) were integrated. Convolutional LSTM predicts the temperature and precipitation based on historical drought conditions. This output along with static variables like elevation, land cover and climate zone is fed into RF to predict drought. The approach is suitable for only large area. Xu et al [18] proposed a hybrid model integrating ARIMA with LSTM model for short term drought forecasting. Hybrid ARIMA-LSTM model was found to provide better accuracy compared to ARIMA, support vector regression and LSTM classifiers. Band et al [19] analyzed the drought index forecast accuracy of different time series models. Seasonal, non-seasonal and combined differencing models were experimented. The study found that combined differencing provides better forecasting accuracy. Rajib et al [20] used one dimensional convolutional neural network for drought assessment based on the complex association between rainfall variation and hydro meteorological parameters of air temperature, surface pressure, wind speed, relative humidity, evaporation, soil moisture and geo potential height. CNN is able to learn the complex association between meteorological parameters to predict rainfall. From the rainfall predictions, SPAI drought index is calculated. The approach does not consider the temporal variations in the complex relationship modeling.Dikshit et al [21] used stacked LSTM architecture for predicting drought measure. Standard Precipitation Evaporation Index was predicted using hydro meteorological and climatic variables. LSTM performance was enhanced by loss measurement and weight updating using gradient descent operator. Mokhtar et al [22] experimented with four different machine learning models of RF, Extreme Gradient Boost (EGB) , CNN and LSTM. Classifiers were trained to predict Precipitation Evapotranspiration Index (SPEI) for two time scale of 3 months (SPEI-3) and 6 months (SPEI-6). The classifiers were trained with precipitation, average temperature, minimum temperature, maximum temperature, wind speed and relative humidity.XGB was found to provide higher forecast accuracy for SPEI even higher than LSTM. XGB performed better without consideration for climatic variability, but in presence of it, the performance of XGB may drop.

B. *Traditional parameters based prediction*

Salite et al[5] explored the traditional indicators for drought prediction in Gaza. Author detailed the eleven traditional indicators based on moon appearance, cloud's appearance, wind direction, stars pattern and plant behavior. Authors also inferred that most of the traditional indicator performed negatively due to large climate and temperature variations. They stressed that technical indicators alone were not sufficient for achieving higher prediction accuracy. Balehegn et al [9] documented the indigenous weather and climate forecasting knowledge among the Afar communities. These communities forecast climate and weather from bio physical variables in their environment. Bo physical variables are extracted from various cues of trees, animals, winds and celestial bodies. Lack of sufficient data make it difficult to prove the validity of traditional knowledge systems statistically. Owen et al [10] identified multiple indigenous indicators used in farming communities of Africa. These indicators provide short time and long time seasonal information necessary for crop planning decisions. Author stressed the need for integrating IK with modern forecasting to improve the reliability of forecasting but no integration models were proposed. Pareek et al [11] explored the IK prevalent in tribal communities of Rajasthan. IK relating to cloud formation, lightning, wind direction, rains, drought, disaster prediction, response, mitigation, and effects of weather on crops were collected and analyzed. It was an exploratory study, it collected the tribal IK on drought but authors did not construct any model based on it. Ngenga et al [12] assessed the IK used by farmers in Kenya on climate forecasting, their perceptions of climate variability and adaptation strategies, and their correlation with conventional approaches. Kenyan farmer's IK were based on behavior of trees, animals, the sky, moon, and wind similar to rest of African countries. The correlation between the IK and conventional is measured using Chi Square statistical tests. The test found that the farmer observations were in agreement with historical observation of rain fall and drought. The author evaluated the observations but did not propose any model or tool to forecast based on the observations. Johnston et al [13] analyzed the IK forecasting practices prevalent among Zimbabwe farmers. The study found that IK practices are in threat due to climate variability and need to be adapted by integrating to modern forecasting methods. Islam [14] documented the IK practices among rural communities in Barind. Most of the practices were based on birds and inspect behavior patterns. Though the study documented the practices, the relevance was not tested and no prediction models were constructed based on the observations.

From the survey, many IK based observations with high correlation to drought were found. But they were not quantitative in terms of drought scales like SPI. Multi-variate LSTM trained with meteorological variable of precipitation, temperature (average/minimum/maximum), relative humidity and wind speed was found to perform better among all classifier in predicting drought. The LSTM provided quantitative results in form SPI scales. Due to higher climatic variability, the baselines data on LSTM based forecasting models built have wide shifts introducing large prediction error. Also some of IK observation does fail due to climatic variations. In the survey, many works have recommended integration of IK with meteorological forecasting methods but to our knowledge no working models have been proposed. The prediction error due to climatic variations can be minimized with integration of IK with meteorological forecasting methods. Integration of IK with LSTM models brings another advantage to extend the LSTM model trained on larger coverage areas to forecast for local spatial area. In the local spatial area, the importance of the meteorological variable varies from that of global coverage level and this creates prediction error.

III. IK fused Attention Networks

The architecture of the proposed solution is given in Figure 1. IK rules based on hydro meteorological parameters are tested for reliability by matching against historical dataset. The rules whose score is more than a threshold are selected and the hydro meteorological parameters covered by the rules are decided as dominating variables. Conditional attention vector is generated based on the dominating variable and value ranges for inputs. This conditional attention vector is provided to the LSTM attention model. LSTM attention model is a multivariate time series model which takes the hydro meteorological parameters as input and provides the drought index value as output.

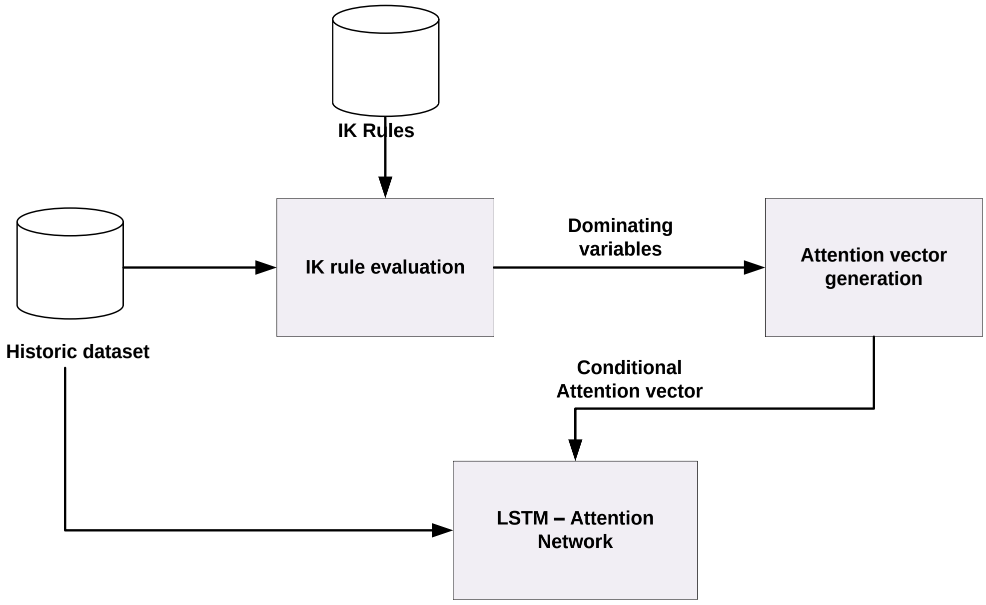

Fig 1 Integrated IK-LSTM Model

This work experiments with SPI drought index. The experiments are conducted for two scales of three and six months.

