Article

Application of Machine Learning and Hydrological Models for Drought Evaluation in Ungauged Basins Using Satellite-Derived Precipitation Data

Anjan Parajuli [1], Ranjan Parajuli [2], Mandip Banjara [3], Amrit Bhusal [4] , Dewasis Dahal [5] and Ajay Kalra [5,*]

1   AECOM, 6000 Fairview Road, Suite 200, Charlotte, NC 28210, USA; anjan.parajuli@aecom.com
2   Kleinfelder, 707 17th Street, Suite 3000, Denver, CO 80202, USA; rparajuli@kleinfelder.com
3   Stantec, 601 Grassmere Park, Suite 22, Nashville, TN 37211, USA; mandip.banjara@stantec.com
4   Arcadis U.S., Inc., 7575 Huntington Park Dr. Suite 130, Columbus, OH 43235, USA;
    amrit.bhusal@arcadis-us.com
5   School of Civil, Environmental, and Infrastructure Engineering, Southern Illinois University,
    1230 Lincoln Drive, Carbondale, IL 62901, USA; dewasis.dahal@siu.edu
*   Correspondence: kalraa@siu.edu; Tel.: +1-(618)-453-7008

Abstract: Drought is a complex environmental hazard to ecosystems and society. Decision-making on drought management options requires evaluating and predicting the extremity of future drought events. In this regard, quantifiable indices such as the standardized precipitation index (SPI), the standardized precipitation evapotranspiration index (SPEI), and the standardized streamflow index (SSI) have been commonly used to characterize meteorological and hydrological drought. In general, the estimation and prediction of the indices require an extensive range of precipitation (SPI and SPEI) and discharge (SSI) datasets in space and time domains. However, there is a challenge for long-term and spatially extensive data availability, leading to the insufficiency of data in estimating drought indices. In this regard, this study uses satellite precipitation data to estimate and predict the drought indices. SPI values were calculated from the precipitation data obtained from the Centre for Hydrometeorology and Remote Sensing (CHRS) data portal for a study water basin. This study employs a hydrological model for calculating discharge and drought in the overall basin. It uses random forest (RF) and support vector regression (SVR) as machine learning models for SSI prediction for time scales of 1- and 3-month periods, which are widely used for establishing interactions between predictors and predictands that are both linear and non-linear. This study aims to evaluate drought severity variation in the overall basin using the hydrological model and compare this result with the machine learning model's results. The results from the prediction model, hydrological model, and the station data show better correlation. The coefficients of determination obtained for 1-month SSI are 0.842 and 0.696, and those for the 3-month SSI are 0.919 and 0.862 in the RF and SVR models, respectively. These results also revealed more precise predictions of machine learning models in the longer duration as compared to the shorter one, with the better prediction result being from the SVR model. The hydrological model-evaluated SSI has 0.885 and 0.826 coefficients of determination for the 1- and 3-month time durations, respectively. The results and discussion in this research will aid planners and decision-makers in managing hydrological droughts in basins.

Keywords: drought prediction; HEC-HMS; random forest; support vector regression; standardized streamflow index; standardized precipitation index

1. Introduction

Drought is an extreme climatic phenomenon that develops slowly but increases in intensity and frequency, causing significant consequences [1]. The drought in Texas in 2011, the U.S. Central Great Plains drought in 2012, and the California drought have caused extreme losses to society, agriculture, and ecosystems, significantly impacting water supply

and crop production [2,3]. As per the United States Department of Agriculture (USDA) Report in 2017, nearly 800 million people were found to be malnourished, with most of this population coming from developing countries, and the primary cause was drought [4]. Drought's known effects are decreased water supply, poor water quality, environmental hazards, forest fires, crop failure, decreased productivity, disturbed riparian habitats, reduced power generation, famine, civil unrest, and suspension of recreational activities. Drought frequency and severity will increase due to climate change (e.g., precipitation pattern change) and socioeconomic changes (e.g., increasing water demand) [1]. As a result, accurate drought prediction based on real-time drought monitoring is necessary for early warning, preparation, and mitigation to lessen its negative impacts [5].

Meanwhile, climate change and temperature rises, which affect drought frequency and severity, are expected to continue [6], making future droughts more severe and prolonged. Thus, evaluating and predicting droughts is critical for early warning and preparing the most vulnerable areas to mitigate drought impacts. Drought prediction at various time scales is required to develop management strategies to minimize negative societal and economic impacts [7]. Previous studies have classified drought into multiple types to characterize and address drought events: meteorological drought, hydrological drought, agricultural drought, and socioeconomic drought [8]. Usually, when there is a shortage of precipitation, it leads to a decrease in water availability. This is known as a meteorological drought, which can result in hydrological drought due to streamflow reduction and invariably spreads over time through a hydrological cycle, causing reduced soil moisture and agricultural drought. Socioeconomic drought occurs when the water supply does not satisfy society's productive and consumptive activities. In addition to the four types of drought, many studies have defined other types, such as groundwater drought, environmental drought, and ecological drought, depending on their target systems [9–11].

Rainfall data are crucial for assessing drought. Observations from conventional precipitation gauges are regarded as the most accurate source of precipitation data for the most part. However, satellite-based precipitation data for drought monitoring are preferred due to the high installation and maintenance cost, short-length data records,, erratic observations, and limited spatial measurements of precipitation gauges [12]. These shortcomings can be resolved using the precipitation data from satellites. In previous drought forecasting studies, researchers used meteorological data based on rain gauge ground data, which were found to be missing in most areas, making drought forecasting complex. Satellite precipitation data are derived from sensors that use infrared. Researchers utilized satellite precipitation data from Climate Hazards Group Infrared Precipitation (CHIRPS) products, available at spatial resolutions of 0.25 and 0.05 degrees, to monitor South Asian droughts [13]. Tote et al. [14] used CHIRPS precipitation data to monitor droughts and floods in Mozambique. Shukla et al. [15] used CHIRPS precipitation data to forecast droughts in East Africa.

Several researchers have used different machine-learning tools for drought prediction recently; machine learning (ML) models such as artificial neural networks (ANNs), random forest (RF), long short-term memory (LSTM), and support vector machines (SVMs) have been used for drought prediction [16–19]. ML models can improve prediction accuracy without being explicitly designed for the purpose [20]. With the large variety of drought prediction models available, it can be arduous for researchers to decide which is the best model for drought prediction research if they are unaware of the available prediction models. Most models have more limited performance with short-term forecasting than with long-term forecasting [21]. This is because the hydrological processes related to short-term drought are complex, including factors such as moisture, temperature, and evapotranspiration after precipitation.

There has been less research on using the RF model in drought prediction [22]. Moreover, RF has not been used for drought prediction near the Folsom Lake basin area. In the study by Chen et al. [23], ARIMA and RF models were used in China's Haihe River basin to predict drought. The study included precipitation, temperature, evaporation, and

surface water content as input variables. The RF model can forecast reliable results without adjusting its parameters, is flexible in capturing time series fundamental relationships, and can generate an ensemble of drought forecasts rather than a mean prediction. According to the study, the model based on RF was found to be more reliable than ARIMA in forecasting droughts, both in the short and long term. Dikshit et al. [24] used RF for drought forecasting in New South Wales, Australia. SPEI3 was found to have a better coefficient of determination (R^2) than SPEI1 in the validation period. Rainfall, minimum, maximum, and mean temperatures, potential evapotranspiration (PET), vapor pressure, and cloud cover were the climatic variables used in the prediction model. Park et al. [25] utilized three machine learning approaches: RF, boosted regression trees, and Cubist for drought prediction in two arid (48 counties of Arizona and New Mexico) and two humid regions (146 counties of North and South Carolina). RF was found to perform best for SPI prediction. Similarly, Park et al. [26] predicted severe drought using satellite images and topography data based on RF in western Korea.

Similarly, among the various ML techniques, SVR can be regarded as one of the most used drought prediction models [5]. The SVR model changes the non-linear relationship between predictor and predictand to linear, where inputs are mapped into higher dimensional space. Mapping is conducted in SVR using the kernel function. SVR can learn from a small set of datasets and has the ability to handle datasets. Borji et al. [27] used SVR and ANNs for hydrological drought forecasting. The study showed better efficiency for the SVR model in long-term drought forecasting than the ANN. Achite et al. [28] found that an SVM produced better results out of four machine learning models used—which were an ANN, an artificial neuro-fuzzy inference system (ANFIS), an SVM, and a decision tree (DT)—in hydrological drought forecasting in the Wadi Ouahrane basin in the northern part of Algeria. A hundred percent accuracy was found in the SVM model for hydrological drought forecasting in the Gidra River [29].

The above studies have the limitation of not comparing the hydrological drought indices resulting from HEC-HMS discharge output, RF model prediction, SVR model prediction, and station gauge data for short-term drought in a basin using drought indices SSI. This research aims to fill this gap by comparing their results with those of RF and SVR for the short term (1 and 3 months). Moreover, the satellite precipitation data from the CHRS data portal are used in our prediction model for hydrological drought indices calculation (SSI), and the process-based hydrological modeling tool (HEC-HMS) is used for discharge calculation, which is needed in hydrological drought indices (i.e., SSI) calculation. The novelty of our research is the use of machine learning and hydrological models for future drought prediction and evaluation, respectively.

This study investigates the feasibility of using satellite precipitation data for evaluating and predicting drought severity. The SSI is used to quantify the severity of hydrological drought, and the HEC-HMS is used to model streamflow. Previous studies have used streamflow from only the gauge stations and measured drought indices. However, in this study, the station gauge data are used, a basin is created, and discharge is calculated at the outlet using the HEC-HMS. The station discharge and HEC-HMS discharge are used for SSI calculation. RF and SVR models are applied using time-lagged SPI, SSI, and climatic data for future hydrological drought prediction (i.e., SSI prediction), indicating the study area's drought severity. The research shows good predictability of hydrological drought (SSI) in a basin using RF and SVR models for short-term drought.

2. Materials and Methods

2.1. Study Area

The evaluation of hydrological drought was applied to the Folsom Lake basin in Northern California, which recently experienced extreme drought from 2012 to 2016 [30]. The drought monitoring study carried out by Mozgovoy [31] with the use of satellite and ground data concluded that there has been desiccation of Folsom Lake during 2011–2015. This area requires special attention and the basin contributing to the runoff towards Folsom

Lake was chosen as a site for this research. The study area is 883 km^2, and the central part of the basin is located at 38.93° latitude and -121.02° longitude. Figure 1 illustrates the location of the designated study area and the gauge station downstream. The study area generally has a Mediterranean climate, with hot, dry summers and cool, wet winters. According to the Western Regional Climate Centre [32], average annual temperatures in the Folsom Lake area range from a low of around 10 degrees Celsius in winter to a high of around 30 degrees Celsius in summer. Maximum temperatures can exceed 38 degrees Celsius during heat waves in the summer, while minimum temperatures can drop below freezing in the winter. Annual precipitation varies significantly, with some years experiencing above-average rainfall and others experiencing drought conditions. On average, the area receives around 500–800 mm of precipitation annually, most falling between October and April. Recently, at the beginning of 2023, intense storms caused flooding, power outages, and landslides in the area. The opposite of the flooding was the extremely low precipitation experienced in 2015 and 2016. Climate change has altered atmospheric patterns, which can also be seen in this area. One extremely dry and one extremely wet about every two decades is experienced in the southern California region [33]. The flooding and drought events in the area have raised the importance of drought and flood management and preparedness strategies.

