

## Article

# Feature-Weighting-Based Prediction of Drought Occurrence via Two-Stage Particle Swarm Optimization

Karpagam Sundararajan <sup>1</sup> and Kathiravan Srinivasan <sup>2,\*</sup> 

<sup>1</sup> School of Information Technology and Engineering, Vellore Institute of Technology, Vellore 632014, India

<sup>2</sup> School of Computer Science and Engineering, Vellore Institute of Technology, Vellore 632014, India

\* Correspondence: kathiravan.srinivasan@vit.ac.in

**Abstract:** Drought directly affects environmental sustainability. Predicting the drought at the earliest opportunity will help to execute drought mitigation plans. Several drought indices are used to predict the severity of drought across different geographical regions. The two main drought indices used in India for meteorological drought are the standardized precipitation index (SPI) and standardized precipitation evapotranspiration index (SPEI). This work is a study to find the ability of above mentioned indices to predict meteorological drought for the state of Tamil Nadu using 62 years of data. The prediction results are evaluated using the performance metrics of precision, recall, f1 score, Matthews correlation coefficient, and accuracy. The dataset is severely imbalanced due to the low number of drought incidence years. Hence there exists a tug of war between precision and recall, so for improving it without affecting one another, a multi-objective optimization process is applied. The prediction performance is improved by using the filter-global-supervised feature weighting and wrapper-global-supervised feature weighting techniques. In the filter-based feature weighting approach, the information gain measure and Pearson correlation coefficient are used as feature weights. For the wrapper-based feature weighting approach, two-stage particle swarm optimization (PSO) is designed to calculate the weights of the features, and the random forest is used as the classifier. This two-stage PSO constructs the best population set for individual objectives and then searches around it to find the best particle so that the multiple contradicting objectives will converge into the best solution easier. When compared to classification without feature weighting, two-stage PSO feature weighting achieves a 45% improvement in precision. However, only a moderate improvement in recall is obtained. According to the findings, SPI3 and SPEI12 should be given more weightage in metrological drought prediction.

**Keywords:** feature weighting; multi-objective optimization; particle swarm optimization (PSO); standardized precipitation index (SPI); standardized precipitation evapotranspiration index (SPEI)



**Citation:** Sundararajan, K.; Srinivasan, K. Feature-Weighting-Based Prediction of Drought Occurrence via Two-Stage Particle Swarm Optimization. *Sustainability* **2023**, *15*, 929. <https://doi.org/10.3390/su15020929>

Academic Editors: Saqib Iqbal Hakak and Thippa Reddy Gadekallu

Received: 6 December 2022

Revised: 23 December 2022

Accepted: 30 December 2022

Published: 4 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

### 1.1. Drought Basics

Drastic climate changes are the major threat to humans in the forthcoming years. Excessive rainfall and excessive heat are recurrently affecting our mother earth nowadays. There exists a need for a better drought monitoring system in India. Machine learning and deep learning algorithms have proved their efficiency in providing solutions for a lot of applications in domains such as space, medicine, agriculture, and many more. Most of the research works done in drought indices prediction use their past values or a combination of other drought indices or climatic indicators [1].

#### 1.1.1. Drought

A drought can be defined as “an extended period—a season, a year, or several years—of deficient precipitation compared to the statistical multi-year average for a region that results in a water shortage for some activity, group, or sector”.

### 1.1.2. Impact of Drought on Environment Sustainability

Drought affects water sources, which is followed by the destruction of the plant community, aquatic community and society. Agriculture plays a pioneering role in the Indian economy. An understanding of extreme climate is needed to help the farming community in planning adaptation strategies. The drought that occurred in Tamilnadu in the year 1982, caused huge losses and even affected the moisture-surplus regions of the state, such as the Nilgiri hills, and resulted in the destruction of more than 6000 hectares of tea plantation. The hydropower generation failed in the Mettur reservoir. In the 1987 drought, the groundwater level fell steeply, up to 11 m. Nearly 77 crores were spent for relief measures.

### 1.1.3. Climatic Indicator

Climatic indicators are variables or parameters used to describe climate conditions. The World Metrological Organization had grouped the climatic indicators under four categories, namely, temperature and energy, atmospheric composition, ocean and water, and cryosphere, e.g., precipitation, temperature, streamflow, groundwater and reservoir levels, soil moisture, and snowpack.

### 1.1.4. Drought Indices

Indices are typically computed numerical representations of drought. A drought index assimilates thousands of data on rainfall, snowpack, streamflow, and other water-supply indicators into a comprehensible picture, Ref. [2], e.g., the standardized precipitation index (SPI), standardized precipitation evapotranspiration index (SPEI), PDSI (Palmer drought severity index, and aridity anomaly index (AAI).

