E-ISSN: 2618-0618
P-ISSN: 2618-060X

www.agronomyjournals.com
2024; SP-7(6): 387-392
Received: 11-03-2024
Accepted: 20-04-2024

Vikas Kumar Singh
Assistant Professor, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Shivam
Assistant Professor, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Mo Akram
M.Tech Students, Department of SWCE, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Sarvda Nand Tiwari
Ph.D. Scholar, Department of SWCE, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Ankit
Ph.D. Scholars, Department of SWCE, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Vipin Kumar Roshan
Ph.D. Scholars, Department of SWCE, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Akanksha Mathur
M. Tech students, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Khwahiz Ali
M. Tech students, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Sakshi Dixit
M. Tech students, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Corresponding Author:
Vikas Kumar Singh
Assistant Professor, MCAET, ANDUAT, Ayodhya, Uttar Pradesh, India

Comprehensive drought risk assessment over the central plain zone of Uttar Pradesh using machine learning approach

Vikas Kumar Singh, Shivam, Mo Akram, Sarvda Nand Tiwari, Ankit, Vipin Kumar Roshan, Akanksha Mathur, Khwahiz Ali and Sakshi Dixit

DOI: https://doi.org/10.33545/2618060X.2024.v7.i6Sf.882

Abstract
This study aimed to assess patterns of drought vulnerability over the central plain zone of Uttar Pradesh, India, using the Standardized Precipitation-Evapotranspiration Index (SPEI). The analysis was conducted for the period of 1981-2020 to identify regions within the central plain zone that exhibit high drought vulnerability. The SPEI analysis revealed the central plain zone was the most susceptible to drought conditions. The data indicates that there were multiple periods of severe drought, with SPEI values below -2, signifying severe drought conditions, in 1987, 1991, 2006, and 2015. Number of years had positive SPEI readings, including 1981, 1985, 1990, 1994, and 1996, indicating times of high precipitation and ideal moisture levels. Machine Learning Regression has developed a comprehensive drought risk assessment model. The values of R^2, RSME and MAE for training and testing set are 0.9917, 0.0898 & 0.0676 and 0.9744, 0.2068 & 0.1368, respectively.

Keywords: Comprehensive, risk, plain, machine, learning, approach

Introduction

Droughts are characterized by protracted, abnormally dry spells of weather that last long enough for the lack of precipitation to drastically reduce moisture content and cause a hydrological imbalance (Mishra & Singh, 2010) [4]. They can also be sustained to the point where there is insufficient water for any particular activity. For example, various experts characterize a drought in different ways (Paulo & Pereira, 2006) [7]. It can be described by a variety of experts: a meteorologist as below-average rainfall; an agriculturist as a lack of moisture in the root zone (Nagarajan, 2010) [5] a hydrologist as below-average water levels in reservoirs, lakes, and other similar areas; an economist as a shortage of water that negatively impacts the established economy (Panu & Sharma, 2002) [6]. Drought can be categorized as a purely meteorological phenomenon (Lloyd-Hughes, 2014) [3]. A meteorological drought is the earliest and most obvious occurrence in the onset and development of drought conditions. Droughts are a common natural phenomenon in India, with regional variations in length and intensity (Kumar *et al.*, 2013) [2]. India's diverse climate, which varies from humid tropical parts to arid and semi-arid regions, makes it vulnerable to droughts (Surendran *et al.*, 2019) [8]. India's drought is caused by a number of factors, including deforestation, unsustainable groundwater exploitation, erratic monsoons, and the consequences of climate change. Droughts can have severe consequences, including crop failures, animal losses, water shortages, and financial hardship, particularly for farmers and the underprivileged (Gautier *et al.*, 2016) [1]. Multiple linear regressions are a statistical technique used to model the relationship between a dependent variable and two or more independent variables (Tranmer & Elliot, 2008) [9]. It is a powerful tool for understanding and predicting complex phenomena, as it allows for the consideration of multiple factors simultaneously.

Study Area

The central plain zone of Uttar Pradesh is located in Uttar Pradesh, a state in northern India. Situated in the center of the Indo-Gangetic Plain, central plain zone of Uttar Pradesh physical location has greatly influenced the region's history and development.