The proposed solution has three important stages: IK rule evaluation, Attention vector generation and LSTM-Attention network model training and forecast. Each of the stages is detailed in below subsections.

A. IK Rule evaluation

The architecture of the IK rule evaluation is given in Figure 2.

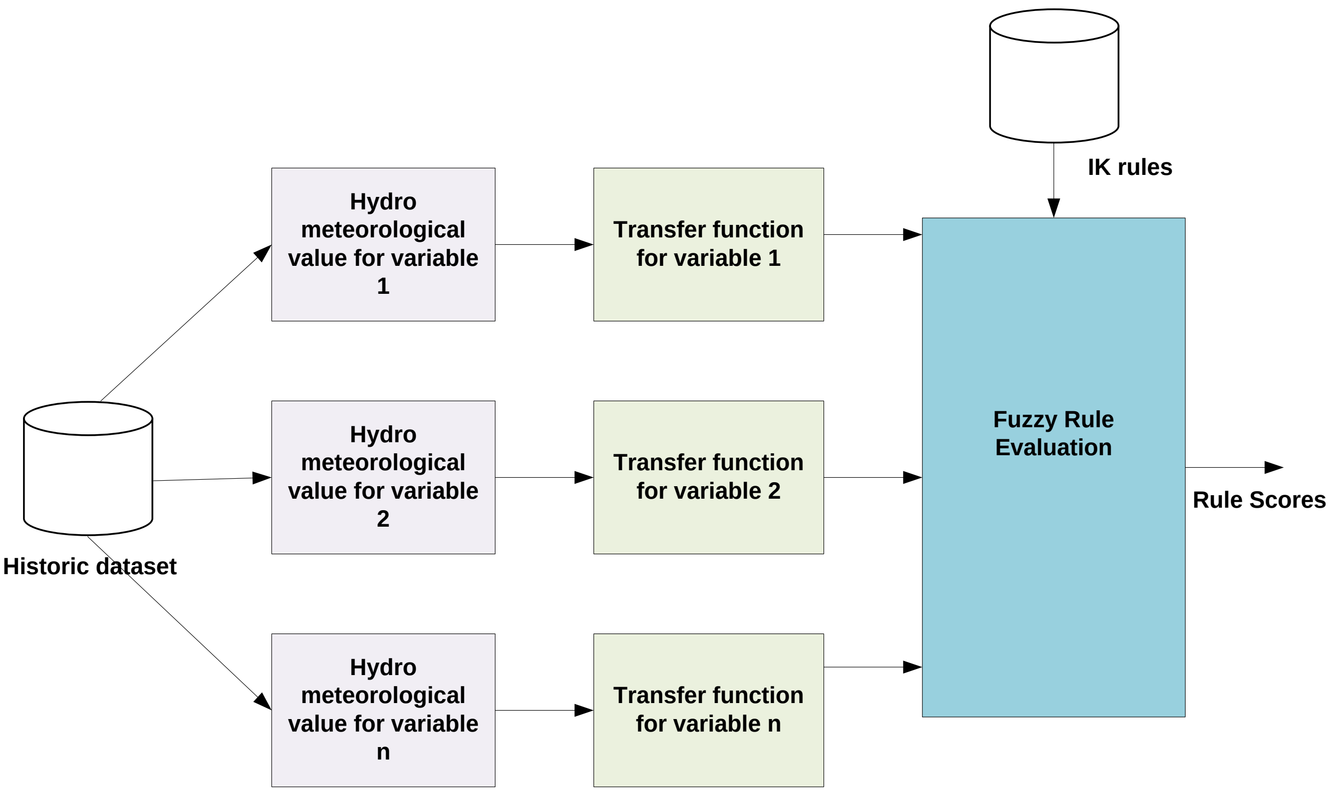

Fig 2 Architecture of IK Rule evaluation

The IK rules are evaluated using fuzzy logic system across the historical data and each rule is scored. The IK rules are fuzzy. They are based on categorical values. Say the rule is "*if wind speed is high in June then low rainfall*", the values for wind speed (high) and rainfall (lower) are categorical instead of numerical.

To evaluate the validity of the rules, meteorological variables in historic dataset must be used. But the values in the dataset are numeric. Thus transformation must be done as the first step. The transformation function mapping each of the meteorological variables to fuzzy categorical variables must be defined. A sample of the transformation function mapping wind speed is designed depending on the spatial location over a period of year. For the wind speed in Chennai given in Figure 3, the maximum value is 5.9 m/s and minimum value is 5.1 m/s. The transformation function constructed with three categorical variables of low, medium and high is shown in Figure 4.

Fuzzy function is used for transformation as the views are not concrete in IK. The transfer function is constructed for SPI with two categorical variables with chance of drought as HIGH or LOW.

For each IK rule, a fuzzy system [28] is constructed with meteorological variables as input and SPI as the output. The fuzzy system is tested for each of the historical values, and ratio of number of historic values rows classified correctly to the total number of historic values is given as validity score.

$$score(rule) = \frac{Number\ of\ historic\ vlaues\ classified\ correcctly\ with\ rule}{Total\ number\ of\ historic\ values}$$

The rules whose score is greater than threshold are decided as valid IF rules. The variables covered in the valid rules are the dominating variables.