### 1.1.5. Standardized Precipitation Index (SPI)

SPI is the most commonly used indicator worldwide for detecting and characterizing meteorological droughts. It was developed by McKee et al., (1993). SPI indicates the precipitation conditions in a specific period within a long time series, by calculating the statistical distribution of rainfall accumulations for defined intervals (e.g., daily, 1 month, 3 months) over a long time series (at least 30 years), To calculate the SPI, first, the precipitation data are fitted to the gamma probability distribution function and then transformed into a standard normal distribution. The strength of SPI is that it uses precipitation only and can characterize both wetness and dryness. It is comparable across different regions and less complex than most of the drought indices. The disadvantage is that it does not account for evapotranspiration and the impacts of runoff. Table 1 lists the SPI value and the corresponding drought severity level [3].

**Table 1.** List of SPI/SPEI values and their respective drought severity level.

S. No.	SPI/SPEI Value	Drought Severity Level
1	Greater than 2	Extremely wet
2	1.5 to 2	Very wet
3	1 to 1.4	Moderately wet
4	−1 to 0.9	Normal
5	−1.5 to −1.1	Moderately dry
6	−2 to −1.6	Severely dry
7	values below −2	Extremely dry

### 1.1.6. Standardized Precipitation Evapotranspiration Index (SPEI)

The SPEI is an extension of SPI which considers the potential evapotranspiration (PET). It measures normalized anomalies in precipitation minus PET. The SPEI drought index

was developed by Vicente-Serrano et al., (2012) The strength of the SPEI is that it uses information about evapotranspiration and the output applies to all climate regimes. The limitation is that more data than precipitation are required and its sensitivity to the method used to calculate PET. Table 1 lists the SPEI value and the corresponding drought severity level [4].

### 1.2. Feature Weighting

In feature selection, whether the feature has relevance or not with the target feature and the prediction model is studied. Thus it will be either included or removed in the prediction process. However, with feature weighting, the magnitude of influence or the importance that the feature would have with respect to the target feature and the prediction model is calculated. This importance score decides how much magnitude of the feature value has to be used in the prediction. Instead of creating a do-or-die situation for indices with feature selection, their partial usefulness is taken into the prediction task [5].

#### Feature Weighting Types

Iratxe [6] has proposed the taxonomy for feature weighting methods as follows:

- (a) Supervised and unsupervised approach;
- (b) Filter and wrapper based approach;
- (c) Local and global based approach.

Depending on the learning approach followed for the classification, the feature weighting can be categorized as supervised and unsupervised.

In the filter-based approach, the strength of the features is studied without any dependency on the classification algorithm. Weightage is given, depending on the general character of the feature. The popular filter-based measures are information gain, chi-square, relief, Pearson correlation coefficient and Fisher score [7].

In the wrapper-based approach, the feature importance is measured by using the performance of the classification task with a specific classification algorithm. The feature importance scores are adjusted based on the feedback of performance measures returned by that specific classification algorithm. The time for execution and the complexity is high.

In the global weighting approach, irrespective of the class label, the weight is allocated to all samples in the feature. In the local weighting approach, the samples are grouped based on the class label and then the feature weights are applied.

In implementing the feature weighting methodology, first, the learning approach needed for the problem statement is determined, and next, it is determined whether the global weighting is enough or there is a need to do local feature weighting. Next, it is decided what feedback metric is needed, either filter-based or wrapper based. Dataset  $D$  is considered with  $X$  samples and  $Y$  features. Let  $D_{ij}$  represents the database value of the  $i$ th sample and  $j$ th feature. The weight calculated for Feature1, Feature2 and Feature  $Y$  are  $w_1$ ,  $w_2$  and  $w_Y$  respectively. The weighted dataset ( $D'$ ) is calculated by multiplying the sample values  $D_{ij}$  with weight  $w_j$ .

### 1.3. Wrapper-Based Measure

In wrapper-based measure, feature weights are determined to work with a specific machine-learning algorithm, so the generalization is missing. The performance metrics of the algorithm are used as the evaluation metric of the wrapper algorithm. The performance metrics for classification are precision, recall, and accuracy; similarly for regression, mean square error, root mean square error, and mean absolute error are used [8].

#### Particle Swarm Optimization (PSO)

The optimization problem refers to the selection of the best solution from the set of feasible solutions. In feature weighting, the algorithms have to work on the continuous value space. The weight for the feature has to be determined to maximize the prediction performance.

Stochastic means involving chance or probability. Stochastic optimization can work on uncertainties using its randomness search nature. It is more suitable for problems with high dimensions that are nonlinear in nature. A random nature is the trump card to escape from the local optima and find the global optimum.

Metaheuristic algorithms have a stochastic nature, and the advantages of stochastic algorithms are they avoid premature convergence and do a wide explorative and efficient search of the solution space [9].