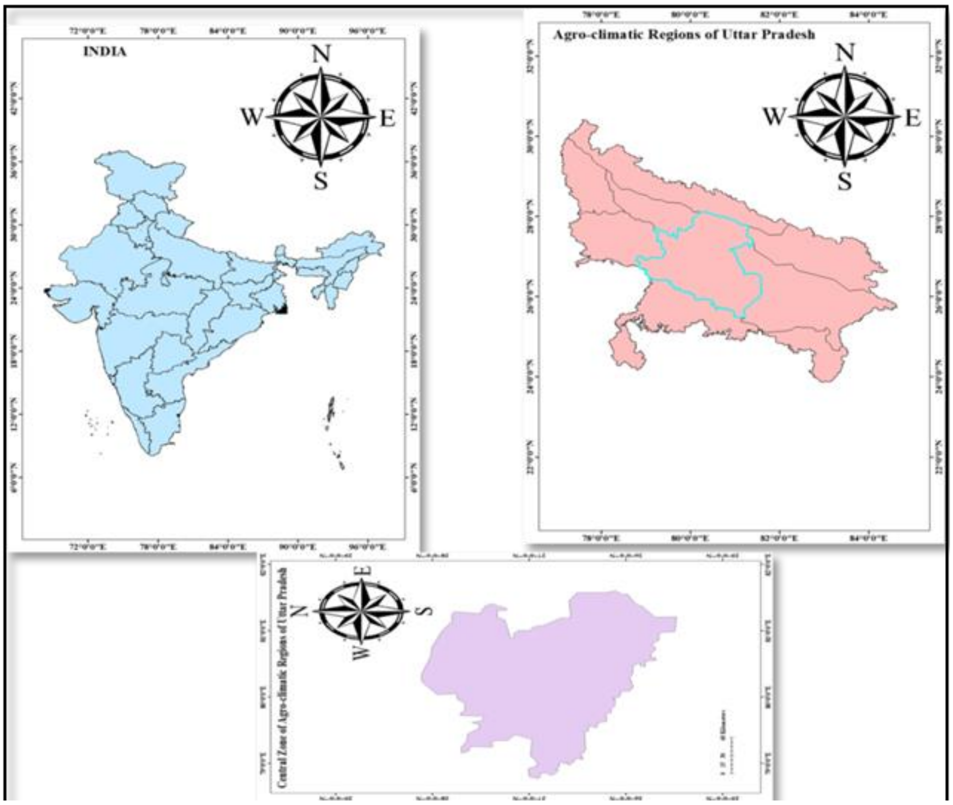

Fig 1: Central Plain Zone of Uttar Pradesh

Central plain zone of Uttar Pradesh is a significant agricultural center because of the region's alluvial soil, which has encouraged the growth of a variety of crops. The engineering, leather, and textile sectors of the region have been major drivers of economic expansion. In addition to its industrial and economic strength, central plain zone of Uttar Pradesh has had to overcome a number of social and developmental obstacles.

Data Collection

Annual rainfall data for central plain zone of Uttar Pradesh was obtained from Indian Meteorological Department Pune from 1981-2020.