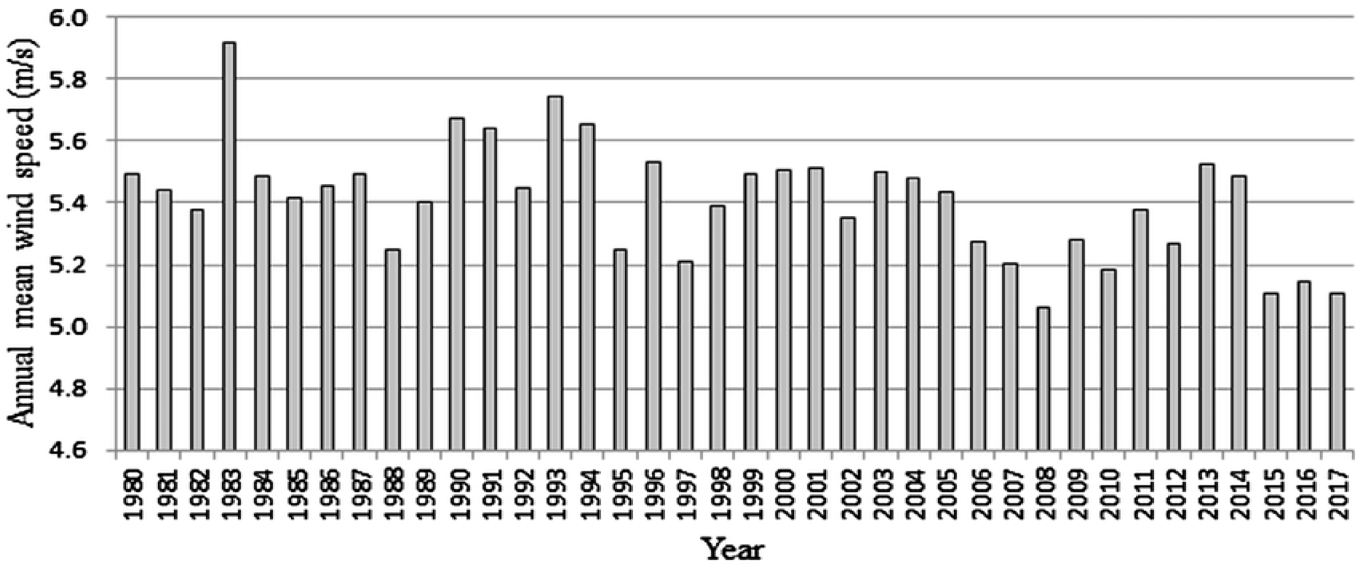

Fig 3 Wind speed in Chennai

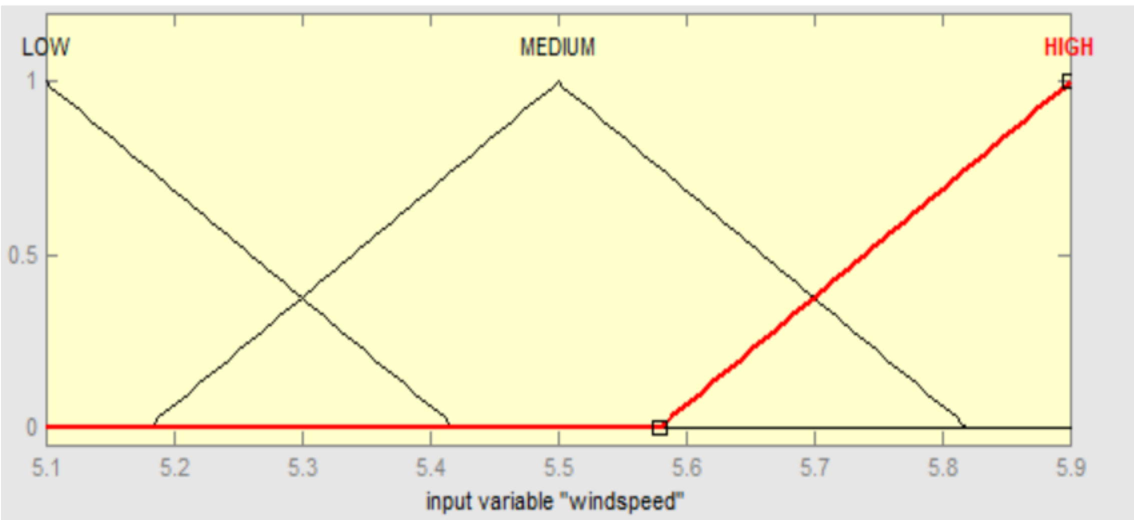

Fig 4 Fuzzy transfer function for wind speed

Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
0	0	0	0	0	1	0	0	0	0	0	0

Full moon	No Moon	Any
0	0	1

Fig 5 Sample positional encoding

B. Conditional Attention vector generation

The drought will be predicted based on multiple hydro meteorological variables using LSTM in this work. But typical LSTM gives equal importance to all the variables. But for localized prediction, certain variables may have higher importance in prediction. To enforce prediction with higher weights to certain features based on IK, attention model is used in this work. The attention mechanism selects the most influencing input features and provides higher Attention vector is created from the IK rules. The query is represented in form of positional encoding information of IK rule. Say for the IK "“*if wind speed is high in June then low rainfall*”, June is the positional information. The positional information is represented in form of binary vector. A sample positional encoding information is given in Figure 5. They key is binary vector with each position

weights to corresponding original sequence. The output of attention model is provided as input to the LSTM model to predict drought in terms of SPI. The attention model maps a query(Q) and set of key value pairs ($< K, V >$) to output (O). With Q, K, V represented as vectors, the output is calculated as a weighted sum of the values, where the weight assigned to each value is computed by a compatibility function of the Q with the corresponding key (K).

representing a hydro meteorological variable and the position is set as 1, when the corresponding meteorological variable is covered in the IK rule else it is marked as 0. The value is a binary vector with each position representing a hydro meteorological variable. The value at the position is score of IK rule when the variable is covered in the IK rule. The query, key and value represent the attention vector.

C. LSTM Attention Network

Multivariate LSTM [25] with attention is proposed in this work for drought index prediction. LSTM has been selected as it is proved to provide higher accuracy of drought prediction compared to seasonal models like ARIMA [26].Attention modeling with LSTM [27] allows for providing different weights to features depending on positional information of time sequence data and its coverage in IK rules.

Given a sequence of meteorological observations $X = (X_1, X_2, \ldots X_T)$ where each X_i is the set of meteorological

variables and its corresponding drought prediction. It is represented as $X_t = \{x_t^1,\ x_t^2, \ldots x_t^N, y_t\}$ where N is the number of meteorological variables ,x_t^i is the meteorological variable and y_t is the drought index. This work uses following variables: precipitation, temperature, relative humidity and wind speed as input. SPI is used as drought index in separate models. The objective is to predict the drought index value based on inputsequence X with weight importance to features given by attention vector. The LSTM -Attention architecture is given in Figure 6.