The bioinspired optimization algorithm was developed by studying the behavior of animals in their groups to work and achieve their task. Consider the task of searching for food. In this case, this food is our optimized result and their searching strategy is the optimization algorithm steps. Effective solutions have been achieved for optimization problems in science and engineering domains using bioinspired algorithms [10].

PSO is a metaheuristic stochastic bioinspired optimization algorithm. The PSO algorithm was proposed in the year 1995 by Eberhart and Kennedy [11], is a population-based algorithm. There is no recombination or mutation between the particles to generate a population. It starts from a population set created at random and then it carries out the self-adjustment in the solution space to find the global optimal solution. The standard PSO algorithm uses Equation (1) to calculate particle position and Equation (2) to calculate velocity and find:

$$X(t+1) = X(t) + V(t+1) \quad (1)$$

$$V(t+1) = V(t) + c1 * rand() * (Xpbest - X(t)) + c2 * rand() * (Xgbest - X(t)) \quad (2)$$

$V(t)$ : Velocity of the particle at time 't'

$X(t)$ : Particle position at time 't'

$c1, c2$ : Learning factor or accelerating factor

$rand()$ : Uniformly distributed random number between 0 and 1

$Xpbest$ : Particle's best position

$Xgbest$ : Global best position

From its inception until now, the algorithm has been under modification and enhancement, but still, there are a sea of opportunities available to enhance the algorithm [12]. The algorithm needs improvement in the following categories

- Selection of the parameter values of inertia weight,  $c1, c2$ ;
- Topology choices;
- Learning strategy improvements;
- Modifying position and velocity update rule;
- Binary and multi-objective optimization;
- Combining with other optimization algorithms.

The term inertia weight was introduced in the year 1998 by Shi and Eberhart [13] to bring in a balance between the global search completed at the early stage and the local search needed for fine-tuning at later stages. Equation (3) holds the inertia weight in the particle velocity calculation.

$$V(t+1) = w * V(t) + c1 * rand() * (Xpbest - X(t)) + c2 * rand() * (Xgbest - X(t)) \quad (3)$$

The linearly decreasing weight value from 0.9 to 0.4 is the commonly followed method. The setting up of  $w$  with the fuzzy system or with random initialization is also suggested.

#### 1.4. Filter-Based Measures

The filter-based measures evaluate the feature's strength and its bonding with the output/target feature and does not include the effect of the classification algorithm [14].

The filter-based measures used in this paper are the information gain and Pearson correlation coefficient.

#### 1.4.1. Information Gain (IG)

Information gain is a measure describing how wealthy information a feature has about the target variable. Entropy plays a major role in the information gain measure. Entropy describes the degree of randomness, uncertainty, and impurity in the data. It describes the probability distribution of data with respect to each class label in the dataset. If it is a skewed distribution, then the entropy is lower. If there is a balanced probability distribution, then the entropy is higher. The surprise of an event is based on its entropy.

Consider the dataset  $M$  for a variable  $n$ . The formula to calculate entropy  $H(M)$  for the dataset  $M$  with  $C$  classes is given in Equation (4) and the IG is given in Equation (5)

$$H(M) = - \sum_{i=1}^C p_i \log_2 p_i \quad (4)$$

$$IG(M, N) = H(M) - \sum_{v \in N} \frac{M_v}{M} \cdot H(M_v) \quad (5)$$

$M$ —Target column

$N$ —A column in the dataset for which the entropy is calculated

$v$ —For each value in  $N$

#### 1.4.2. Pearson Correlation Coefficient (PCC)

In studying the nature of feature relevance and redundancy, two types of measure used are consistency and correlation. Correlation evaluates the strength of the feature by means of its relevancy with the target variable and other input features. There should be high relevance with the target variable and low relevance with other input features. The Pearson correlation coefficient is suggested as a good measure in many works. The formula to calculate PCC is given in Equation (6)

$$PCC(X_i, X_j) = \frac{\sum (X_i - X_{\text{mean}})(Y_i - Y_{\text{mean}})}{\sqrt{\sum (X_i - X_{\text{mean}})^2 \sum (Y_i - Y_{\text{mean}})^2}} \quad (6)$$

$X_i$ —Value of the  $i$ th index  $X$  variable.

$Y_i$ —Value of the  $i$ th index  $Y$  variable.