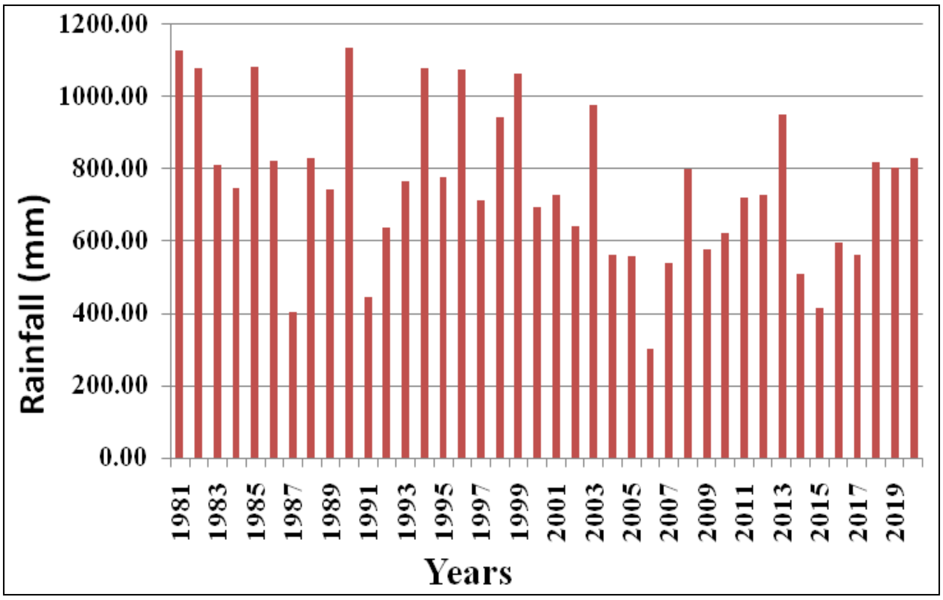

Fig 2: Annual rainfall data of central plain zone of Uttar Pradesh

This dataset provides annual rainfall data for a location over a 40-year period from 1981 to 2020. The rainfall values range from a low of 299.26 mm in 2006 to a high of 1,130.81 mm in 1990 indicating significant variability in the annual precipitation at central plain zone of Uttar Pradesh location.

Materials and Methods

Standardized Precipitation Evapotranspiration Index (SPEI): Standardized Precipitation Evapotranspiration Index (SPEI) was used in the current study to quantify the deficiency of precipitation in various time scales. These periods cover both transient and persistent abnormalities in precipitation. While long-term anomalies are used for groundwater, stream flow, and reservoir storage studies, short-term anomalies are typically used in soil moisture investigations (Beguería *et al.*, 2014) [10]. All categories of SPEI value are mention in Table 1. The SPEI is calculated as the difference between precipitation (P) and potential evapotranspiration (PET), divided by the standard deviation (σ) of this difference for a given time scale.
Equation 1 is used to calculate Standardized Precipitation Evapotranspiration Index:

$$\text{SPEI} = \frac{P-PET}{\sigma}$$

Eq^n… 1
Where,
P = Precipitation if current months
PET = Potential Evapotranspiration
σ = Standard deviation

Table 1: Categorization of SPEI value

Sr. No.	Category	SPEI range
1	Extremely wet condition	2 or more
2	Severely wet condition	1.5 to 1.99
3	Moderately wet condition	1 to 1.49
4	Mildly wet condition	0 to 0.99
5	Mildly dry condition	0 to -0.99
6	Moderately dry condition	-1 to -1.49
7	Severely dry condition	-1.5 to -1.99
8	Extremely dry condition	-2 or less

Modelling of drought using machine learning approach (multiple learning regression)

A statistical method based on machine learning approach for simulating the relationship between a dependent variable and two or more independent variables is called multiple linear regressions. At its core, multiple linear regression aims to find the best-fitting linear equation that describes the relationship between the dependent variable and the independent variables.

The general form of the multiple linear regression equation is

$$Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + ... + \beta_nX_n + \varepsilon,$$

Where,
Y is the dependent variable, $X_1, X_2, ..., X_n$ are the independent variables, β_0 is the y-intercept, $\beta_1, \beta_2, ..., \beta_n$ are the regression coefficients that represent the change in the dependent variable for a unit change in the corresponding independent variable, and ε is the error term that accounts for the unexplained variation in the dependent variable Additionally, multiple linear regressions provides measures of the overall model fit, such as the coefficient of determination (R^2), which indicates the proportion of the variance in the dependent variable that is explained by the independent variables (Yang *et al.*, 2020) [11].

Results and Discussion

This dataset provides the Standardized Precipitation-Evapotranspiration Index (SPEI) values for the same location over the same 40-year period from 1981 to 2020.

Analysis of SPEI data

The SPEI values shown in Figure 2 and dataset range from a minimum of -3.38 in 2006 to a maximum of 1.21 in 1990, indicating a wide variation in the drought conditions experienced at this location over the years.