Similarly, the area is a blend of urban, agricultural, and natural land. Forest and agricultural land dominate the land cover, including orchards, vineyards, and row crops. Natural areas, such as grasslands, oak woodlands, and chaparral, are also home to various wildlife species. Overall, the land use and cover are a mix of urban, agricultural, and natural areas. The area has diverse hydrogeological features blended by both natural watersheds and managed groundwater systems. Major rivers like the American and Truckee play crucial roles in surface and groundwater systems. The study region extends from the Sierra Nevada mountains to the Sacramento valley and comprises complex geology, including ancient metamorphic rocks, mineral resources, and seismic features. [34]

The North Fork American River contributes its runoff toward Folsom Lake. The Basin is created upstream of Folsom Lake in the American Fork River tributary. It is one of Northern California's most extended branches of the American River. The flow of this river fluctuates depending on the precipitation amount and time of the year. The flow is high in the rainy season and may decrease during dry months. Basin-contributing runoff on Folsom Lake is critical, as water from it is used for recreation, drinking water, irrigation, and hydroelectric power plants [35]. Figure 2 indicates the historic drought conditions in California, as obtained from the US drought monitor. This shows the vulnerability of California to drought and the need of its preparedness for future drought conditions. Here, the figure indicates extreme drought occurrences in California state overall from 2014 to 2017 and later in different months of 2021, 2022, and 2023. This has significance with the drought index value that we evaluated with the precipitation data in that time interval, although our study area is confined to a part of California which is basically the watershed area contributing towards the runoff of Folsom Lake. Table 1 represents the gauge information of the watershed used in our research study.

Table 1. Gauge information involved in our study area.

Water Network and Its Location	Gauge id	Latitude	Longitude	Elevation (m.a.s.l)
Lake Valley canyon near North Fork American river	11426190	39°17′56″	120°38′31″	1341
North Fork American river at North fork dam (study outlet)	11427000	38°56′10″	121°01′22″	579
Onion Creek tributary no.3 near Soda springs	11426110	39°17′04″	120°21′20″	1099.5
Onion Creek tributary no.5 near Soda springs	11426120	39°17′04″	120°20′44″	564
Onion Creek tributary no.2 near Soda springs	11426130	39°16′34″	120°21′57″	457
Onion Creek tributary no.1 near Soda springs	11426140	39°16′30″	120°21′58″	406
Onion Creek near Soda springs	11426150	39°16′02″	120°21′50″	1828
Onion Creek tributary no.7 near Soda springs	11426160	39°15′58″	120°21′19″	300
NF Forbes Creek near Dutch flat	11426200	39°08′37″	120°45′30″	1163
North Shirttail Creek near Dutch flat	11426400	39°07′49″	120°47′44″	1110
North Fork American river near Colfax	11426500	39°02′25″	120°54′06″	671
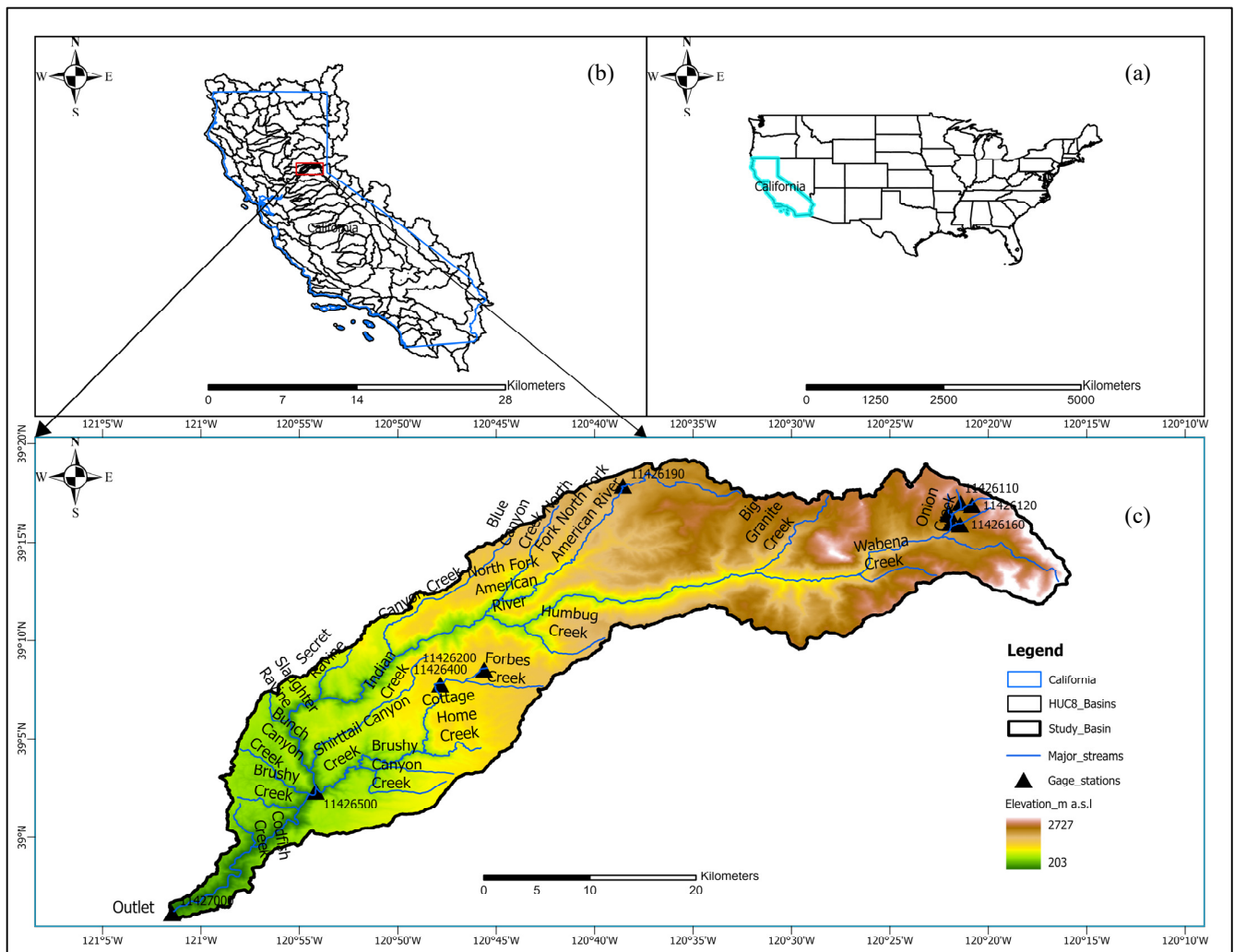
Figure 1. (**a**) Map of the United States with California state; (**b**) map of California with every watershed; and (**c**) map of the watershed with gauge stations and water line.
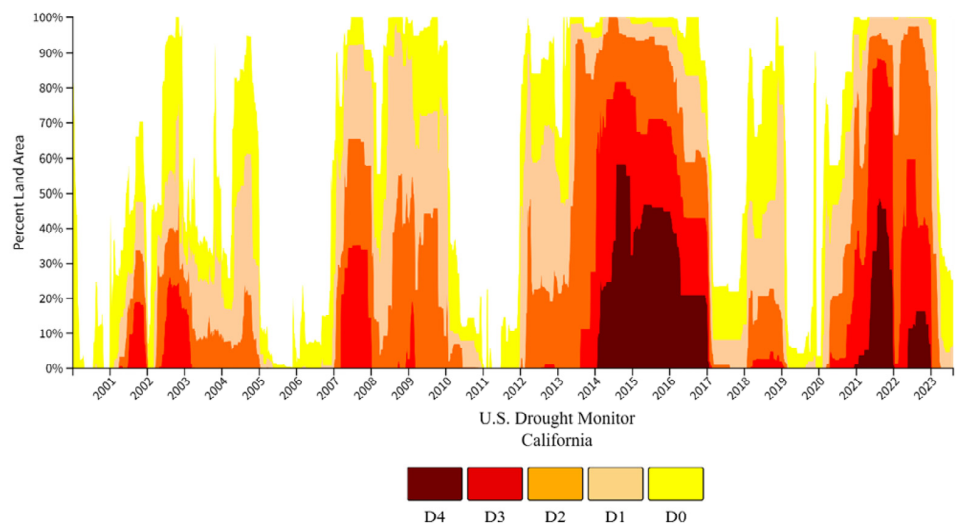
Figure 2. Historical drought conditions in California (D0, D1, D2, D3, and D4 indicates Abnormally Dry, Moderate Drought, Severe Drought, Extreme Drought and Exceptional Drought conditions, respectively) [36].

2.2. Data and Processing

This study applied different data types to prepare a comprehensive hydrological model using the HEC-HMS and machine learning. These data sources included a digital elevation model (DEM), which provides valuable information on elevation and slope characteristics across the study location. The curve number (CN) grid, a necessary component of hydrological analysis, was generated by integrating data from land use, land cover (LULC), and soil group data. This grid allowed for a detailed assessment of potential runoff within the study region. The impervious data set was crucial in identifying non-permeable surfaces, encompassing urban infrastructure such as buildings and road networks. These datasets were integrated within the Arc-GIS platform, and the resultant integration was employed to formulate a comprehensive hydrological model utilizing the HEC-HMS framework.

Furthermore, the research also incorporated precipitation data obtained from the CHRS data portal, which served as a fundamental input in multiple modeling techniques, including the HEC-HMS, RF, and SVR models. This precipitation data was crucial in enhancing the precision and reliability of the hydrological assessments. Moreover, discharge data from the United States Geological Survey (USGS) gauging station were diligently incorporated into the research, further enriching the hydrological analysis with real-world, ground-truth data. The monthly average discharge and precipitation values for 21 years dated from 2001 to 2021 were used for this study. This holistic approach, amalgamating various datasets and integrating observed discharges, bolstered the scientific rigor and comprehensiveness of the study's hydrological modeling and predictions. The data used and their sources are mentioned below in Table 2.

Table 2. Illustration of data used and their sources.