$X_{\text{mean}}$ —Mean of values in the variable  $X$

$Y_{\text{mean}}$ —Mean of values in the variable  $Y$

#### 1.5. Contributions to the Work

There are numerous studies on feature selection methods to enhance classification performance, but rather than omitting a feature, one might limit or amplify it depending on its significance. The creation of the wrapper-based feature weighting with PSO was prompted by the power of the PSO algorithm in optimization tasks. Precision and recall are the key performance indicators for a dataset that is imbalanced, but it is difficult to increase both of these metrics, necessitating multi-objective optimization. The main contribution of this work is

- To develop a meteorological drought-occurrence prediction system for the state of Tamil Nadu using a machine learning algorithm and weighted SPI and SPEI, since most of the works are on the indices' prediction only.
- To develop the feature weighting model with the multi-objective PSO algorithm to improve the precision and recall performance of an imbalanced dataset.
- To modify the population initialization methodology for the multi-objective PSO model.

## 2. Literature Review

The previous works in the areas of drought prediction with machine algorithms and the construction of various PSO variants are studied. The review work done on the drought prediction is given in Table 2.

**Table 2.** Review of the use of machine learning techniques in drought.

S.No.	Author	Data Used	Contributions	Shortcomings
1	Mosaad Khadr [15]	Precipitation data, LST, NDVI, and soil moisture.	The short to medium term SPI was forecasted using different types of homogenous hidden Markov models. Real-time multivariate Madden–Julian oscillation (RMM MJO) index-based model predicts well using the random forest. Good in forecasting SPI for various time-series one month ahead. The SPEI1 and SPEI3 prediction using long short-term memory (LSTM) was done. One LSTM Layer and one dense layer are used and with the performance metrics coefficient of determination and root mean square error, the model is evaluated. Spatial analysis is also performed in addition to the statistical measures	The one-month lead time prediction performance of SPI3, SPI6, SPI12 is comparatively lower than other leadtimes. SPI12 prediction shows a minimum RMSE value.
2	Dikshit [16]	SPEI1, SPEI3, hydro-metrological variables, namely, rainfall, maximum temperature, minimum temperature, mean temperature, vapour pressure, and cloud cover.	The features used in SPEI1, SPEI3, SPEI6, and SPEI12 prediction are evaluated using the cross-correlation function and a distributed lag nonlinear model (DLNM) methods. Then the SPEI1–6 forecasting was executed using DLNM, an artificial neural network model and an XGBoost model. The forecasting results are compared and the highest prediction accuracy was returned by XGBoost model.	Experiments were not carried out on longer timescales. Interpretation of the results achieved has to be completed.
3	Zhang [17]	Seasons, metrological factors, and climatic indicators.	Time series decomposition was performed as a preprocessing step in SPEI1, SPEI6, and SPEI12 prediction with generalized additive model (GAM), GAM integration with ensemble empirical mode decomposition, and autoregressive integrated moving average (ARIMA), and forecast quantile regression averaging (fQRA) as classifiers. The fQRA forecast was best at forecasting 12 month scale.	Instead of solar radiation data, the sunshine hour was considered. Making use of tree boosters will give better results.
4	Mathivha [18]	SPEI with PET was calculated using Hagreaves’ and Samani’s temperature-based method.	The SPI3 and SPI6 prediction was performed with ANN and M5P for the Maharastra State of India. The number of SPI of varying timescale combinations suitable for SPI3 and SPI6 are 7 and 4. M5P returns the best RMSE value.	Decomposition does not give great improvements in one-month time scale.
5	Pande Chaitanya [19]	SPI	The SPI and SPEI effectiveness is compared for the mountainous and plateau region present in the northern part of Burundi. The timescales of 2,4,24 and 48 are used. The results show that SPEI is a better measure for meteorological data than SPI.	The metrological drought determination is done based on the SPI value alone.
6	Jean Marie [20]	SPI and SPEI		The author also suggests the usage of multiple drought indices in predicting drought.

Wangil Yang [21] in his work had experimented with linearly decreasing weight, and non-linearly decreasing weight using upward parabola, downward parabola, and exponential formulas, and recommended that the exponential decreasing inertia weights have more convergence precision. The Bayesian technique is used to adjust the inertia weight and with the help of Cauchy mutation, the balance between exploration and exploitation is achieved by Limin Zhang [22]. The inertia weight is updated dynamically with respect to the average particle spacing and a slowly varying function is used in updating the velocity in the work of Liang H. [23].

Wangil yang [21] worked on setting up the acceleration coefficient using the linear methods based on synchronous and asynchronous nature formulas and the nonlinear methods based on the sine and cosine functions. The larger value of  $c_1$  and smaller value of  $c_2$  give good exploration capacity in its earlier stages and its vice versa smaller  $c_1$  and larger  $c_2$  help to catch up with the best global optimum in the later stages. Visakan Kadirkamanathan [24] worked on stability analysis and suggested that 'w' should be between 0.8 and 0.9 and the value of  $c_1 + c_2$  has to be decreased.