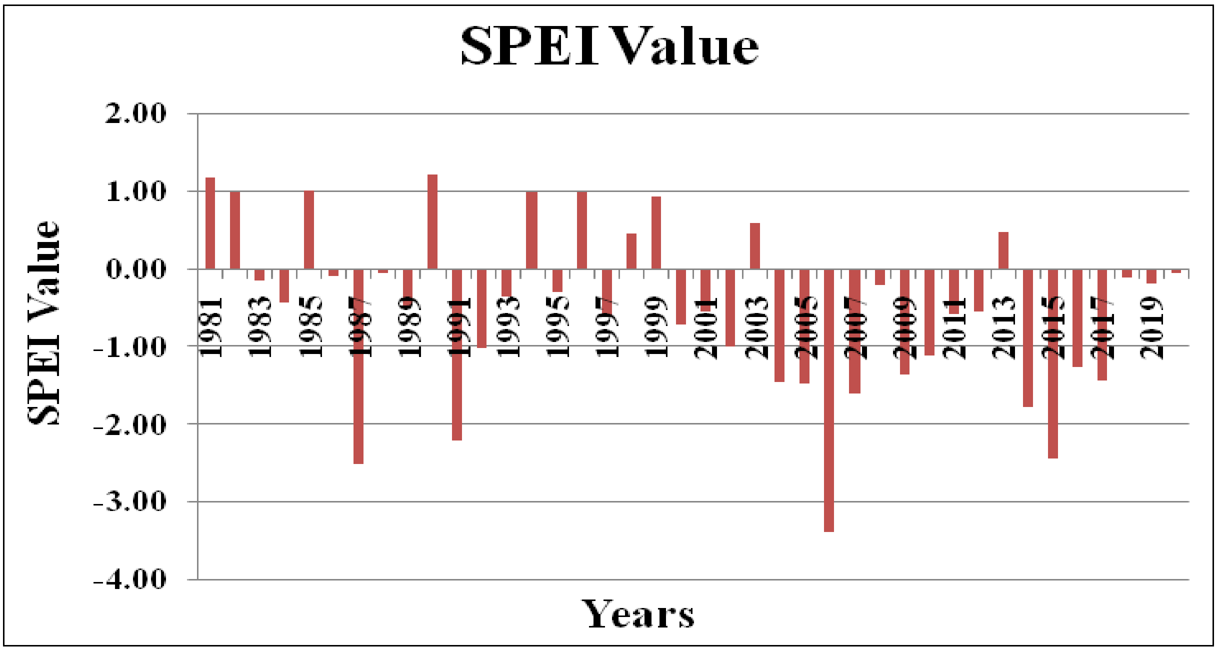

Fig 2: SPEI Drought Indices

Negative SPEI values correspond to drier-than-normal conditions, while positive values indicate wetter-than-normal conditions.
The data shows several periods of significant drought, such as in 1987, 1991, 2006, and 2015, where the SPEI values were below -2, indicating severe drought conditions. Conversely, there were also several years with positive SPEI values, such as 1981, 1985, 1990, 1994, and 1996, suggesting periods of abundant precipitation and favorable moisture conditions. The fluctuations in the SPEI values over time highlight the dynamic nature of the

local climate and the importance of considering both precipitation and evaporative demand when assessing drought risk and water resource management in the region.