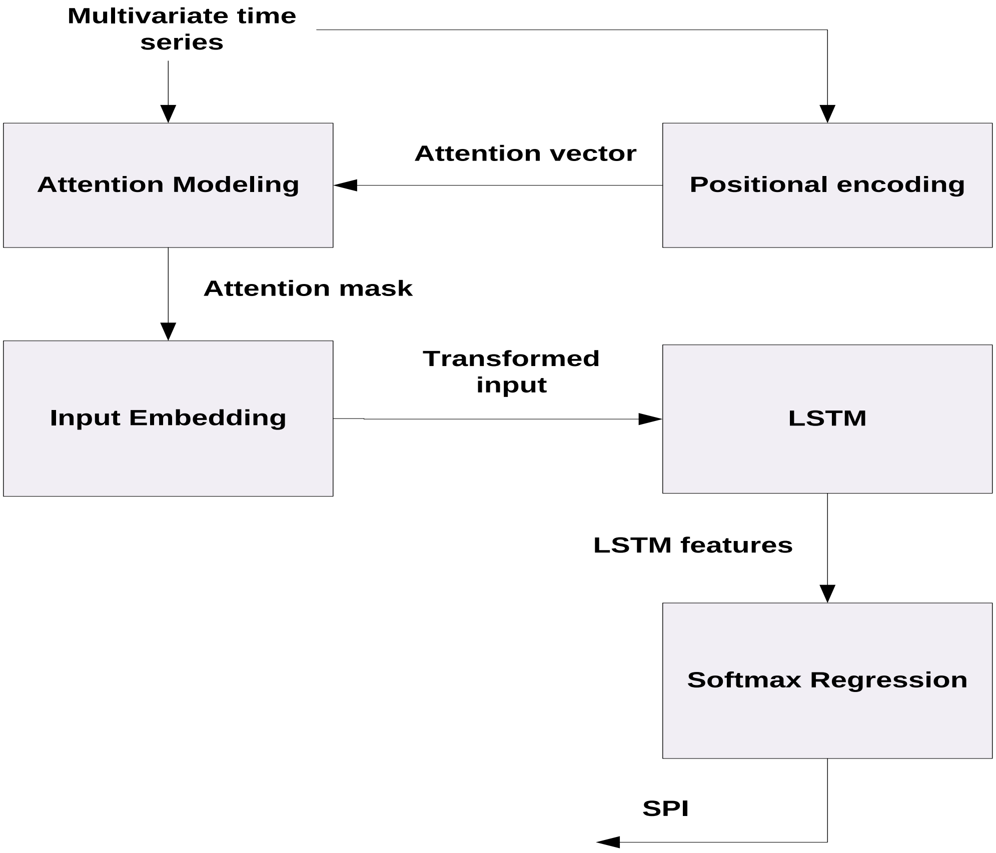

Fig 6 LSTM Attention architecture

For each input sequence$X = (X_1, X_2, \dots X_T)$, the positional encoding is constructed by creating the time domain specific encoding (as given in Figure 5) for each of the X_t. With positional encoding a query to the Attention modeling, the attention weight values providing differential importance to each of the features of X_i is obtained as mentioned by attention vector generation procedure given in Section B as $\{\beta_1, \beta_2, \dots \beta_N\}$ The input $X_t = \{x_t^1, x_t^2, \dots x_t^N\}$ is transformed to Z_i where

$$Z_i = \{\beta_1 x_t^1, \beta_2 x_t^2, \dots . \beta_N x_t^N\}$$

The transformed input is passed to input embedding. In typical NLP architecture, words in sentence are mapped to high dimensional vector to capture dependencies across different words without considering temporal information. In the same way, the dependencies across the variables in Z_i are learnt using input embedding. The embedding is done by passing Z_i to a 1D convolutional neural network to get a d dimensional embedding for Z_i.

The sequence of input embedding are provided as input to the multivariate LSTM model to predict the drought index.

LSTM is an extension of RNN (Recurrent Neural Networks). It has gating mechanism and a cell activation state, in addition to the existing hidden state, since the network learns when to forget long-term information and when to incorporate new information. Separating the hidden state with the cell activation state also allows for the network to learn controlling how much of the cell activation it outputs. The structure of LSTM is given below

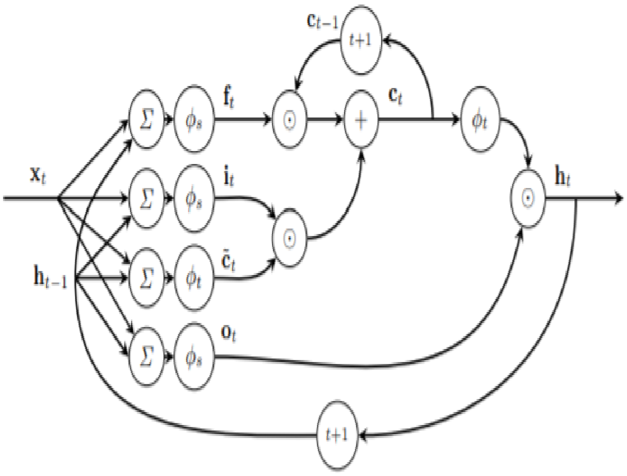

Fig 7 LSTM structure

A combination of an input vector x and the previous hidden state is taken as input by an LSTM node.

A new candidate cell activation ˜c is calculated by the LSTM. It is calculated as the weighted sum of the inputs and bias b.

The result is then passed to a hyperbolic tangent activation function as given below

$$c_t = \emptyset_t(W_c x_t + U_c h_{t-1} + b_c)$$

c_t is the candidate cell activation. x_t is the input vector. W and U are the weight matrices.h_{t-1} is the hidden state vector at the previous time step and b_c is the bias. the gates control how much of activation must be retained and how much must be forgot. Input gate control how must activation to retain and forget gate decided how much cell activation must be forgot. The final gate is incorporated to calculate the hidden state.

$$f_t = \emptyset_s(W_f x_t + U_f h_{t-1} + b_f)$$

$$i_t = \emptyset_s(W_i x_t + U_i h_{t-1} + b_i)$$

$$o_t = \emptyset_s(W_o x_t + U_o h_{t-1} + b_o)$$

f_t is the forgot gate vector. i_t is the input gate vector.o_t is the output gate vector.

The architecture of drought prediction LSTM is given in Figure 7.

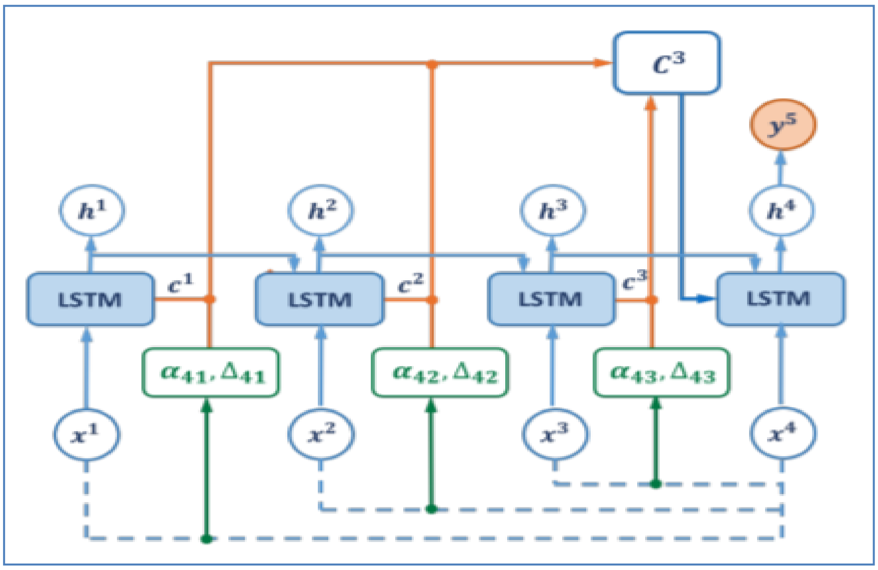

Fig 8 LSTM architecture

It takes the $Z = (Z_1, Z_2, \ldots Z_T)$, where T observation are used to predict the drought index at time T+1 and each Z_i is the input embedding of the transformed original sequence $X = (X_1, X_2, \dots X_T)$. The final LSTM layer output is passed to a Softmax classifier in regression setting [29]. In regression setting, softmax classifier the LSTM output to one of possible value of drought indicator. Say there are K drought indicator values $\{1,2, \dots K\}$, the softmax classifier must estimate the probability for each of the K values. The output of the softmax classifier is the K dimensional vector providing the K estimated probabilities. The loss function for training the softmax regression classifier is given as

$$L = -[\sum_{i=1}^{m} \sum_{k=0}^{1} 1\{y^{(i)} = k\} \log P(y^{(i)} = k | z^{(i)}; \theta)]$$