The chaos-driven random number generator is used in the velocity update formula with linear decreasing inertia weight, and various chaotic maps such as the dissipative standard map, Lozi map, and Arnold's cat map can be used. The dissipative stand map random generator performed well in avoiding premature convergence in Michal Pluhacek's work [25].

Wei Li worked on multi-population cooperative approaches, and the application of different learning strategies for the different sets of populations based on their fitness value ranking was followed. An additional mutation operator was also used to increase the exploration capability of the population. The two learning strategies used are dynamic segment-based mean learning strategy and multidimensional comprehensive learning strategy [26].

The dynamic learning strategy was developed by Wensing Ye with a prime focus on increasing the information exchange between particles. The particles are divided into two groups, namely, communication and ordinary particles. Ordinary particles complete the exploitation process and the communication particles perform the exploration work. A modification in the velocity formula is done by using the average of best fitness scores in the subswarm [27].

The modifications in Harris Hawk optimization (HHO) were carried out to improve the exploration and exploitation capability. The hybrid variants of Harris Hawk optimization perform well in faster convergence and computational accuracy. HHO proved its efficiency in feature selection also [28].

Dimensionality reduction techniques linear discriminant analysis (LDA) and principal component analysis (PCA) performance on diabetic retinopathy and intrusion detection datasets are studied with four machine learning algorithms namely decision tree induction, support vector machine (SVM), naive Bayes classifier and random forest. PCA with random forest and SVM is best for the intrusion detection dataset. For the diabetic retinopathy dataset, the performance is decreased with dimensionality reduction, and hence both LDA and PCA fail. The author suggests performing it in bigger datasets such as images and text [29].

### 3. Materials and Methods

#### 3.1. Data Used

The global SPI and SPEI drought indices data were downloaded in Netcdf file (.nc) format. SPI source was the Research Data Archive (RDA) of the National Center for Atmospheric Research (NCAR) in Boulder, Colorado [30]. NCAR used precipitation data (1949–2012) from the original University of East Anglia Climate Research Unit (CRU) global time series, TS3.21, 1.0 degree by 1.0 degree grids. The source of SPEI was the Higher Council for Scientific Research (CSIC). Ref. [31] CSIC utilized the monthly precipitation and potential evapotranspiration data from the CRU (current version 4.05 of the CRU TS dataset).

The SPI considers only the precipitation, but the SPEI considers the evapotranspiration factor also. The shape file for the state of Tamilnadu was downloaded from the India Administrative webpage. Using python Xarray, the set of drought indices values for the region of interest (ROI) specified by the shape file of Tamilnadu alone was transferred to a .csv file. For each month, the average of the value of the ROI was calculated and used. The time period taken for the study was from 1950 to 2012. The metrological drought years of Tamilnadu as given in Drought Network News by K.K.Nathan [32], Indian Agricultural Research Institute are 1951, 1965, 1966, 1967, 1972, 1979, 1980, 1982, 1983, 1986, 1987, 1989, 2002, 2009, and 2017. Features used for the drought occurrence prediction in this work are monthly SPI3, SPI6, SPI12, SPEI3, SPEI6, and SPEI12 values from 1950 to 2012. The class labels are drought and non-drought. The drought years are marked as 1 and the non-drought as 0.