Drought condition analysis: The Table 2 presents the Standardized Precipitation Evapotranspiration Index (SPEI) values and corresponding drought conditions for the years 1981 to 2020. The SPEI is used to determine the intensity and duration of droughts, with positive values indicating normal or non-drought (ND) conditions and negative values indicating varying levels of drought severity. From 1981 to 1990, the data indicates a mixture of normal and mild drought conditions, with notable normal conditions (SPEI = 1.19) in 1981, 1982 (SPEI = 1.00), 1985 (SPEI = 1.01), and 1990 (SPEI = 1.21). Severe drought conditions (SPEI = -2.51) were observed in 1987, while 1986 and 1988 to 1989 experienced mild droughts (SPEI = -0.09, -0.04, -0.46, respectively). In the 1990, there were fluctuating conditions with a severe drought (SPEI = -2.21) in 1991 and a moderate drought (SPEI = -1.02) in 1992. However, years like 1994 and 1996 experienced normal conditions (SPEI = 1.00, 0.99), while 1997 (SPEI = -0.61) and 1999 (SPEI = 0.94) indicated mild and normal conditions, respectively. This decade highlights a relatively stable climatic condition with occasional droughts that did not persist over multiple years. This period also includes years with normal conditions, suggesting episodic drought events rather than prolonged dry periods. The early 2000 were marked by significant drought conditions, with moderate droughts in 2002 (SPEI = -0.99), 2004 (SPEI = -1.44), 2005 (SPEI = -1.46), and 2010 (SPEI = -1.10). An extreme drought occurred in 2006 (SPEI = -3.38) and a severe drought in 2007 (SPEI = -1.59). However, 2003 experienced normal conditions (SPEI = 0.60). From 2011 to 2020, there were periods of moderate and severe droughts. 2011 (SPEI = -0.58), 2012 (SPEI = -0.53), 2016 (SPEI = -1.25), and 2017 (SPEI = -1.43) were marked by moderate droughts. An extreme drought was recorded in 2015 (SPEI = -2.43). In contrast, normal conditions were observed in 2013 (SPEI = 0.48) and 2019 (SPEI = 0.46). Overall, the data reveals significant variability in drought conditions over the four decades, with periods of both severe and mild droughts interspersed with normal conditions.

Table 2: SPEI based drought condition analysis for central zone of Uttar Pradesh

Years	SPEI Value	Condition	Years	SPEI Value	Condition
1981	1.19	ND	2001	-0.54	MD
1982	1.00	ND	2002	-0.99	MD
1983	-0.13	MD	2003	0.60	ND
1984	-0.43	MD	2004	-1.44	MOD
1985	1.01	ND	2005	-1.46	MOD
1986	-0.09	MD	2006	-3.38	ED
1987	-2.51	ED	2007	-1.59	SD
1988	-0.04	MD	2008	-0.19	MD
1989	-0.46	MD	2009	-1.36	MOD
1990	1.21	ND	2010	-1.10	MOD
1991	-2.21	ED	2011	-0.58	MD
1992	-1.02	MOD	2012	-0.53	MD
1993	-0.35	MD	2013	0.48	ND
1994	1.00	ND	2014	-1.77	SD
1995	-0.29	MD	2015	-2.43	ED
1996	0.99	ND	2016	-1.25	MOD
1997	-0.61	MD	2017	-1.43	MOD
1998	0.46	ND	2018	-0.10	MD
1999	0.94	ND	2019	-0.18	MD
2000	-0.70	MD	2020	-0.05	MD

Multiple learning Regression (MLR)

Table 3: Model Performance Evaluation Parameter

Model Performance Training Set		
R^2	RSME	MAE
0.9917	0.0898	0.0676
Model Performance Testing Set		
R^2	RSME	MAE
0.9744	0.2068	0.1394

Training Set Performance: R-squared (R^2) value of 0.9917 indicates that the model explains 99.17% of the variance in the training data, which is an exceptionally high goodness-of-fit. Root Mean Squared Error (RMSE) of 0.0898 suggests the model has low error in predicting the training set values. Mean Absolute Error (MAE) of 0.0676 further confirms the high accuracy of the model's predictions on the training data.

Testing Set Performance: The R^2 value of 0.9744 on the testing set is slightly lower than the training set, but still indicates the model explains 97.44% of the variance in the unseen testing data. The RMSE of 0.2068 and MAE of 0.1394 on the testing set are higher than the training set, but still relatively low, indicating good generalization performance. The high R-squared values low RMSE, and low MAE on both the training and testing sets suggest that the model has achieved excellent fit and predictive performance shown in Table 3. The slightly lower, but still strong, performance on the testing set compared to the training set is expected and indicates the model is able to generalize well to new, unseen data.

Modelling Graph of Training Set: This figure 3a and 3b presents a graphical comparison between simulated data (shown in red) and drought conditions (shown in blue) based on the number of SPEI (Standardized Precipitation-Evapotranspiration Index) values. The SPEI is a widely used drought index that combines precipitation and temperature data to assess the intensity of dry and wet conditions.