Data Used	Sources
Digital Elevation Model (DEM)	National Map Viewer (The National Map Viewer|U.S. Geological Survey)
Precipitation	CHRS Data Portal (CHRS Data Portal)
Station Discharge	USGS USGS Current Water Data for the Nation
Land Use and Land Cover (LULC)	National Land Cover Database USGS (LULC)
Watershed Boundary	USGS Stream stat ("https://www.usgs.gov/streamstats accessed on 4 April 2024")
Soil	USDA (USDA—National Agricultural Statistics Service—Quick Stats)

2.3. Drought Indices

2.3.1. Standardized Precipitation Index (SPI)

The SPI measures the severity of meteorological drought using the probability distribution function of precipitation at multiple time scales, directly indicating drought due to a lack of precipitation. Thus, SPI's evaluation considers only precipitation data for the development of the probability distribution. The higher time frame precipitation data are adjusted with a specific (typically gamma) distribution, which is then altered into a standard normal distribution using an equal probability transformation approach [37]. SPI values below zero indicate dry conditions, while those above zero signify wet conditions [38]. The SPI value indicates the deviation of the total precipitation deficit from the normalized mean value [39]. It can be evaluated for various durations like 1, 3, 6, 12, 24, and 48 months. The SPI was chosen as a meteorological index for comparing drought severity by the participants in the inter-regional workshop on indices and early warning [40]. This study uses the SPI as a predictor for hydrological drought prediction in a machine learning

model. The formula for calculating the SPI for a given month 'i' is generally expressed by Equation (1),

$$SPI_i = \frac{X_i - X_{mean}}{\sigma} \tag{1}$$

where, X_i is the observed precipitation for the month i, X_{mean} is the mean precipitation and σ is the standard deviation of the precipitation data. In this study, the SPI is computed for periods of 1 month and 3 months, which are denoted as SPI1 and SPI3, respectively. Similarly, Table 3 presents the SPI classification for drought severity, as outlined by Fitchett in 2019. The SPI values are categorized into various conditions ranging from extremely wet to extremely dry. This classification provides a qualitative framework for interpreting SPI values, aiding in assessing moisture conditions based on historical precipitation data.

Table 3. SPI classification to represent drought severity [41].

SPI Range	Conditions
≥ 2.0	Extremely wet
$1.5 \geq 1.99$	Very wet
$1.0 \geq 1.49$	Moderately wet
$-0.99 \geq 0.99$	Near Normal
$-1.0 \geq -1.49$	Moderately dry
$-1.5 \geq -1.99$	Severely dry
≤ -2	Extremely dry

2.3.2. Standardized Streamflow Index (SSI)

The SSI is another extensively used drought index, especially to evaluate the severity of hydrological drought [42]. The computation process for the SSI closely resembles that of SPI; the SSI calculation uses observed or simulated streamflow data, while SPI evaluation requires precipitation data. The frequencies of monthly streamflow values are found through a specific probability distribution function. Then the quantiles of the standard normal distribution corresponding to the frequencies are calculated as the SSI values [43]. The classification of the SSI is presented in Table 4. The SSI can also be evaluated at different time scales. This study considered 1- and 3-month SSI evaluations to characterize and compare seasonal and short-term droughts with SPI evaluations. The formula for calculating the SSI for a given month 'i' is generally expressed by Equation (2),

$$SSI_i = \frac{Q_i - Q_{mean}}{\sigma} \tag{2}$$

where, X_i is the discharge for the month i, X_{mean} is the mean discharge, and σ is the standard deviation of the discharge value. This study computed SSIs for 1 month and 3 months, denoted as SSI1 and SSI3, respectively. Similarly, the SSI range classification, as in Table 4, provides a comprehensive overview of hydrological conditions based on probability within specified intervals. Ranging from extremely wet to extreme drought, the SSI values are associated with distinct conditions and their corresponding probabilities. For instance, an SSI value equal to or greater than 2.0 signifies extremely wet conditions with a 2.3% probability, while an SSI value of -2 or lower indicates extreme drought conditions, also with a 2.3% probability. These classifications, incorporating condition and probability, contribute to a nuanced understanding of streamflow variability, offering valuable insights into the likelihood of different hydrological states.

Table 4. Hydrological drought classification based on SSI [44].

SSI Range	Condition	Probability
≥2.0	Extremely wet	2.3%
1.5 ≥ 1.99	Severe wet	4.4%
1.0 to 1.5	Moderate wet	9.2%
−1 to 1.0	Near Normal	68.2%
−1.5 to −1.0	Moderate drought	9.2%
−2.0 to −1.5	Severe drought	4.4%
≤−2	Extreme drought	2.3%

2.4. Hydrological Model

This study focuses on drought indices, which require a long-term series of discharge data in the study area. However, such data may not always be available in regions where data are scarce. To overcome this issue, the study employs the HEC-HMS model to simulate the river discharge at the outlet of the study watershed. The simulated data are then utilized to calculate the drought indices. To evaluate the discharge at the basin outlet, the hydrological analysis uses the HEC-HMS model. This model is typically applied to evaluate discharge with precipitation, taking into account the meteorological and topographic properties of the basin [45]. The HEC-HMS simulates the process of converting rainfall into runoff during an event while considering critical factors that control the runoff [46]. The HEC-HMS model consists of distinct loss, transformation, and watershed routing modules. The routing, loss, and transformation methods chosen are the Muskingum routing method, Soil Conservation Service (SCS) Curve Number, and SCS unit hydrograph methods, due to their widespread use, stability, and data availability. The HEC-HMS model integrates the basin model, meteorological model, control specification, and input data as its components.

A basin model explains the basin's physical attributes to simulate runoffs and rainfall abstraction [47]. The inputs used for the model are the watershed characteristics such as land use, soil types, and other relevant parameters like impervious area. Variable data sources like USGS and USDA were utilized. DEM, LULC, and soil group data of the area of interest were clipped using an arc map. This process aimed to estimate the extent of impervious surfaces within the research area's geographical boundaries. The basin model used the SCS curve number (CN) method to calculate runoff and abstraction [48].

At the beginning of the project, Arc-Hydro, an integrated part of Arc-GIS Pro software with version 3.2.1, extracts the network of rivers and the terrain data. We used various Arc Hydro tools to obtain various datasets that represent the drainage patterns of the catchment. These datasets included stream definition and segmentation, flow direction and accumulation, and watershed delineation generated using raster analysis. Next, the sub-basin parameters were derived using arc hydro tool-generated raster data. The curve number is a hydraulic attribute that we generated utilizing the soil and land use database in Arc Hydro. Moreover, the HEC-GeoHMS allowed us to import lag time and impervious percentages like hydrological attributes, which were crucial in our study.

The meteorological model considers various parameters, including rainfall and discharge. For this study, rainfall data for the research area were obtained from the Precipitation Estimation from Remotely Sensed Information using ANN—Dynamic Infrared Rain Rate near real-time (PDIR-Now). This source allows the use of up-to-date, high-resolution ($0.04° \times 0.04°$ pixel) satellite rainfall data from all around the world. On the other hand, runoff data for the part of the specific river being studied were taken from the USGS database. At the downstream endpoint, there is just one USGS gauging station (USGS 11427000) within the study area. The accessed discharge data covers observations from 2001. Furthermore, daily rainfall and discharge data have been collected within a grid for 21 years, from 2001 to 2021.

2.5. Machine Learning Models

2.5.1. Random Forest

In this investigation, the RF machine learning method was employed to predict drought severity in the study watershed. RF stands out as a supervised machine learning method renowned for its predictive and classification capabilities. Its computational efficiency, resilience to instability, and capacity to model non-linear relationships render it particularly appealing [49]. The RF model constructs multiple decision trees based on a bootstrapped training data sample. Notably, a random subset of predictors is considered for forming binary splits at each decision tree node, providing diversity to the individual trees. To derive the anticipated response, one navigates through the tree starting from its main node to the specific end node. The overall prediction is then calculated by averaging the predictions made by each individual tree. The selection of the best combination of decision trees contributes to the model's robustness [50].

To enhance the efficiency and predictive performance of the simulation, focused consideration was given to features with high predictive potential. A random forest regressor was used to mitigate overfitting, and each tree's depth was restricted to 100. The nodes were continuously expanded until the number of samples in the leaves became lower than the specified amount. Additionally, parameters such as "max_features" were set to 'auto' and "max_leaf_nodes" were set to 'None', allowing the model to expand based on optimal fitting requirements. The model produces its output by calculating the average of the outputs generated by each individual tree. This ensemble approach, incorporating multiple decision trees, not only guards against overfitting but also synergistically leverages the diverse insights of each constituent tree to enhance the overall predictive accuracy of the RF model.

2.5.2. Support Vector Regression

The SVR model is a machine learning model with a regression pattern similar to the SVM model. It is a supervised ML algorithm proposed by Vladimir and is widely used for nonlinear problems [51]. SVMs are widely used ML techniques for both classification and regression. When employing an SVM with a kernel function (such as radial, linear, sigmoid, or polynomial), the approach involves transforming a nonlinear problem into a higher-dimension space. In this elevated space, the initially nonlinear problem is converted into a linear one, which is then addressed through SVM techniques [5].

In SVR, the function used to model the data points is formed by linearly combining "kernels" centered around each input point [52]. In the SVR model used, a linear kernel relates input features and the target variable linearly. SVR develops the decision boundary or optimal separating hyperplane.

2.6. Selection of Input Variables

The input variable selection process for the machine learning models in this study involves a rigorous process to identify the most influential predictors. Encompassing climatic and hydrological factors, the chosen predictors include discharge data, satellite precipitation data, SPI, and SSI data spanning 1 to 5 months ahead, along with the cumulative sum of 5 months of precipitation, specific humidity, and dew point data. This selection process begins with a comprehensive correlation analysis among all variables, emphasizing the identification of pairs exhibiting significant correlations. Variables with higher correlation coefficients are prioritized, signifying stronger relationships. The goal is to retain a subset of predictors optimized for capturing variability without overfitting, ensuring that the chosen variables collectively maximize predictive power while minimizing redundancy, ultimately enhancing the efficiency and interpretability of the hydrological and climatic predictive models.

The correlation matrix is shown in Figure 3. It was found that discharge, precipitation, SPI, and SSI data from 1 to 5 months ahead and the sum of 5 months of precipitation

data correlate better with SSI1 and SSI3. Based on this good correlation, they are used as predictors for SSI1 and SSI3 in the machine learning models.
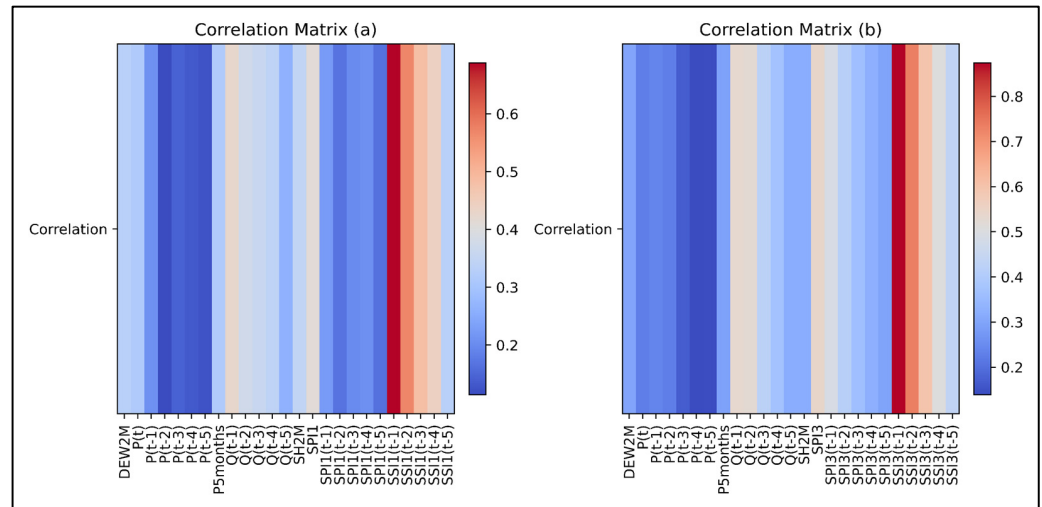

Figure 3. The correlation matrix: (**a**) SSI1 (**b**) SSI3.