### 3.2. Methodology

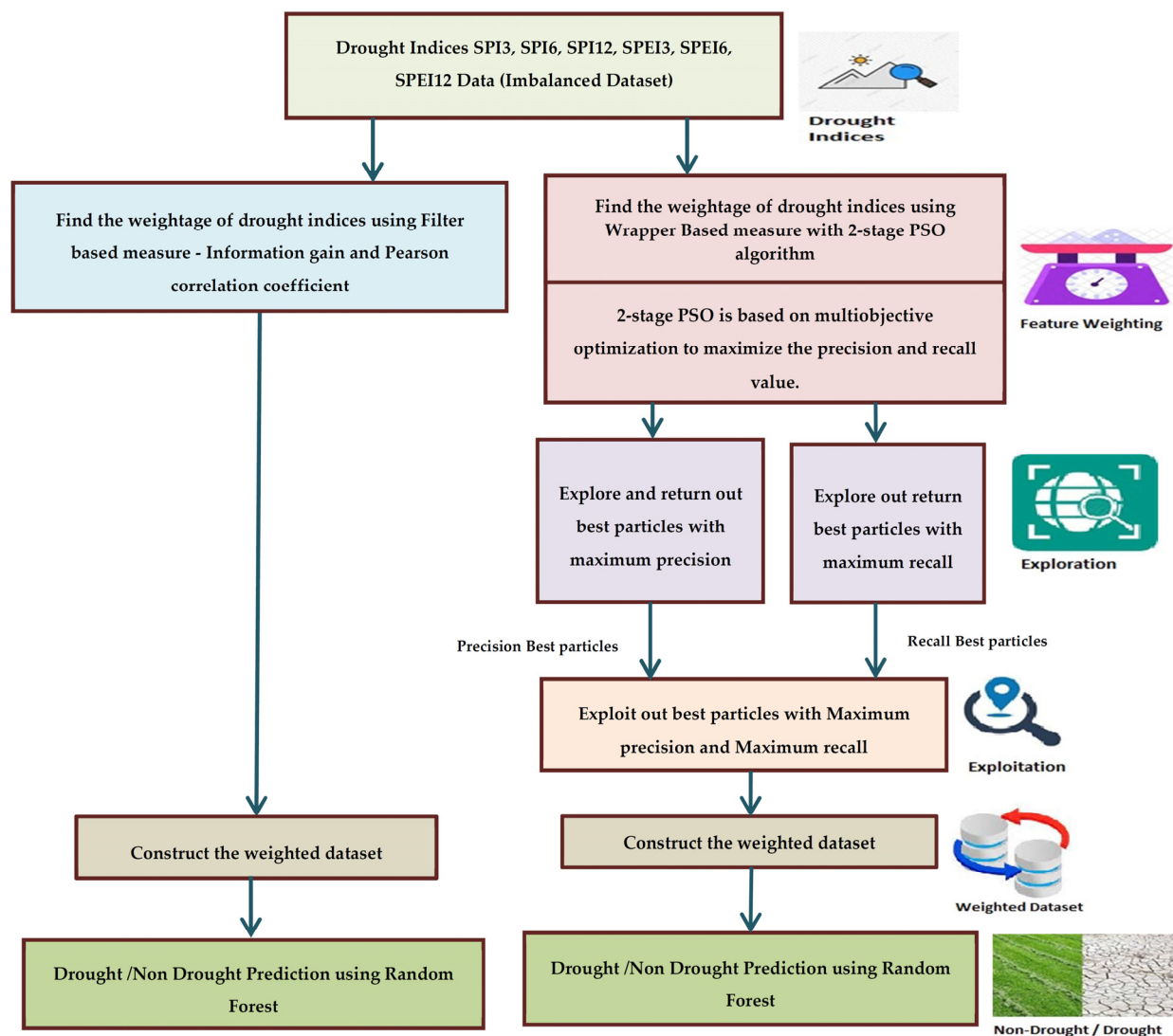
The drought indices used in this work are recognized as the best drought indicators by the drought forecasting department of various countries.

These indices have been used for the past 50 years. However, depending on the geographical location, the importance and power of these indices can vary. Hence there is a need to find the importance of the information these indices provide in drought forecasting of our geographical locations. In this work, the drought occurrence was predicted with the help of the two drought indices SPI and SPEI, their month-wise data with varying timescales of 3, 6, and 12 was used. Most of the research works are on the usage of machine learning techniques to predict the drought indices from their past values and with some additional information from other indices and climate indicators. However, there is not much work on drought occurrence prediction. The drought occurrence years are minimum over the time period taken for study, which makes the dataset a highly imbalanced one. Hence there is a need for special techniques to improve the precision and recall in an imbalanced dataset [33]. The monthly SPI 3, SPI 6, SPI 12, SPEI 3, SPEI 6 and SPEI 12 are used as input variables and the drought occurrence is the predicted target variable. The drought occurrence prediction without any feature weighting, with wrapper-global-supervised feature weighting [34], and with filter global supervised approaches were implemented and their results were compared.

The wrapper global supervised approach makes use of random forest and the filter global supervised approach was implemented with the filter-based measures information gain and Pearson correlation coefficient. The drought prediction framework developed and the overall work done in this paper are depicted in Figure 1.

#### 3.2.1. Global-Filter-Supervised Feature Weighting

The information gain values for the input features (drought indices) SPI 3, SPI 6, SPI12, SPEI 3, SPEI 6, and SPEI 12 were calculated and normalized to a 0 to 1 scale. These values were used as the weight of the feature and the weighted dataset was constructed as given in Algorithm 1. Using the random forest, the drought classification was carried out on the weighted dataset. The drought classification performance before and after the feature weighting was calculated. The same process was repeated with weight calculated using the Pearson correlation coefficient.



**Figure 1.** Drought prediction Framework with weighted Features.

### 3.2.2. Global-Wrapper-Supervised Feature Weighting

In the global-wrapper-supervised method, the feature weight is found using the two-stage PSO algorithm. Using the calculated weight, the weighted dataset was constructed as given in Algorithm 1. The random forest was the evaluator and the feedback was received, and the weight values were adjusted accordingly.