Where

$$P\left(y^{(i)} = k \middle| z^{(i)}; \theta\right) = \frac{\exp(\theta^{(k)} z^{(i)})}{\sum_{j=1}^{K} \exp\left(\theta^{(k)} z^{(i)}\right)}$$

Where $\theta^{(1)}$,$\theta^{(2)}$,…$\theta^{(k)}$ are the parameters of the model and $\exp(\theta^{(k)} z^{(i)})$ is the normalization of parameter with the input feature values. The proposed solution has two important process- training and prediction. The process flow of training and prediction is given as flowchart in Figure 9 and Figure 10.

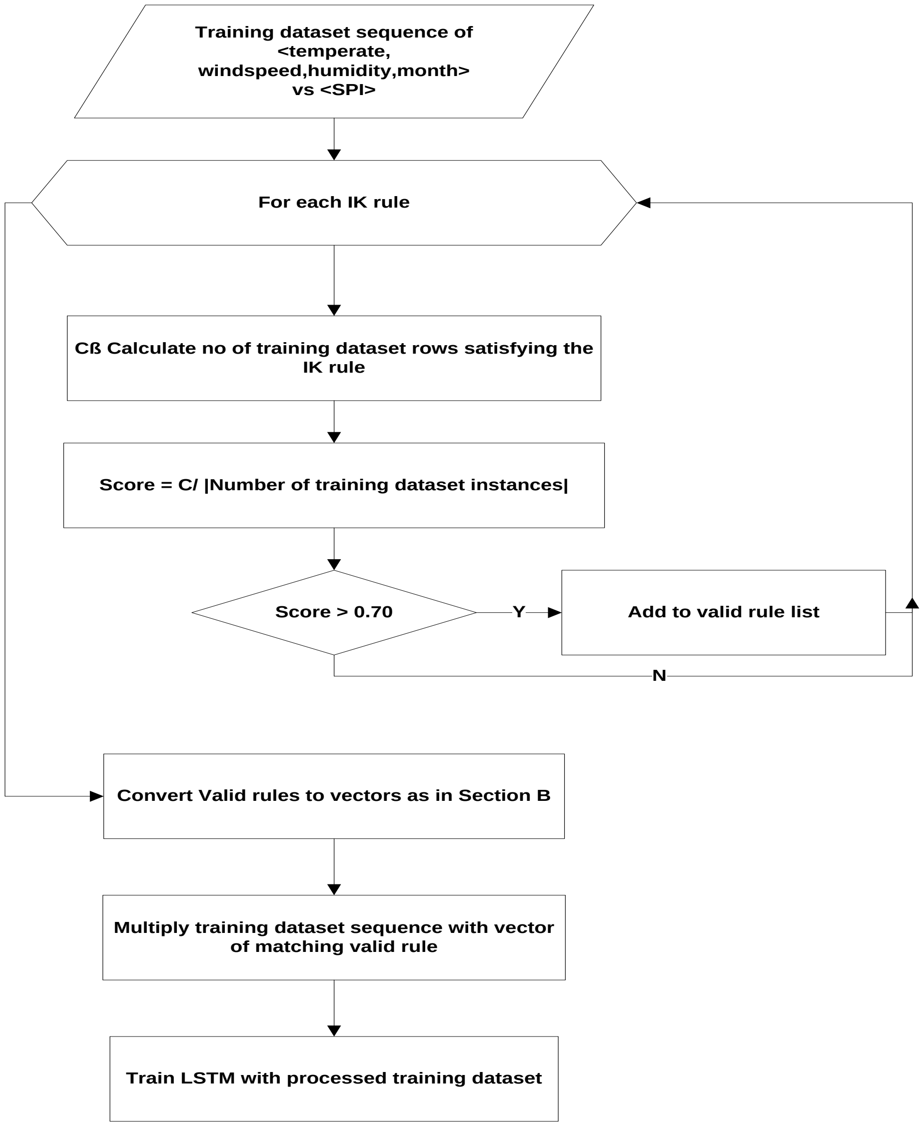

Fig 9 Training process flow

In the training process, the training dataset instances is processed by multiplying with the conditional vectors generated for matching IK rules and a LSTM regressor is trained with the processed dataset.

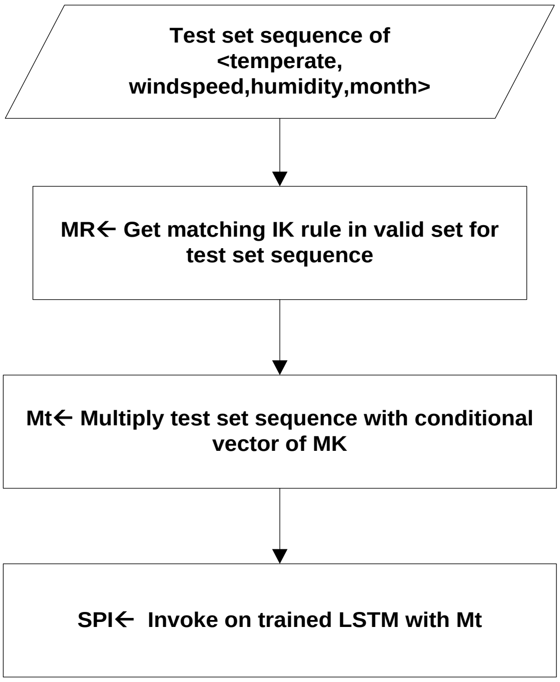

Fig 10 Prediction process flow

The prediction process takes the test set sequence of meteorological variables. It then multiples the test set sequence with the conditional vector of matching IK rule. The processed test set sequence is passed to LSTM regressor to predict the SPI value.