First, the meteorological drought index (SPI) and hydrological drought index (SSI) are calculated from the satellite precipitation data and gauge station, respectively, for a duration of 1 and 3 months. Later, the standardized streamflow index is calculated by evaluating the discharge from the HEC-HMS model. Figure 4 indicates the different steps involved in hydrological modeling using the HEC-HMS.

Machine learning models take 75% of the selector's predictor dataset as a training set and 25% of the data as a testing dataset. The predictors used here are mentioned in Table 5 Using these predictors, the SSI is predicted for two-time durations, 1 and 3 months. Later, the value of the coefficient of determination, root mean square error, and mean absolute error are evaluated in the Python platform for the training, testing, and overall datasets.
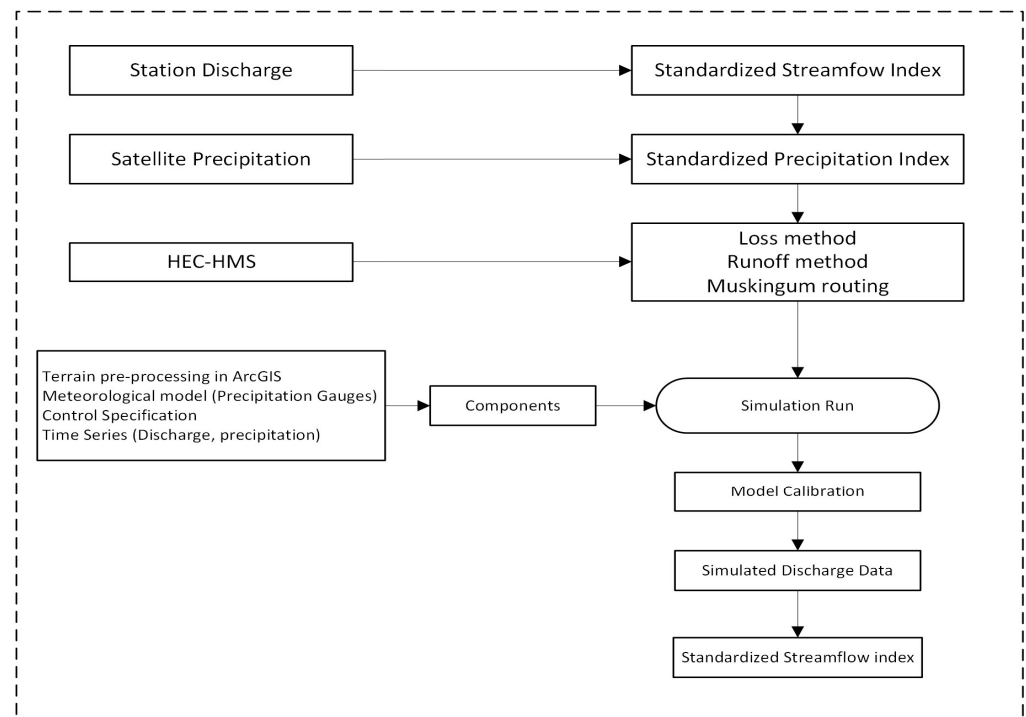

Figure 4. Flowchart of Hydrology analysis and drought index calculation.

Table 5. Input and output variables for machine learning models.

S. N	Input	Output
1	Q(t-1), Q(t-2), Q(t-3), Q(t-4), Q(t-5), P(t), P(t-1), P(t-2), P(t-3), P(t-4), P(t-5), P5months, SPI1, SH2M, DEWP2M, SPI1(t-1), SPI1(t-2), SPI1(t-3), SPI1(t-4), SPI1(t-5), SSI1(t-1), SSI1(t-2), SSI1(t-3), SSI1(t-4), SSI1(t-5)	SSI1
2	Q(t-1), Q(t-2), Q(t-3), Q(t-4), Q(t-5), P(t), P(t-1), P(t-2), P(t-3), P(t-4), P(t-5), P5months, SPI3, SH2M, DEWP2M, SPI3(t-1), SPI3(t-2), SPI3(t-3), SPI3(t-4), SPI3(t-5), SSI3(t-1), SSI3(t-2), SSI3(t-3), SSI3(t-4), SSI3(t-5)	SSI3

2.7. Evaluation Parameters

The observed and predicted data were analyzed, and the most suitable model was chosen using the coefficient of determination (R^2) value and root mean square error (RMSE) as criteria. RMSE measures the variance of errors between the actual and predicted values, whereas R^2 determines the fitness between the predicted and original values.

$$R^2 = \frac{\sum_{i=1}^{N}(\hat{Y}_i - \overline{Y}_i)}{\sum_{i=1}^{N}(Y_i - \overline{Y}_i)^2} \quad (3)$$

$$\overline{Y}_i = \frac{1}{N}\sum_{i=1}^{N} Y_i \quad (4)$$

where Y_i and $\hat{Y}_i$ are the observed and predicted values, $\overline{Y}_i$ is the mean value, and N is the amount of data.

$$RMSE = \sqrt{\frac{SSE}{N}} \quad (5)$$

$$SSE = \sum_{i=1}^{N}(\hat{Y}_i - Y_i)^2 \quad (6)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|\hat{Y}_i - Y_i| \quad (7)$$

where RMSE, SSE, and MAE indicate the root mean square error, sum of squared error, and mean absolute error, respectively. A higher R^2 value implies a better prediction capacity of the model, and if the value is 1, then there is a perfect correlation between the predicted and observed values. Usually, values above 0.5 are considered acceptable [53]. This gives a foundation for determining whether the model is suitable for prediction. Similarly, RMSE quantifies the difference in the actual and predicted values. A lower RMSE value indicates the closeness of the actual values with predicted values. A Zero RMSE value is taken as a perfect fit. RMSE values that are lower than half the standard deviation of the measured data can be deemed low, making them suitable for assessing the model's performance.

3. Results

3.1. Verification of HEC-HMS Model

In this study, the HEC-HMS hydrological model was verified to assess its performance in simulating streamflow. R^2 was employed to evaluate the model's accuracy in replicating the observed hydrological data. as the results presented in Figure 5 demonstrate the model's capability. The R^2 value of 0.73 suggests that the HEC-HMS model strongly correlates with the actual flow. Furthermore, Figure 6 provides a scatter plot that visually illustrates the relationship between the HEC-HMS model's simulated values and the observed data points. This plot represents how well the model aligns with the actual hydrological measurements, visually confirming the model's accuracy in depicting both high and low streamflow values.

The noteworthy outcome of this study is that the HEC-HMS model exhibits a remarkable ability to accurately represent a wide range of streamflow conditions, encompassing both high-flow and low-flow situations. This versatility underscores the model's reliability and suitability for various water resource management applications, making it a valuable

tool for drought prediction. An R^2 value lying between 0.7 and 0.9 implies a strong correlation, while a value between 0.5 and 0.7 represents a moderate correlation. Furthermore, a value ranging from 0.3 to 0.5 indicates a relatively weak correlation [54]. For the forecast of SPI1 and SPI3, the results of SVR are better than those of RF. To calibrate and validate the hydrological model, parameters like loss method, transform, baseflow, and routing method were adjusted to ensure that the simulated values matched the observed values at the outlet. The validation was performed using a single hydrological event from 4 January to 14 January 2018. The key statistical metric R² was calculated for both events, with the results of R² being greater than 0.7.
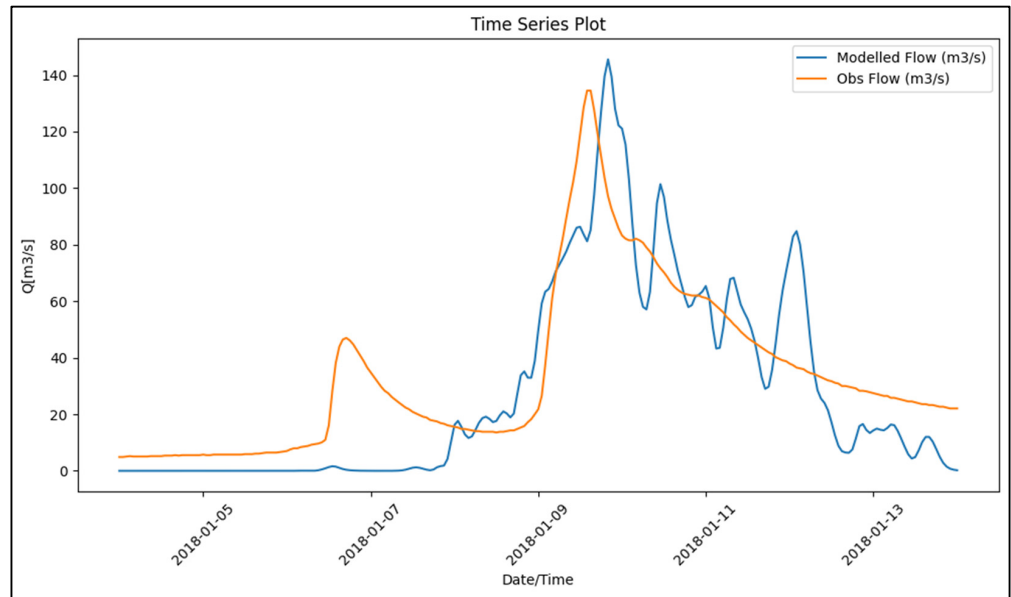
**Figure 5.** Modeled flow vs. observed flow from 4 January to 14 January 2018.
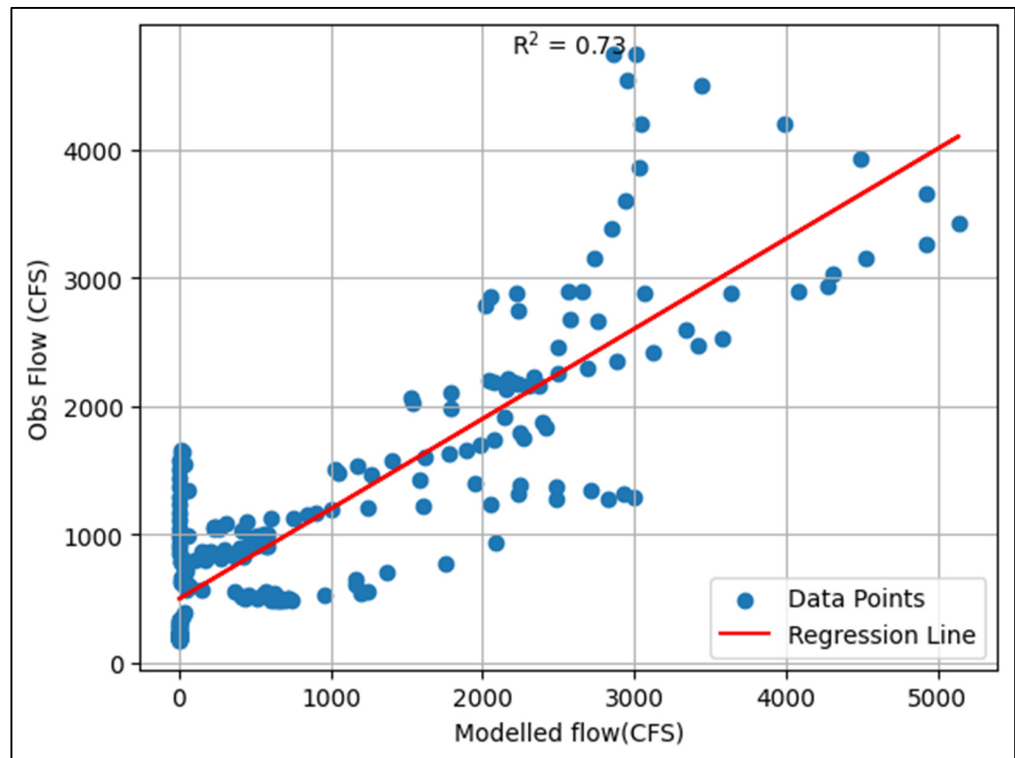
Figure 6. Correlation graph showing the scatter plot of modeled Vs observed flow.