### 3.3. Two-Stage PSO Algorithm (PSO 2-Stage)

The two-stage PSO aims to enhance the multi-objective optimization by tuning the population (particles) from one that is objective-specific into one meeting the multiple objectives. The multi-objective optimization means that their objectives/goals are contradicting each other [35], hence solving the precision–recall tradeoff with multi-objective optimization was performed in this paper. When the multi-objective optimization module works on a random population, it struggles to converge to the best solution. So instead of passing a random population directly into the multi-objective optimization module, a fine-tuned population was fed into the multi-objective optimization module. For that, the populations best serving the individual objectives were explored. From this fine-tuned population the best particles satisfying the contradictory goals and achieving the multi-objective optimization were exploited.

**Algorithm 1.** Pseudocode for weighted dataset construction**Input:** Features  $F = [\text{Month}, \text{SPI } 3, \text{SPI } 6, \text{SPI } 12, \text{SPEI } 3, \text{SPEI } 6, \text{SPEI } 12]$ Weight of features  $W = [w_1, w_2, w_3, w_4, w_5, w_6, w_7]$ **Output:**  $F'_{ij}$  represents the weighted feature value of the  $i$ th feature for the  $j$ th month with weight  $w_i$ **Auxiliary Variables:**' $i$ ' represents Features numbering from 1 to 7' $j$ ' represents Months from Jan-1950 to Dec-2012 $F_{ij}$  represents the feature value of the  $i$ th feature for the  $j$ th month $W_i$  represents the weight calculated for the  $i$ th feature**Initialization:** $IG[i]$  = Information gain value for Feature in  $F[i]$  $PCW[i]$  = Pearson correlation coefficient for the Feature in  $F[i]$  $PSOWT[i]$  = Particle position values, given by PSO 2-stage for the Feature in  $F[i]$ **Begin:**WeightedDatasetConstruction( $W[i]$ ) {

1. **If**  $F_{ij} > 0$
2.      $F'_{ij} = F_{ij} \times W_i$
3. **Else**
4.      $F'_{ij} = F_{ij} / W_i$
5. **End If**
6. **Do** Drought Classification on  $F'_{ij}$  using RandomForest
7. Calculate performance metrics // Weight Calculation
8. **Return**  $F'$ —weighted dataset
9. **End** WeightedDatasetConstruction( $W[i]$ ) }
10. Call WeightedDatasetConstruction( $IG[i]$ )
11. Call WeightedDatasetConstruction( $PCW[i]$ )
12. Call WeightedDatasetConstruction( $PSOWT[i]$ )

To implement the above idea, the two-stage PSO algorithm was designed. In stage 1 the exploration was carried out for each and every objective independently and the best particles were filtered out separately. In the second stage, the exploitation was carried out starting with the best particles returned for each objective in the exploration stage. This involved more concentration on exploration in stage 1 to find the objective-specific best population and exploitation in stage 2 to find the best particles satisfying multi-objective criteria. Exploration is the capability of searching for the best solutions in the global space. Exploitation is the convergence of the search with optimal solutions. When the spatial diversity is higher, then it is an indication that the exploration has reached the stagnation level and the exploitation can be initiated.

The main challenge in the PSO is in achieving the balance between exploration and exploitation. Ref. [36] This exploration and exploitation balance was achieved by fine tuning the three main controlling parameters, namely, inertia weight ( $w$ ), cognitive ratio ( $c1$ ), social ratio ( $c2$ ), and the position update formula.

The inertia weight was determined based on the iteration and is given in Equation (7)

$$w_i = w_{max} - \left( \frac{w_{max} - w_{min}}{\max\_iteration} \right) * i \quad (7)$$

$W_{max} = 0.9$

$w_{min} = 0.2$

$w_i$  = weight at iteration ' $i$ '

$\max\_iteration$  = Maximum Iteration

The values of  $c1$  encourage the particle to learn on its own and  $c2$  encourages the particle to learn from its neighbours. So having a higher  $c1$  value means the learning can be around the local search, and if the  $c2$  is higher it means the learning is from the global

best. But there should be a balance in the difference between  $c_1$  and  $c_2$  values, they should not be much deviated [37].

In the PSO algorithm, the position update formula contributes more to increasing or decreasing the exploration phase. Ref. [38] Exploration refers to the search for optimal values carried out in the global search space. The exploration should be made efficient by making it search in every nook and corner of the solution space. So, the position update formula of the PSO algorithm plays a major contribution to exploration. In the standard PSO algorithm, the weight factor is not taken into the position update formula, but then some authors stressed the importance of adaptive weight and including the weight in the position update formula. Ref. [39] As suggested by Hao Liu [40], updating the weight with respect to the iteration movement is performed. The formula used for position update during exploration is given in Equation (8). The values of  $c_2$  are kept higher and  $c_1$  lower, to improve exploration.

$$X = wX + (1 - w)V \quad (8)$$

$X$ —Particle Position

$w$ —Inertia weight

$V$ —Velocity of the particle

Performing too much exploitation will make the PSO algorithm suffer from the local premature convergence problem. The best formula for particle new position calculation enhancing the exploitation capability is given in Equation (9).

$$X = X + V \quad (9)$$

The value of  $c_1$  is kept higher and  $c_2$  as lower, to improve exploitation.