IV. Results

The performance of the proposed solution was tested by implementing in Python. Keras and Tensor flow modules are used to realize the attention vector mechanism and integration with LSTM. The performance is tested by collecting the meteorological variables temperature, humidity, wind speed and precipitation over a period of years in interval of every month for certain location. For every month SPI value is calculated. The sequence of input (temperature, humidity and wind speed) is arranged in window interval (3, 6, 12 months) and for each sequence SPI value is associated as output. This dataset is split into 80:20 ratio with 80% for training the classifier and 20% is used for testing the classifier. The SPI value predicted by the classifier is compared against the actual SPI value to measure the effectiveness of the proposed solution. The performance of the proposed solution is measured in terms of : Nash–Sutcliffe model efficiency coefficient (NSE) [30], the mean square error (MSE), the mean absolute error (MAE) , mean bias error (MBE) and correlation coefficient (R). The metrics are calculated as below

$$NSE = 1 - \frac{\sum(P_i - A_i)^2}{\sum(\bar{A} - A_i)^2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|A_i - P_i|$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)$$

$$R = \frac{\sum_{i=1}^{n}(\bar{A} - A_i)((\bar{P} - P_i)}{\sqrt{\sum_{i=1}^{n}(\bar{A} - A_i)^2}\sum_{i=1}^{n}(\bar{P} - P_i)^2}$$

In the above equations, n is the number of test observations, A is the actual value and P is the predicted value.The performance of the proposed solution is compared against XGBoost, random forest (RF) and LSTM which are trained only using meteorological variables without consideration for IK attention vectors.

The drought is measured in terms of SPI in time scales of 3, 6, 9 and 12 months. SPI (McKee et al 1993) is metric to quantize the precipitation scarcity on different time scales. SPI is calculated in terms of current precipitation (Y), average precipitation $(\bar{Y})$, standard deviation of precipitation (σ) over the period of time as

$$SPI = \frac{Y - \bar{Y}}{\sigma} \times 100$$

Based on the SPI values, drought is classified to five classes as shown in Table 1.

Table 1 SPI values to Drought class mapping

Drought class	SPI Values
Wet	>1.5
Slightly wet	1.0 to 1.49
Normal	-0.99 to +0.99
Slightly Dry	-1 to -1.49
Dry	<1.5

The multivariate time series is arranges in 3 month interval and 6 month interval to predict. SPI value at end of the interval is converted to Drought class and the sequence of training set is prepared. Chitradurga district of Karnataka (Figure 9) was selected for testing the drought prediction accuracy of proposed solution. This region is situated in Agro climatic region-10 with average rainfall of 592.5mm and average humidity of 58-76%. The region has 32 rainy days with usual showers in June to September. South west monsoon is the major contributor to the rainfall in this region. The temperature ranges from 21°C to 31.8°C. Meteorological data for Chitradurga district were collected from year 1987 to 2017 from Karnataka state natural disaster monitoring center.

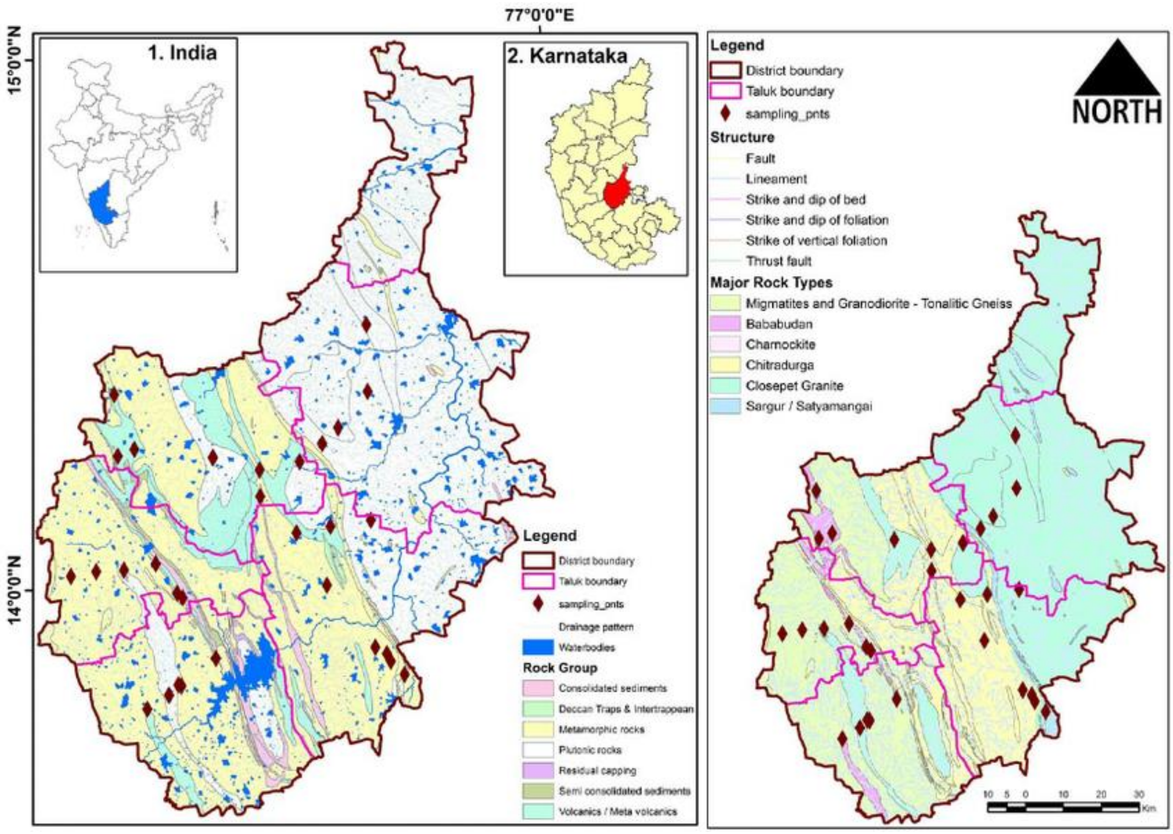

Fig 11 Chitradurga district

The IK knowledge for drought were extracted from [24] and translated for this work. The translated IK are given in Table 2.

IK	Translated IK
Wind flow on continuous interval over months reduces rain.	If wind flow is high, then slightly dry
A higher heat in may month may cause high monsoon	If temperature is high in in May then Wet
Higher rain follows Holi rain.	It rain in March, then Wet

The translated IK is used for attention vector generation and improving the LSTM attention model. The comparison of proposed LSTM with attention model (trained with fused IK attention vector and meteorological variables) with machine learning models (trained only with meteorological variables) are presented below.

Table 2 Comparison of MSE

	MSE			
	Proposed LSTM with attention	**XGBoost**	**RF**	**LSTM**
SPI-3	0.11	0.17	0.18	0.14
SPI-6	0.09	0.15	0.14	0.13
SPI-9	0.08	0.11	0.12	0.11
SPI-12	0.07	0.10	0.11	0.10
Average	0.08	0.13	0.13	0.12

The MSE is compared across the solutions for all SPI scales and the result is given in Table 2. Lower the value of MSE, better is the prediction accuracy. The proposed solution has 62% lower MSE compared to XGBoost & RF and 50% lower MSE compared to LSTM.