3.2. Drought Evaluation Using HEC-HMS Model

The HEC-HMS-evaluated drought conditions correlate at 0.89 for SSI-1 and 0.84 for SSI-3. The errors are also significantly low, with MAE values of 0.115 and 0.131 and RMSE values of 0.290 and 0.361 for SSI-1 and SSI-3, respectively. These results indicate a greater significance of HEC-HMS-produced discharge for drought evaluation. Table 6 and Figure 7 contain the evaluation parameters with their values and the regression line of the HEC-HMS vs. observed SSIs for 1 and 3 months, respectively. The result indicates the hydrological model's capability to depict the drought condition using the modeled streamflow values in the basin.
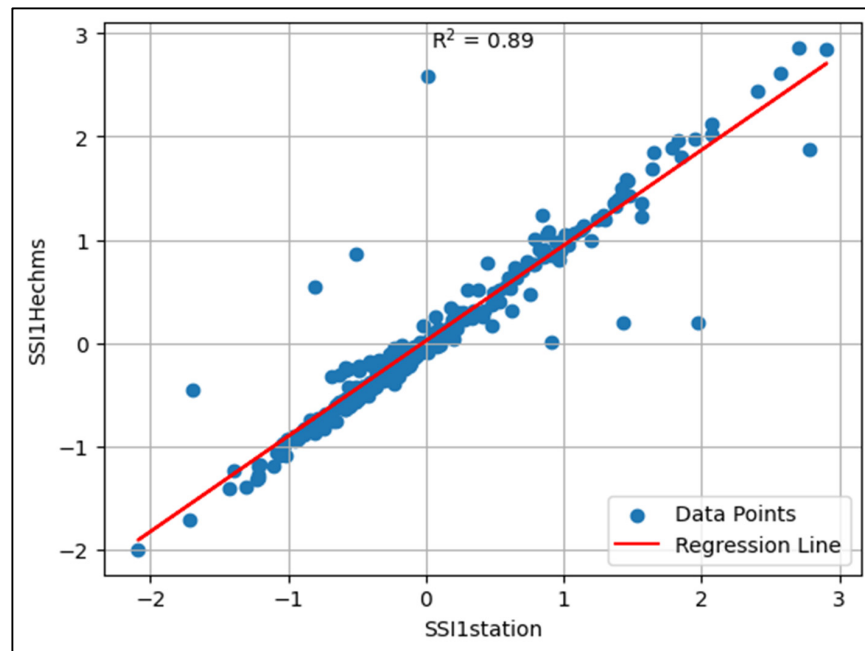
(a)
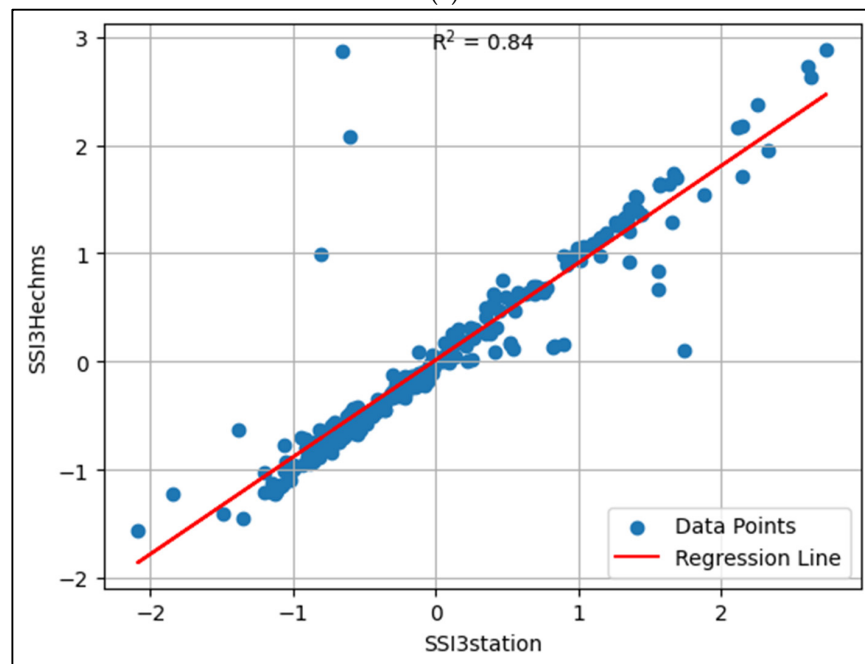
(b)

Figure 7. Regression analysis of the standardized streamflow index for the given months: (**a**) 1 month and (**b**) 3 months from the observed flow in the basin to the SSI from the HEC-HMS.

Table 6. Evaluation Parameters (HEC-HMS Model).

Evaluation Parameters	SSI-1	SSI-3
MAE	0.115	0.131
RMSE	0.290	0.361
R^2	0.89	0.84

### 3.3. Drought Prediction Evaluation Using Machine Learning Models

In this study, the verification of the random forest and the support vector regression model was conducted to assess its performance in hydrological drought prediction. Three essential performance metrics (RMSE, MAE, and R^2) were employed to evaluate the relationship between the predicted and occurred drought.

3.3.1. Random Forest Model

The result from Table 7 and Figure 8a,b show a high correlation value of 0.823 in higher-duration drought SSI-3 as compared to low-duration drought SSI-1, with a correlation value of 0.628 in the testing dataset and similar in the overall dataset. Similarly, the testing datasets for a 3-month duration drought have low MAE and RMSE values of 0.323 and 0.434, respectively, and the MAE and RMSE values are higher, with values of 0.485 and 0.605, respectively, in a 1-month duration drought. This result proves the better capability of the RF model for short-term drought prediction.
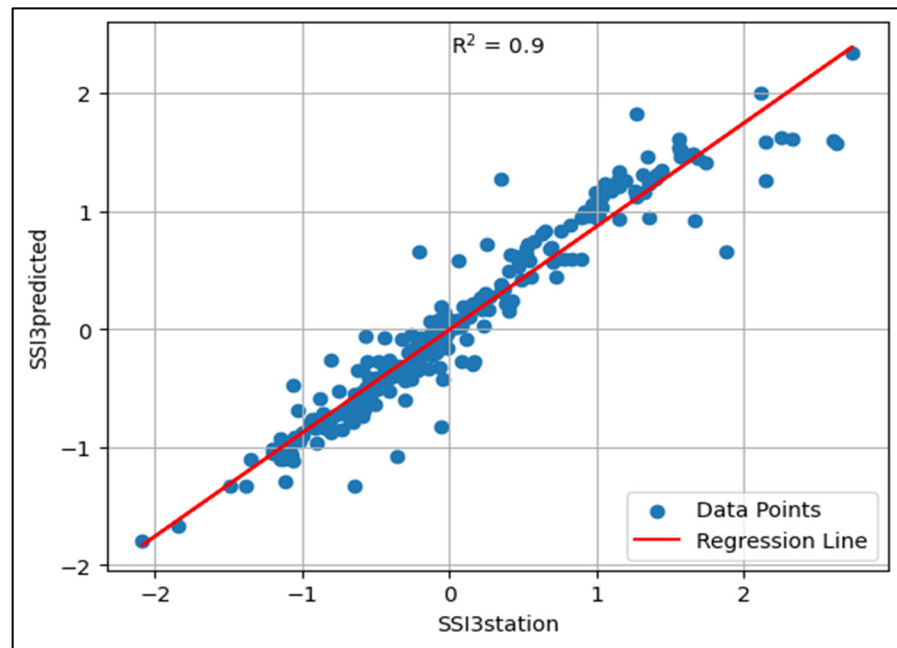
Table 7. Evaluation parameters for SSI-1 and SSI-3.