The red line represents the simulated data, indicating the hypothetical or modeled SPEI values, while the blue line represents the actual drought conditions observed. The observed and simulated SPEI data ranges from 1.19 & -2.51 and 1.38 & -

2.29. The graph illustrates the fluctuations in SPEI values over time, with both the simulated data and the observed drought conditions exhibiting periods of positive and negative SPEI values. Positive SPEI values indicate wetter than average conditions, while negative SPEI values indicate drier than average conditions.

The graph allows for a visual comparison between the simulated data and the observed drought conditions, highlighting the similarities and differences in the patterns and magnitudes of the SPEI values. This type of data visualization can be useful for understanding the performance of drought simulation models, as well as identifying periods of drought and wet conditions in a given region or time period.

Modelling Graph of Testing Set: A visual comparison of simulated and observed drought levels for a range of SPEI (Standardized Precipitation Evapotranspiration Index) values is shown in Figure 4a and 4b. SPEI, a commonly used indicator of drought conditions, has negative values that reflect conditions that are drier than usual. The simulated data points are represented by the red line, and the observed drought values are represented by the blue line.

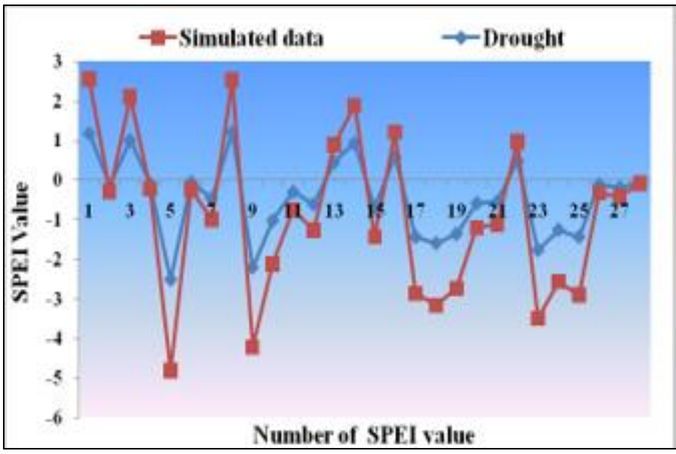

Fig 3a: Line graph of training dataset

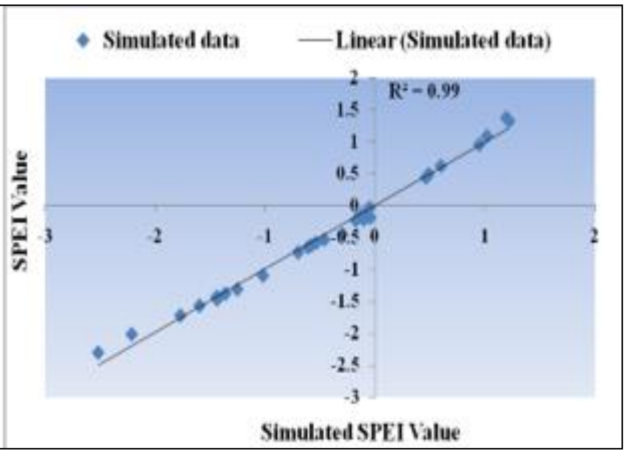

Fig 3b: Line graph of training dataset

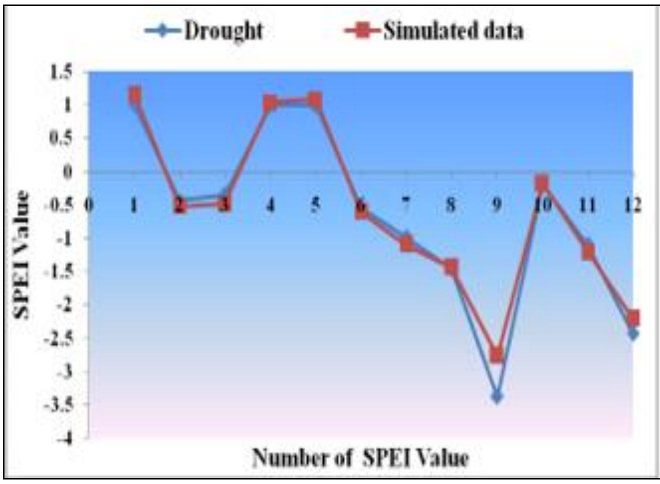

Fig 4a: Line graph of training dataset

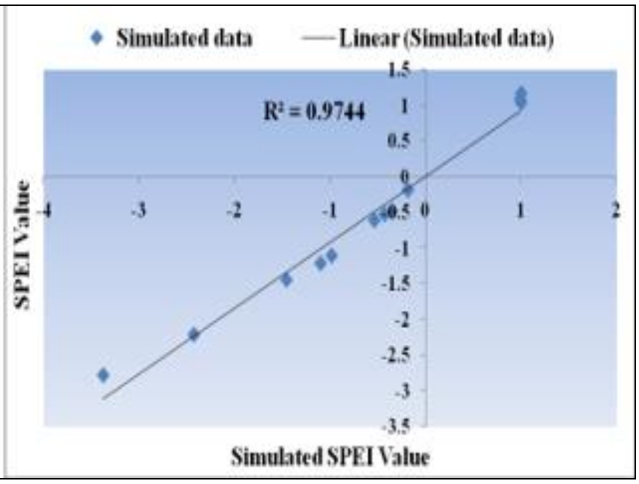

Fig 4b: Line graph of training dataset

The observed and simulated SPEI data ranges from 0.99 & -3.37 and 1.17 & -2.77 shown in Fig 4a and 4b. The plot shows that, overall, the simulated data points and the observed drought levels follow a similar trend, with peaks and troughs appearing at different times. This shows that the general patterns and dynamics of the reported drought conditions are well captured by the simulated data.

Conclusions