Table 3 Comparison of NSE

	NSE			
	Proposed LSTM with attention	XGBoost	RF	LSTM
SPI-3	0.80	0.73	0.73	0.79
SPI-6	0.81	0.75	0.75	0.80
SPI-9	0.82	0.77	0.79	0.81
SPI-12	0.85	0.79	0.80	0.83
Average	0.82	0.76	0.77	0.81

The NSE is compared across the solutions for all SPI scales and the result is given in Table 3. Higher the value of NSE better is the prediction accuracy. The proposed solution has 7% higher NSE compared to XGBoost, 6% higher NSE compared to RF and 1.2 % higher NSE compared to LSTM.

Table 4 comparison of MAE

	MAE			
	Proposed LSTM with attention	**XGBoost**	**RF**	**LSTM**
SPI-3	0.29	0.35	0.36	0.31
SPI-6	0.26	0.34	0.35	0.27
SPI-9	0.23	0.31	0.32	0.24
SPI-12	0.21	0.29	0.29	0.22
Average	0.24	0.32	0.33	0.26

The MAE is compared across the solutions for all SPI scales and the result is given in Table 4.Lower the value of MAE, better in the prediction accuracy. The proposed solution has 33% lower MAE compared to XGBoost, 37% lower MAE compared to RF and 8% lower MAE compared to LSTM.

Table 5 Comparison of MBE

	MBE			
	Proposed LSTM with attention	**XGBoost**	**RF**	**LSTM**
SPI-3	-0.16	0.04	0.05	-0.12
SPI-6	-0.023	-0.01	0.03	-0.08
SPI-9	-0.02	-0.02	-0.06	-0.03
SPI-12	-0.01	-0.05	-0.08	-0.04
Average	-0.05	-0.01	-0.015	-0.0675

The MBE is compared across the solutions for all SPI scales and the result is given in Table 5. The proposed solution has 35% lower MBE compared to LSTM.

Table 6 Comparison of R

	R			
	Proposed LSTM with attention	XGBoost	RF	LSTM
SPI-3	0.88	0.83	0.82	0.86
SPI-6	0.89	0.84	0.84	0.87
SPI-9	0.90	0.86	0.86	0.89
SPI-12	0.92	0.87	0.87	0.90
Average	0.897	0.85	0.847	0.88

The R is compared across the solutions for all SPI scales and the result is given in Table 6. Higher the R value, better is prediction accuracy. The R value in proposed solution is 5.2% higher compared to XGBoost, 5.5% higher compared to RF and 1.8% higher compared to LSTM.

Integrating IK with LSTM has improved the prediction accuracy in the proposed solution as seen from the metrics results. Machine learning models trained only with meteorological variables had higher prediction error compared to the proposed model which is trained with fused meteorological variables with IK based condition vectors. Compared to traditional parameter for drought prediction methods discussed in the literature, the proposed solution provides quantitative results and it is more scientific as results are backed with support of meteorological variables.

The proposed solution is easily extensible for any number of IK rules. This work has proposed a mechanism to convert IK rules expressed in linguistics term to quantitative vector using fuzzy logic. To the best of our knowledge, there has no earlier work on translation of linguistic IK rules to quantitative feature vector and fusion of modern meteorological variables with IK quantitative feature vector.

V. Conclusion

This work proposed a LSTM attention model integrating indigenous knowledge for meteorological drought prediction. As part of the work, IK rules in fuzzy domain are evaluated against historical meteorological data. The IK rules with higher significance are used for attention vector generation.

Attention vector generated from IK rules is merged into LSTM for providing differential weights to the features. The differential weighted features are classified by the multivariate LSTM to SPI drought prediction index. The performance of the proposed solution is tested for meteorological data of Chitradurga district of Karnataka. SPI prediction accuracy was higher in proposed solution with NSE value greater atleast by 1.2% and MAE lower atleast by 33% compared to existing works. The proposed solution is able to provide better prediction accuracy compared to multivariate LSTM without attention. Integration of IK with LSTM has been demonstrated in this work and results looks promising. But the IK rules considered in this work are limited. Testing against large number of IK rules and creating attention vector from IK rules with more named entities is in scope of future work.

References

[1] National Climatic Data Center(NCDC)(2006) Van Nostrand’s scientific Encyclopedia,10.1002/0471743984.vse8648

[2] Wilhite, D. A. (1990). The Enigma of Drought. Drought assessment, management, and planning: Theory andcase studies, vol. 2, pp. 3–15. Boston: Kluwer Academic Publishers

[3] Victor, M. P. (1999). Hydrological drought. http://ponce.sdsu.edu/..

[4] Sundararajan, Karpagam & Garg, Lalit & Srinivasan, Kathiravan & Bashir, Ali & Kaliappan, Jayakumar & Ganapathy, Ganapathy & Selvaraj, Senthil Kumaran & Thiruvadi, Meena. (2021). A Contemporary Review on Drought Modeling Using Machine Learning Approaches. Computer Modeling in Engineering and Sciences. 128. 447-487. 10.32604/cmes.2021.015528.

[5] Salite, Daniela. (2019). Traditional prediction of drought under weather and climate uncertainty: analyzing the challenges and opportunities for small-scale farmers in Gaza province, southern region of Mozambique. Natural Hazards. 96. 10.1007/s11069-019-03613-4.

[6] Jena, Pravat & Azad, Sarita. (2021). A novel method to detect drought and flood years in Indian rainfall associated with weak and strong monsoon. Theoretical and Applied Climatology. 145. 1-15. 10.1007/s00704-021-03652-7.

[7] Zhang, Linqi & Liu, Yi & Ren, Jun & Teuling, Adriaan & Zhu, Ye & Wei, Linyong & Zhang, Linyan & Jiang, Shanhu & Yang, Xiaoli & Fang, Xiuqin & Yin, Hang. (2021). Analysis of Flash Drought in China using Artificial Intelligence models. 10.5194/hess-2021-541.

[8] Xiang, Zhang & Obringer, Renee & Wei, Chehan & Chen, Nengcheng & Niyogi, Dev. (2017). Droughts in India from 1981 to 2013 and Implications to Wheat Production. Scientific Reports. 7. 44552. 10.1038/srep44552.

[9] Balehegn, M., Balehey, S., Fu, C. et al. Indigenous weather and climate forecasting knowledge among Afar pastoralists of north eastern Ethiopia: Role in adaptation to weather and climate variability. Pastoralism 9, 8 (2019).

[10] Owen, Mafongoya & Mafongoya, Paramu & Mudhara, Maxwell. (2021). Using Indigenous Knowledge Systems in Seasonal Prediction and Adapting to Climate Change Impacts in Bikita District in Zimbabwe. The Oriental Anthropologist: A Bi-annual International Journal of the Science of Man. 21. 0972558X2199766. 10.1177/0972558X21997662.

[11] Pareek, Aparna & Trivedi, P.. (2011). Cultural values and indigenous knowledge of climate change and disaster prediction in Rajasthan, India. Indian Journal of Traditional Knowledge. 10. 183-189.