	SSI-1			SSI-3		
Evaluation Parameters	Training	Testing	Overall	Training	Testing	Overall
MAE	0.137	0.485	0.224	0.099	0.323	0.156
RMSE	0.20	0.605	0.347	0.150	0.434	0.252
R^2	0.984	0.628	0.85	0.968	0.823	0.9
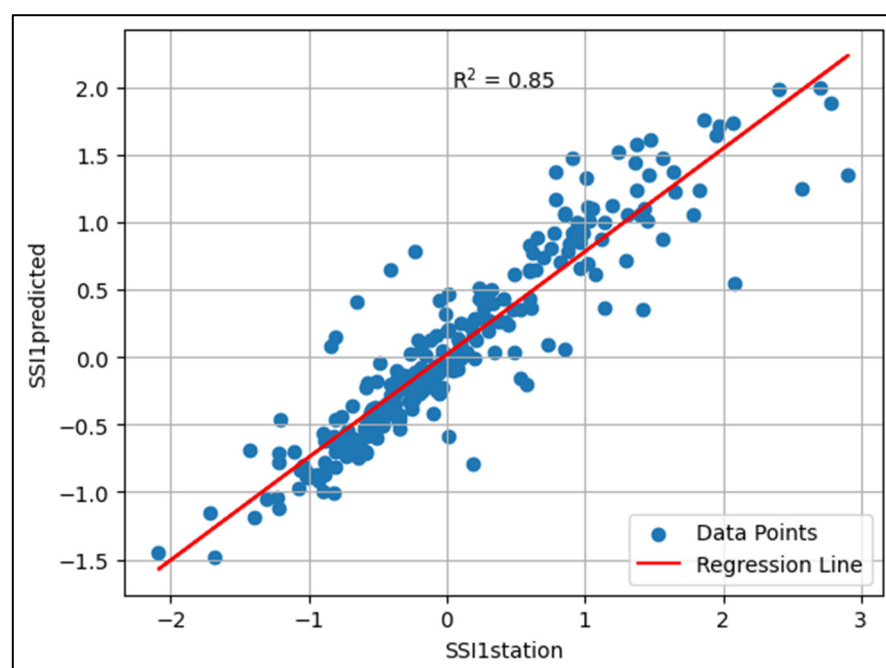
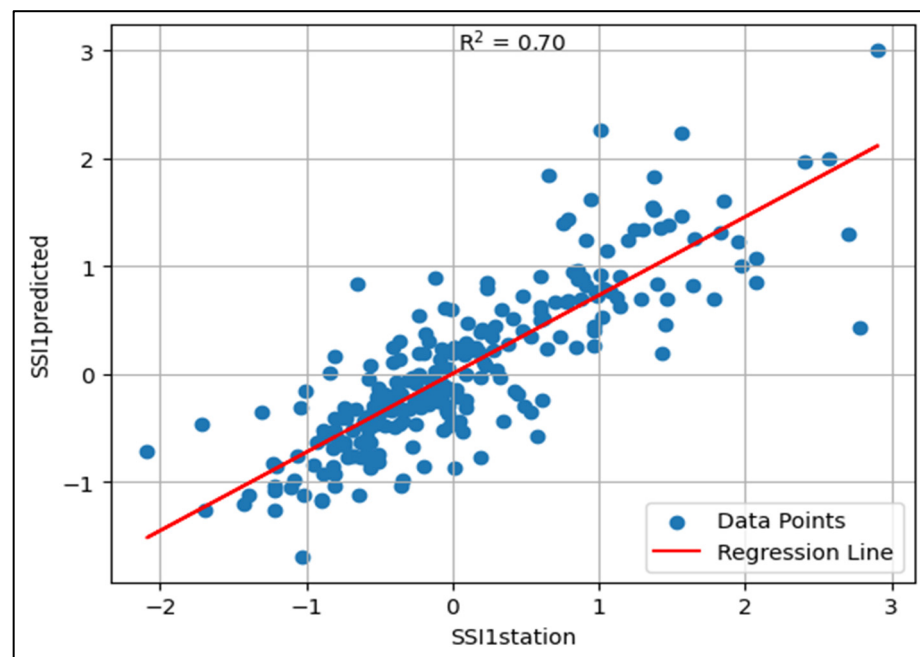(**a**)

Figure 8. *Cont*.

(**b**)

Figure 8. Correlation graph of observed and random forest predicted standardized streamflow index: (**a**) SSI1 (overall data) and (**b**) SSI3 (overall data).
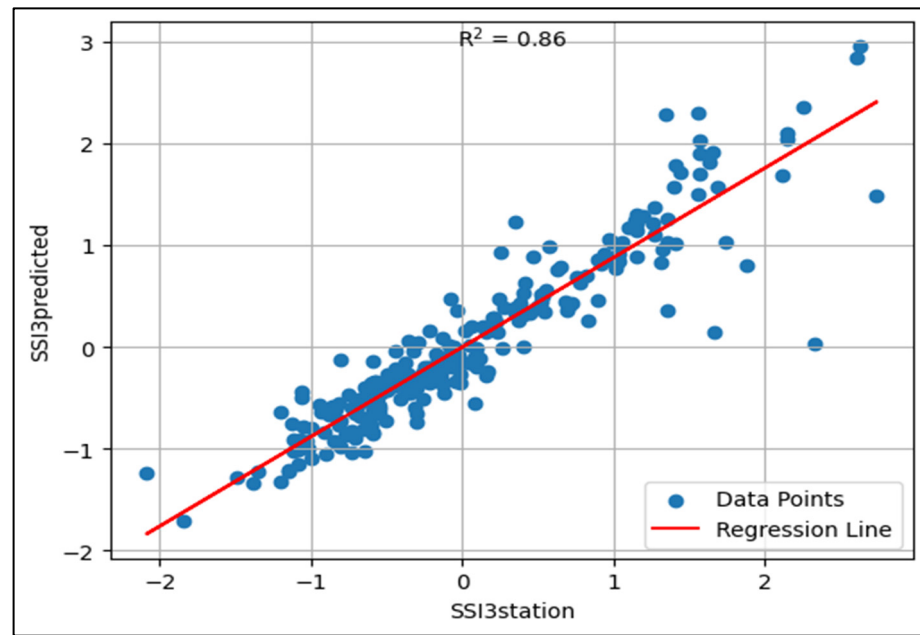
3.3.2. Support Vector Regression Model

The result from Table 8 and Figure 9a,b show a high correlation value of 0.903 in the higher-duration drought SSI-3 as compared to the low-duration drought SSI-1, with a correlation value of 0.629 in the testing and overall datasets. Similarly, the testing datasets for a 3-month duration drought have low MAE and RMSE values of 0.209 and 0.287, respectively, and the MAE and RMSE values are higher, with values of 0.392 and 0.500, respectively, in a 1-month duration drought. This result proves the capability of the support vector regression model to predict drought better.

(**a**)

Figure 9. *Cont.*

(**b**)

Figure 9. Correlation graph of observed and support vector regression estimated. Standardized streamflow index: (**a**) SSI1 (overall data) and (**b**) SSI3 (overall data).

Table 8. SSI-1 and SSI-3 evaluation parameters (support vector regression model).

| Evaluation Parameters	SSI-1			SSI-3		
	Training	**Testing**	**Overall**	**Training**	**Testing**	**Overall**
MAE	0.331	0.392	0.346	0.212	0.209	0.211
RMSE	0.480	0.500	0.482	0.340	0.287	0.328
R^2	0.70	0.629	0.696	0.847	0.903	0.862

3.4. Standardized Streamflow Index Variation

Figure 10a shows a graph of the standardized streamflow index for a one-month duration obtained from the observed streamflow data, hydrologically modeled streamflow data, and SVR and RF model-predicted data. We evaluated the streamflow index from the monthly discharge data found in the station and the discharge data we obtained from the HEC-HMS hydrological modeling tool. Using the climatic variables and SPI1 and SSI1 values of 5 months before, we used random forest to predict the SSI1 values. These three evaluated SSI1 values were plotted in Figure 10a. The graphs have many similarities with each other in the initial years. The HEC-HMS obtained SSI1 values which indicate a higher value than the actual, whereas the SSI1 values from the prediction model under-evaluated the actual SSI1 values in most peak conditions. It is observed that the predicted SSI1 graph follows a similar pattern to the actual data. In 2015, SSI1 values from the prediction model are higher than those from the HEC-HMS and the station. During the years 2011 and 2021, the obtained SSI1 value from the HEC-HMS was higher than the predicted SSI1 value.

Figure 10b shows the graph of the standardized streamflow index for a 3-month duration. As for the one-month drought, three SSI3 values were derived from hydrologically modeled discharge, station discharge, and random forest SVR-predicted discharge. The predicted and actual SSI3 values obtained are close to each other most of the time, whereas the HEC-HMS-derived SSI3 values over-evaluated the SSI3 values. Despite this, the hydrologically modeled SSI3 values and the random forest-predicted SSI3 values nearly coincide with the actual value. In the month of 2011, the HEC-HMS graph exceeded both stations

and the predicted SSI3 values. Also, in 2015, the drought condition was undervalued by the HEC-HMS and the predicted SSI3 data.

In comparing the statistical results derived from both machine learning models, time-lagged SPI and SSI values, along with precipitation, discharge, specific humidity, and dew points have provided a strong basis for SSI prediction. When there is increased precipitation, more will be discharged into the basin, which reduces the chance of drought. This shows the inverse relation of precipitation and discharge in drought occurrence, and it is a similar case for humidity. Low dew point values often indicate dry atmospheric conditions and are taken as a parameter for causing drought conditions.
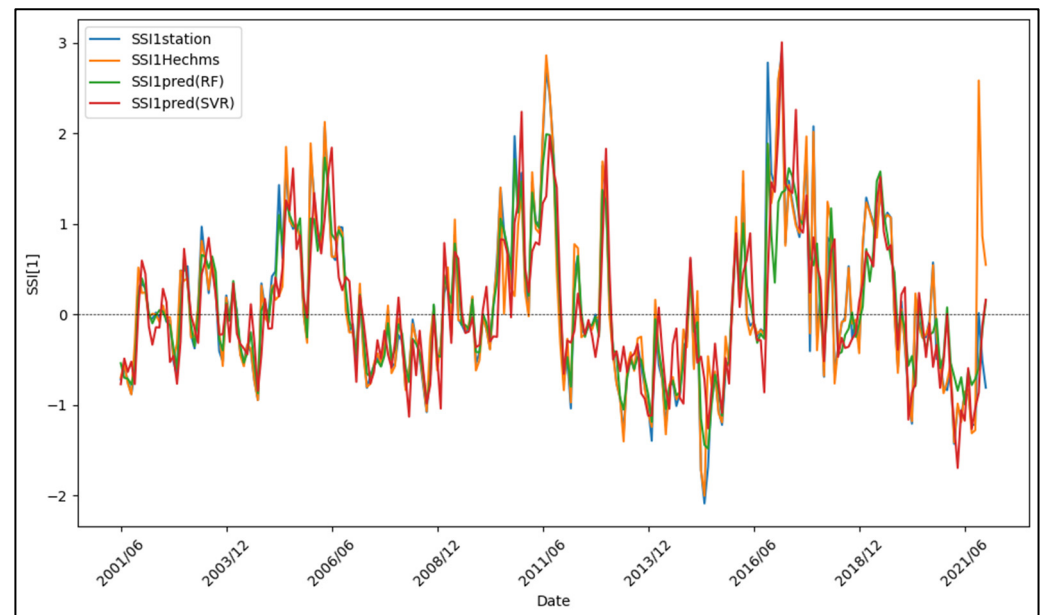
(**a**)
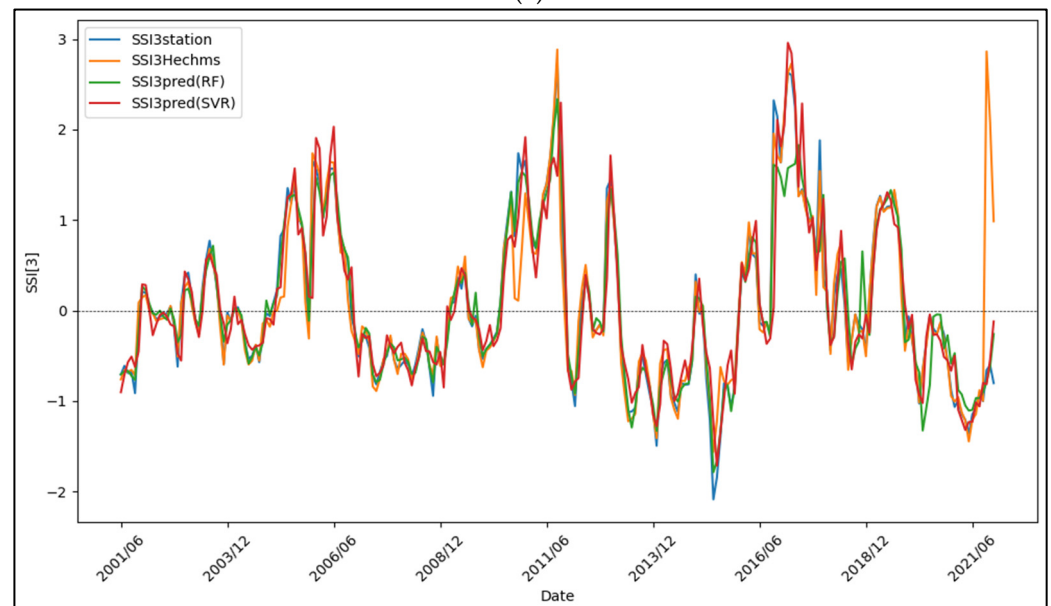
(**b**)

Figure 10. (**a**) SSI1 and (**b**) SSI3.