The study utilized the Standardized Precipitation-Evapotranspiration Index (SPEI) to analyze the patterns of drought vulnerability in the central plain zone of Uttar Pradesh, India over the period of 1981-2020. The SPEI analysis identified several regions within the central plain zone that exhibit high drought vulnerability, with the some parts being the most susceptible have no drought. Machine learning techniques (multiple learning regressions), were employed to develop a comprehensive drought risk assessment model that integrates various biophysical and socioeconomic factors. The drought risk assessment model demonstrated high accuracy in predicting drought vulnerability, providing valuable insights for targeted drought mitigation and adaptation strategies. The findings of this study can inform policymakers and resource managers in developing effective drought management plans, focusing on the most vulnerable regions and prioritizing appropriate interventions.

References

1. Gautier D, Denis D, Locatelli B. Impacts of drought and responses of rural populations in West Africa: a systematic review. Wiley Interdisciplinary Reviews: Climate Change. 2016;7(5):666-681.
2. Kumar KN, Rajeevan M, Pai DS, Srivastava AK, Preethi B. On the observed variability of monsoon droughts over

India. Weather and Climate Extremes. 2013;1:42-50.
3. Lloyd-Hughes B. The impracticality of a universal drought definition. Theoretical and Applied Climatology. 2014;117:607-611.
4. Mishra AK, Singh VP. A review of drought concepts. Journal of Hydrology. 2010;391(1-2):202-216.
5. Nagarajan R. Drought assessment. Springer Science & Business Media. 2010.
6. Panu US, Sharma TC. Challenges in drought research: some perspectives and future directions. Hydrological Sciences Journal. 2002;47(S1).
7. Paulo AA, Pereira LS. Drought concepts and characterization: Comparing drought indices applied at local and regional scales. Water International. 2006;31(1):37-49.
8. Surendran U, Anagha B, Raja P, Kumar V, Rajan K, Jayakumar M. Analysis of drought from humid, semi-arid and arid regions of India using DrinC model with different drought indices. Water Resources Management. 2019;33:1521-1540.
9. Tranmer M, Elliot M. Multiple linear regression. The Cathie Marsh Centre for Census and Survey Research (CCSR). 2008;5(5):1-5.
10. Beguería S, Vicente-Serrano SM, Reig F, Latorre B. Standardized precipitation evapotranspiration index (*SPEI*) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. International Journal of Climatology. 2014;34(10):3001-3023.
11. Yang X, Wu W, Lu L, Yan B, Zhang L, Liu K. Multiple regressions based image super-resolution. Multimedia Tools and Applications. 2020;79:8911-8927.