[12] Mugi-Ngenga EW, Kiboi MN, Mucheru-Muna MW, Mugwe JN, Mairura FS, Mugendi DN, Ngetich FN (2021). Indigenous and conventional climate-knowledge for enhanced farmers’ adaptation to climate variability in the semi-arid agro-ecologies of Kenya. Environmental Challenges 5 (2021) 100355

[13] Johnston, Peter & Olagbegi, Oladapo & Simpson, Nicholas & Trisos, Christopher & Zvobgo, Luckson. (2022). Role of Indigenous and local knowledge in seasonal forecasts and climate risk preparedness: a case study of smallholder farmers in Chiredzi, Zimbabwe. 10.21203/rs.3.rs-1436068/v1.

[14] Islam, Md. (2017). Local knowledge and practices can help in drought prediction and extreme weather management. Indian journal of traditional knowledge. 16.

[15] Zhang, N., Dai, X., Ehsan, M.A., Deksissa, T. (2020). Development of a Drought Prediction System Based on Long Short-Term Memory Networks (LSTM). In: Han, M., Qin, S., Zhang, N. (eds) Advances in Neural Networks – ISNN 2020. ISNN 2020. Lecture Notes in Computer Science(), vol 12557. Springer, Cham

[16] Selvaraj, Poornima & Marudappa, Pushpalatha. (2019). Drought prediction based on SPI and SPEI with varying timescales using LSTM recurrent neural network. Soft Computing. 23. 1-14. 10.1007/s00500-019-04120-1.

[17] Park, Sumin & Im, Jungho & Han, Daehyeon & Rhee, Jinyoung. (2020). Short-Term Forecasting of Satellite-Based Drought Indices Using Their Temporal Patterns and Numerical Model Output. Remote Sensing. 10.3390/rs12213499.

[18] Xu, D., Zhang, Q., Ding, Y. et al. Application of a hybrid ARIMA-LSTM model based on the SPEI for drought forecasting. Environ Sci Pollut Res 29, 4128–4144 (2022).

[19] Band, Shahab & Karami, Hojat & Jeong, Yong-Wook & Moslemzadeh, Mohsen & Farzin, Saeed & Bateni, Sayed & Mosavi, Amir. (2022). Evaluation of Time Series Models in Simulating Different Monthly Scales of Drought Index for Improving Their Forecast Accuracy. Frontiers in Earth Science. 10. 10.3389/feart.2022.839527.

[20] Maity, Rajib & Khan, Mohd & Sarkar, Subharthi & Dutta, Riya & Maity, Subhra Sekhar & Pal, Manali & Chanda, Kironmala. (2021). Potential of Deep Learning in drought assessment by extracting information from hydrometeorological precursors. Journal of Water and Climate Change. 12. 10.2166/wcc.2021.062.

[21] Dikshit, Abhirup & Pradhan, Biswajeet & Alamri, Abdullah. (2021). Long lead time drought forecasting using lagged climate variables and a stacked long short-term memory model. Science of The Total Environment. 755. 142638. 10.1016/j.scitotenv.2020.142638.

[22] A. Mokhtar *et al*., "Estimation of SPEI Meteorological Drought Using Machine Learning Algorithms," in *IEEE Access*, vol. 9, pp. 65503-65523, 2021

[23] Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L.; and Polosukhin, I. 2017. Attention is all you need.

[24] Biradar, Nagaratna & Kerur, Akshata & Tirlapur, Laxmi & Chand, Khem & Anjinappa, Bheemappa. (2021). Analysis of precipitation trend and farmers' practices to predict precipitation in North- Karnataka. Journal of Pharmacognosy and Phytochemistry. 10. 406-411.

[25] R. Ghanbari and K. Borna, "Multivariate Time-Series Prediction Using LSTM Neural Networks," *2021 26th International Computer Conference, Computer Society of Iran (CSICC)*, 2021, pp. 1-5

[26] Achite M, Bazrafshan O, Azhdari Z, Wałęga A, Krakauer N, Caloiero T. Forecasting of SPI and SRI Using Multiplicative ARIMA under Climate Variability in a Mediterranean Region: Wadi Ouahrane Basin, Algeria. *Climate*. 2022; 10(3):36,

[27] Zhang, Xuan & Liang, Xun & Li, Aakas & Zhang, Shusen & Xu, Rui & Wu, Bo. (2019). AT-LSTM: An Attention-based LSTM Model for Financial Time Series Prediction. IOP Conference Series: Materials Science and Engineering. 569. 052037. 10.1088/1757-899X/569/5/052037.

[28] Perfiljeva, Irina. (2006). Fuzzy transforms: Theory and applications. Fuzzy Sets and Systems. 157. 993-1023. 10.1016/j.fss.2005.11.012.

[29] A. Sangari and W. Sethares, "Convergence Analysis of Two Loss Functions in Soft-Max Regression," in *IEEE Transactions on Signal Processing*, vol. 64, no. 5, pp. 1280-1288, March1, 2016

[30] Nash, J. E.; Sutcliffe, J. V. (1970). "River flow forecasting through conceptual models part I — A discussion of principles". Journal of Hydrology. 10 (3): 282–290

[31] Roy, A.K. and Hirway, I. 2007. Multiple Impacts of Droughts and Assessment of Drought Policy in Major Drought Prone States in India. Project Report Submitted to the Planning Commission, Government of India, New Delhi

[32] Palmer, W. C. 1965 Meteorological Drought. Res. Paper No. 45.Weather Bureau, Washington, DC.

[33] McKee, T. B., Doesken, N. J. & Kleist, J.1993 The relationship of drought frequency and duration to time scales. In: Proceedings of the 8th Conference on Applied Climatology, pp. 179–183.

[34] Russo, S., Dosio, A., Sterl, A., Barbosa, P. & Vogt, J.2013 Projection of occurrence of extreme dry-wet years and seasons in Europe with stationary and nonstationary Standardized Precipitation Indices. J. Geophys. Res.

Atmos.118, 7628–7639.

[35] Kao, S. C. & Govindaraju, R. S.2010 A copula based joint deficit index for droughts. J. Hydrol. 380, 121–134.

[36] Won, J., Choi, J., Lee, O. & Kim, S. 2020 Copula-based joint drought index using SPI and EDDI and its application to climate change. Sci. Total Environ. 744, 140701.

[37] Amrit, K., Pandey, R. P. & Mishra, S. K. 2018 Characteristics of meteorological droughts in northwestern India. Nat. Hazards 94, 561–582

[38] Das, P. & Chanda, K. 2020 Bayesian Network based modeling of regional rainfall from multiple local meteorological drivers. J. Hydrol. 591, 125563

[39] Rehana, S. & Monish, N. T. 2020 Characterization of regional drought over water and energy limited zones of India using potential and actual evapotranspiration. Earth Sp. Sci. 7.