The result as indicated in the graphs from Figure 10a,b is found to depict the real drought scenario. The extreme drought condition is predicted to occur in 2015 and 2021, which has happened in the given time. This proves the better drought predictability

of all the models and the higher accuracy prediction capability of SVR as a machine learning model.

4. Discussion

The results of the study explain and predict drought occurrences in a watershed near Folsom Lake. Both hydrological and machine learning models use satellite precipitation data from the PDIR-Now dataset. PDIR-Now was found to be the best precipitation product in the study of Huang et al. [55] among different PERSIANN family products. The satellite-based rainfall estimates are advantageous regarding accuracy, timeliness, spatial coverage, and cost efficiency [56]. It is efficient for real-time rainfall monitoring and showing the development of drought conditions. Satellite precipitation data have demonstrated the capability to accurately depict the spatiotemporal fluctuations in precipitation across most global regions with exceptional precision [57]. Nonetheless, the precision of satellite-derived precipitation is influenced by numerous factors. A significant drawback of satellite precipitation data is their limited historical data availability. Currently, only PERSIANN-CDR and CHIRPS offer data records exceeding 30 years of coverage, leaving many research studies constrained by the existing dataset [58].

This study used hydrological modeling to evaluate droughts. Trambauer et al. [59] reviewed previous studies and found the hydrological model suitable for drought forecasting. Xing et al. [60] conducted research on the adaptability of hydrological models for the purpose of simulating and forecasting droughts. The study concluded that the HEC-HMS model is suitable for forecasting hydrological droughts due to its adaptable structure. These studies concur with our results showing the better performance of the HEC-HMS model for drought evaluation.

This study showed the high accuracy of the SVR and RF models in predicting hydrological drought. In a study by Jehanzaib et al. [61], it was concluded that the SVM model showed better performance than the RF model. The SVM model holds significance in the realm of hydrological variables because of its effectiveness in handling high-dimensional spaces. The predicted results for a 3-month duration show a comparatively precise value for 1 month. This coincides with the drought prediction result of Belayneh and Adamowsk [54], where the SPI6 forecasts were found to be more accurate than SPI3. The only difference is that our study included SSI1 and SSI3 instead of SPI3 and SPI6. The result of SSI3 is more accurate than SSI1 in both machine learning models. This is because of the greater randomness in the weather over a shorter time, which makes accurate prediction challenging. The drawback of both models used here is that neither have taken other types of hydrological data into account, such as groundwater level and temperature, which increase the chances of drought occurrence.

The study employed the HEC-HMS model as the hydrological model, which models several processes, including baseflows, infiltration, and rainfall runoff, using empirical techniques that might not accurately reflect the system's behavior. Since the HEC-HMS is intended for natural watersheds, it might not yield reliable results in urban watersheds. Furthermore, future climatic changes like temperature and precipitation patterns are not taken into consideration by the HEC-HMS. Integration with other models, such as SWAT or MODFLOW, can improve the model. The model can be further enhanced using real-world field data, such as soil wetness. Better HEC-HMS input could use climatic models to generate future climate scenarios. With noisy datasets, RF and SVR can perform poorly, which may result in overfitting. Hybrid modeling, hyperparameter adjustment, and feature selection can all help overcome this constraint. Effective drought forecasting and preparedness are essential, and this can be achieved by implementing measures to conserve water supplies, such as building dams. According to Kazakis et al., mall dams help retain water for groundwater recharge in the summer [62]. On the other hand, their quantity and location are determined by their future uses, such as fish farming, leisure activities, or hydropower research. Additional management techniques, such as tiered

water pricing, rainwater collecting, and intelligent irrigation, can be used to lessen the severity of droughts.

5. Conclusions

The results indicated that the RF and SVR models predict the SSI more accurately for three months of drought than only one month of drought conditions. This is because there is higher fluctuation in the climatic parameters used in the prediction model in the shorter time duration compared to the longer time duration. The results also indicate the better performance of the SVR model than the RF model in drought prediction for this particular study basin area. Despite slight differences in the prediction results, RF can also be used for drought forecasting with significant accuracy. Unlike previous research, our research has closely examined the performance of the hydrological model using satellite-based precipitation data to evaluate the drought index. The result is satisfactory compared to the actual condition. This allows using a hydrologically modeled drought index when discharge data is unavailable in the basin due to the absence of gauge stations.

The study compared the results based on the statistical parameter (R^2, MAE, and RMSE) values. The study used the present and previous months' lagged values of SSI, SPI, precipitation, discharge, dew point, and specific humidity as the machine learning model's input. Further research is suggested to investigate how different drought indices can be used and how they impact the precision of machine learning models in different climatic conditions. The variety of training and testing dataset proportions can be implemented in future studies. Moreover, if available, the model's performance can be enhanced by including more correlated parameters, varying the lead times, and with more refined data from a reliable source. Because of its complex nature, drought is influenced by factors like climatic conditions, land use, and socio-economic factors, so accurate prediction is not possible [30]. For the HEC-HMS model, numerous input variables are required, which can influence the model's output and, ultimately, the assessment of hydrological drought. Future research can incorporate other AI methods as prediction models and select the best for reliability, robustness, and accuracy. This study can be extended to multiple drought-affected basins with different climatic and physical conditions and evaluate the performance variations. Moreover, this study only focuses on hydrological drought using the standardized stream flow index. Future research can investigate other types of droughts, their severity, and frequency using different indices. Further, the hydrological drought result from other hydrological models rather than the HEC-HMS can be evaluated. The different machine learning models can be hybridized and tested for enhanced model performance in upcoming research.

References

1. Haile, G.G.; Tang, Q.; Li, W.; Liu, X.; Zhang, X. Drought: Progress in Broadening Its Understanding. *WIREs Water* **2020**, *7*, e1407. [CrossRef]
2. Hao, Z.; Singh, V.P.; Xia, Y. Seasonal Drought Prediction: Advances, Challenges, and Future Prospects. *Rev. Geophys.* **2018**, *56*, 108–141. [CrossRef]
3. Nielsen-Gammon, J.W. The 2011 Texas Drought. *Tex. Water J.* **2012**, *3*, 59–95. [CrossRef]
4. USDA. Available online: https://www.usda.gov/ (accessed on 24 March 2024).
5. Khan, N.; Sachindra, D.A.; Shahid, S.; Ahmed, K.; Shiru, M.S.; Nawaz, N. Prediction of Droughts over Pakistan Using Machine Learning Algorithms. *Adv. Water Resour.* **2020**, *139*, 103562. [CrossRef]
6. Thakur, B.; Parajuli, R.; Kalra, A.; Ahmad, S.; Gupta, R. Coupling HEC-RAS and HEC-HMS in Precipitation Runoff Modelling and Evaluating Flood Plain Inundation Map. In Proceedings of the World Environmental and Water Resources Congress 2017, Sacramento, CA, USA, 21–25 May 2017; pp. 240–251.
7. Brunner, M.I.; Slater, L.; Tallaksen, L.M.; Clark, M. Challenges in Modeling and Predicting Floods and Droughts: A Review. *Wiley Interdiscip. Rev. Water* **2021**, *8*, e1520. [CrossRef]
8. Dikshit, A.; Pradhan, B.; Huete, A. An Improved SPEI Drought Forecasting Approach Using the Long Short-Term Memory Neural Network. *J. Environ. Manag.* **2021**, *283*, 111979. [CrossRef]
9. Mishra, A.K.; Singh, V.P. A Review of Drought Concepts. *J. Hydrol.* **2010**, *391*, 202–216. [CrossRef]
10. Slette, I.J.; Post, A.K.; Awad, M.; Even, T.; Punzalan, A.; Williams, S.; Smith, M.D.; Knapp, A.K. How Ecologists Define Drought, and Why We Should Do Better. *Glob. Chang. Biol.* **2019**, *25*, 3193–3200. [CrossRef]
11. Vicente-Serrano, S.M.; Quiring, S.M.; Peña-Gallardo, M.; Yuan, S.; Domínguez-Castro, F. A Review of Environmental Droughts: Increased Risk under Global Warming? *Earth Sci. Rev.* **2020**, *201*, 102953. [CrossRef]
12. Xie, W.; Yi, S.; Leng, C. Impacts of Gauge Data Bias on the Performance Evaluation of Satellite-Based Precipitation Products in the Arid Region of Northwestern China. *Water* **2022**, *14*, 1860. [CrossRef]
13. Aadhar, S.; Mishra, V. High-Resolution near Real-Time Drought Monitoring in South Asia. *Sci. Data* **2017**, *4*, 170145. [CrossRef] [PubMed]
14. Toté, C.; Patricio, D.; Boogaard, H.; Van der Wijngaart, R.; Tarnavsky, E.; Funk, C. Evaluation of Satellite Rainfall Estimates for Drought and Flood Monitoring in Mozambique. *Remote Sens.* **2015**, *7*, 1758–1776. [CrossRef]
15. Shukla, S.; McNally, A.; Husak, G.; Funk, C. A Seasonal Agricultural Drought Forecast System for Food-Insecure Regions of East Africa. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 3907–3921. [CrossRef]
16. Bourdin, D.R.; Fleming, S.W.; Stull, R.B. Streamflow Modelling: A Primer on Applications, Approaches and Challenges. *Atmos. Ocean* **2012**, *50*, 507–536. [CrossRef]
17. Deo, R.C.; Şahin, M. Application of the Extreme Learning Machine Algorithm for the Prediction of Monthly Effective Drought Index in Eastern Australia. *Atmos. Res.* **2015**, *153*, 512–525. [CrossRef]
18. Fahimi, F.; Yaseen, Z.M.; El-shafie, A. Application of Soft Computing Based Hybrid Models in Hydrological Variables Modeling: A Comprehensive Review. *Theor. Appl. Climatol.* **2017**, *128*, 875–903. [CrossRef]
19. Rhee, J.; Im, J. Meteorological Drought Forecasting for Ungauged Areas Based on Machine Learning: Using Long-Range Climate Forecast and Remote Sensing Data. *Agric. For. Meteorol.* **2017**, *237*, 105–122. [CrossRef]
20. Bhusal, A.; Parajuli, U.; Regmi, S.; Kalra, A. Application of Machine Learning and Process-Based Models for Rainfall-Runoff Simulation in DuPage River Basin, Illinois. *Hydrology* **2022**, *9*, 117. [CrossRef]
21. Park, S.; Im, J.; Han, D.; Rhee, J. Short-Term Forecasting of Satellite-Based Drought Indices Using Their Temporal Patterns and Numerical Model Output. *Remote Sens.* **2020**, *12*, 3499. [CrossRef]
22. Sundararajan, K.; Garg, L.; Srinivasan, K.; Bashir, A.K.; Kaliappan, J.; Ganapathy, G.P.; Selvaraj, S.K.; Meena, T. A Contemporary Review on Drought Modeling Using Machine Learning Approaches. *Comput. Model. Eng. Sci.* **2021**, *128*, 447–487. [CrossRef]
23. Chen, J.; Li, M.; Wang, W. Statistical Uncertainty Estimation Using Random Forests and Its Application to Drought Forecast. *Math. Probl. Eng.* **2012**, *2012*, e915053. [CrossRef]
24. Dikshit, A.; Pradhan, B.; Alamri, A.M. Short-Term Spatio-Temporal Drought Forecasting Using Random Forests Model at New South Wales, Australia. *Appl. Sci.* **2020**, *10*, 4254. [CrossRef]
25. Park, S.; Im, J.; Jang, E.; Rhee, J. Drought Assessment and Monitoring through Blending of Multi-Sensor Indices Using Machine Learning Approaches for Different Climate Regions. *Agric. For. Meteorol.* **2016**, *216*, 157–169. [CrossRef]
26. Park, H.; Kim, K.; Lee, D. kun Prediction of Severe Drought Area Based on Random Forest: Using Satellite Image and Topography Data. *Water* **2019**, *11*, 705. [CrossRef]
27. Borji, M.; Malekian, A.; Salajegheh, A.; Ghadimi, M. Multi-Time-Scale Analysis of Hydrological Drought Forecasting Using Support Vector Regression (SVR) and Artificial Neural Networks (ANN). *Arab. J. Geosci.* **2016**, *9*, 725. [CrossRef]
28. Achite, M.; Jehanzaib, M.; Elshaboury, N.; Kim, T.-W. Evaluation of Machine Learning Techniques for Hydrological Drought Modeling: A Case Study of the Wadi Ouahrane Basin in Algeria. *Water* **2022**, *14*, 431. [CrossRef]
29. Almikaeel, W.; Čubanová, L.; Šoltész, A. Hydrological Drought Forecasting Using Machine Learning—Gidra River Case Study. *Water* **2022**, *14*, 387. [CrossRef]
30. Bhusal, A.; Thakur, B.; Kalra, A.; Benjankar, R.; Shrestha, A. Evaluating the Effectiveness of Best Management Practices in Adapting the Impacts of Climate Change-Induced Urban Flooding. *Atmosphere* **2024**, *15*, 281. [CrossRef]

31. Mozgovoy, D.K. Monitoring of the Droughts Consequence by High Resolution Satellite Images. *Ecol. Noospherology* **2016**, *27*, 89–95. [CrossRef]
32. Western Regional Climate Center. Available online: https://wrcc.dri.edu (accessed on 24 March 2024).
33. MacDonald, G.M. Severe and sustained drought in southern California and the West: Present conditions and insights from the past on causes and impacts. *Quat. Int.* **2007**, *173–174*, 87–100. [CrossRef]
34. Berg, N.; Hall, A. Increased interannual precipitation extremes over California under climate change. *J. Clim.* **2015**, *28*, 6324–6334. [CrossRef]
35. News, A.B.C. At California's Folsom Lake, a Stark Image of State's Drought Disaster. Available online: https://abcnews.go.com/US/californias-folsom-lake-stark-image-states-drought-disaster/story?id=78209909 (accessed on 18 August 2023).
36. Home | Drought.Gov. Available online: https://www.drought.gov/ (accessed on 24 March 2024).
37. Naresh Kumar, M.; Murthy, C.S.; Sesha Sai, M.V.R.; Roy, P.S. On the Use of Standardized Precipitation Index (SPI) for Drought Intensity Assessment. *Meteorol. Appl.* **2009**, *16*, 381–389. [CrossRef]
38. Zargar, A.; Sadiq, R.; Naser, B.; Khan, F.I. A Review of Drought Indices. *Environ. Rev.* **2011**, *19*, 333–349. [CrossRef]
39. Ghimire, A.B.; Faruk, O.; Shadia, N.; Parajuli, U.; Shin, S. Correlation of Drought Indices with Climatic and Socio-Economic Factors in San Diego, USA. *J. Environ. Eng. Sci.* **2023**, *19*, 120–131. [CrossRef]
40. Hayes, M.; Svoboda, M.; Wall, N.; Widhalm, M. The Lincoln Declaration on Drought Indices: Universal Meteorological Drought Index Recommended. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 485–488. [CrossRef]
41. Fitchett, J. On Defining Droughts: Response to 'The Ecology of Drought—A Workshop Report'. *S. Afr. J. Sci.* **2019**, *115*, 1. [CrossRef]
42. Shukla, S.; Wood, A.W. Use of a Standardized Runoff Index for Characterizing Hydrologic Drought. *Geophys. Res. Lett.* **2008**, *35*, 2. [CrossRef]
43. Shamshirband, S.; Hashemi, S.; Salimi, H.; Samadianfard, S.; Asadi, E.; Shadkani, S.; Kargar, K.; Mosavi, A.; Nabipour, N.; Chau, K.-W. Predicting Standardized Streamflow Index for Hydrological Drought Using Machine Learning Models. *Eng. Appl. Comput. Fluid Mech.* **2020**, *14*, 339–350. [CrossRef]
44. Lai, C.; Zhong, R.; Wang, Z.; Wu, X.; Chen, X.; Wang, P.; Lian, Y. Monitoring Hydrological Drought Using Long-Term Satellite-Based Precipitation Data. *Sci. Total Environ.* **2019**, *649*, 1198–1208. [CrossRef]
45. Dahal, D.; Magar, B.A.; Aryal, A.; Poudel, B.; Banjara, M.; Kalra, A. Analyzing Climate Dynamics and Developing Machine Learning Models for Flood Prediction in Sacramento, California. *Hydroecology Eng.* **2024**, *1*, 10003. [CrossRef]
46. Khan, N.; Shahid, S.; Ahmed, K.; Ismail, T.; Nawaz, N.; Son, M. Performance Assessment of General Circulation Model in Simulating Daily Precipitation and Temperature Using Multiple Gridded Datasets. *Water* **2018**, *10*, 1793. [CrossRef]
47. Scharffenberg, W.A.; Fleming, M.J. *Hydrologic Modeling System-HEC-HMS-User's Manual*, Version 2.0; US Army Corps of Engineers Hydrologic Engineering Center: Davis, CA, USA, 2016.
48. Mockus, V. Estimation of Direct Runoff from Storm Rainfall. *Chapter* **1972**, *10*, 79.
49. Jyolsna, P.; Kambhammettu, B.V.N.; Gorugantula, S. Application of Random Forest and Multi Linear Regression Methods in Downscaling GRACE Derived Groundwater Storage Changes. *Hydrol. Sci. J.* **2021**, *66*, 874–887. [CrossRef]
50. Grömping, U. Variable Importance Assessment in Regression: Linear Regression versus Random Forest. *Am. Stat.* **2009**, *63*, 308–319. [CrossRef]
51. Thai, N.T. Building Early Drought Forecasting Model in the Dak Dak Province Using Machine Learning Algorithms. *IOP Conf. Ser. Earth Environ. Sci.* **2023**, *1170*, 012002. [CrossRef]
52. Sadri, S.; Burn, D.H. Nonparametric Methods for Drought Severity Estimation at Ungauged Sites. *Water Resour. Res.* **2012**, *48*. [CrossRef]
53. Ghimire, A.; Banjara, M.; Bhusal, A.; Kalra, A. Evaluating the Effectiveness of Low Impact Development Practices against Climate Induced Extreme Floods. *Int. J. Environ. Clim. Chang.* **2023**, *13*, 288–303. [CrossRef]
54. Belayneh, A.; Adamowski, J. Drought Forecasting Using New Machine Learning Methods. *J. Water Land Dev.* **2013**, *18*, 3–12. [CrossRef]
55. Huang, W.-R.; Liu, P.-Y.; Hsu, J. Multiple Timescale Assessment of Wet Season Precipitation Estimation over Taiwan Using the PERSIANN Family Products. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *103*, 102521. [CrossRef]
56. Vernimmen, R.R.E.; Hooijer, A.; Mamenun; Aldrian, E.; van Dijk, A.I.J.M. Evaluation and Bias Correction of Satellite Rainfall Data for Drought Monitoring in Indonesia. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 133–146. [CrossRef]
57. Gao, F.; Zhang, Y.; Ren, X.; Yao, Y.; Hao, Z.; Cai, W. Evaluation of CHIRPS and Its Application for Drought Monitoring over the Haihe River Basin, China. *Nat. Hazards* **2018**, *92*, 155–172. [CrossRef]
58. Hinge, G.; Mohamed, M.M.; Long, D.; Hamouda, M.A. Meta-Analysis in Using Satellite Precipitation Products for Drought Monitoring: Lessons Learnt and Way Forward. *Remote Sens.* **2021**, *13*, 4353. [CrossRef]
59. Trambauer, P.; Maskey, S.; Winsemius, H.; Werner, M.; Uhlenbrook, S. A Review of Continental Scale Hydrological Models and Their Suitability for Drought Forecasting in (Sub-Saharan) Africa. *Phys. Chem. Earth Parts A/B/C* **2013**, *66*, 16–26. [CrossRef]
60. Xing, Z.; Ma, M.; Su, Z.; Lv, J.; Yi, P.; Song, W. A Review of the Adaptability of Hydrological Models for Drought Forecasting. *Proc. IAHS* **2020**, *383*, 261–266. [CrossRef]

61. Jehanzaib, M.; Shah, S.A.; Son, H.; Jang, S.-H.; Kim, T.-W. Predicting Hydrological Drought Alert Levels Using Supervised Machine-Learning Classifiers. *KSCE J. Civ. Eng.* **2022**, *26*, 3019–3030. [CrossRef]
62. Kazakis, N.; Karakatsanis, D.; Ntona, M.M.; Polydoropoulos, K.; Zavridou, E.; Voudouri, K.A.; Busico, G.; Kalaitzidou, K.; Patsialis, T.; Perdikaki, M.; et al. Groundwater Depletion. Are Environmentally Friendly Energy Recharge Dams a Solution? *Water* **2024**, *16*, 1541. [CrossRef]