### 3.3.1. Exploration Phase (Stage 1 Process)

The stage 1 process flow diagram is given in Figure 2; it is completely dedicated to the exploration of the best particles serving to improve precision and recall separately. Each particle holds the feature weights of the features SPI 3, SPI 6, SPI 12, SPEI 3, SPEI 6, and SPEI 12 as their particle position in  $X$ , and holds particle velocity in  $V$ . The random population initialization was performed and the fitness objective was to improve the precision metric and recall metric independently. The particle position  $X$ , i.e., the feature weights, was updated at each iteration using Equation (8), the weight used in velocity was updated using Equation (7), and the velocity was updated using Equation (3). The  $K$  best particles with higher precision rate were selected. This stage 1 process was repeated for the fitness function with the objective of improving the recall metric. The exploration phase pseudocode is given in Algorithm 2.

### 3.3.2. Exploitation Phase (Stage 2 Process)

The stage 2 process flow diagram is given in Figure 3; its task is to exploit the best solution from the best particles returned from stage 1. The initial population is the best  $K$  particles with higher precision and the best  $K$  particles with higher recall returned by stage 1. The fitness function is multi-objective, which aims to improve precision and recall. The particle position  $X$  (i.e., the feature weights) were updated at each iteration using Equation (9), the weight used in velocity was updated using Equation (7) and the velocity was updated using Equation (3). The pseudocode for stage 2 is given in Algorithm 3.

**Algorithm 2.** Pseudocode for PSO Stage 1**Input:** S, D, Max\_Iter, Particle P, Velocity V, Position X**Output:** Position X =  $[X_1, X_2 \dots X_D]$  //Position—Represents the weightage of the features**Auxiliary Variables:** Iter = 0, fval\_precision, Precision\_Above75, fval\_Recall, Recall\_Above1, pBest, gBest**Initialization:** S = Population Size, D = Dimension, Max\_Iter = Maximum Iteration, Initialize the particles P with a random population  $P = [P_1, P_2, P_3 \dots P_S]$ , Velocity  $V = [V_1, V_2, V_3 \dots V_D]$ , Position  $X = [X_1, X_2 \dots X_D]$ **Begin: PSO Stage 1 Algorithm**

// Construction of particle population with the best precision

```

1. While (Iter < Max_Iter)
2.   For each particle  $P_i$ 
3.     fval_precision = Find_fitnessPrecision( $P_i$ )
4.     If fval_precision > pBest
5.       pBest = fval_precision
6.     End If
7.     If fval_precision > 0.75
8.       Precision_Above75[ ] =  $P_i$ 
9.     End If
10.  End For
11.  gBest = Max (fval_precision of  $P_i$  in P)
12.  For each particle  $P_i$  in P
13.    Find weight 'w' using Equation (7)
14.    Calculate particle velocity using Equation (3)
15.    Update particle position according to Equation (8)
16.  End For
17.  Increment Iter
18. End While
19. // Construction of particle population with best recall
20. Initialize the particles with random population
21. While (Iter < Max_Iter)
22.   For each particle  $P_i$ 
23.     fval_Recall = Find_fitnessRecall( $P_i$ )
24.     If (fval_Recall > pBest)
25.       pBest = fval_Recall
26.     End If
27.     If (fval_Recall > 0.1)
28.       Recall_Above1[ ] =  $P_i$ 
29.     End If
30.  End For
31.  gBest = Max (fval_Recall of  $P_i$  in P)
32.  For each particle  $P_i$  in P
33.    Find weight 'w' using Equation (7)
34.    Calculate particle velocity 'V' using Equation (3)
35.    Update particle position 'X' according to Equation (3)
36.  End For
37.  Increment Iter
38. End While
39. Return Precision_Above75, Recall_Above1

```

**End PSO Stage 1 Algorithm**

The global-wrapper-supervised feature weighting algorithm is the fitness function. The fitness functions used for the PSO 2-stage algorithm were Find\_fitnessPrecision( $P_i$ ), Find\_fitnessRecall( $P_i$ ) and Find\_fitness, and their task was to construct the weighted dataset as given in Figure 1 for particle  $P_i$ , carry out classification, calculate performance, and return the required performance metric. The global best particle holds the best optimal weights. Based on the precision and recall value achieved by random forest, the particle's new position was set and the process was repeated for maximum iterations.

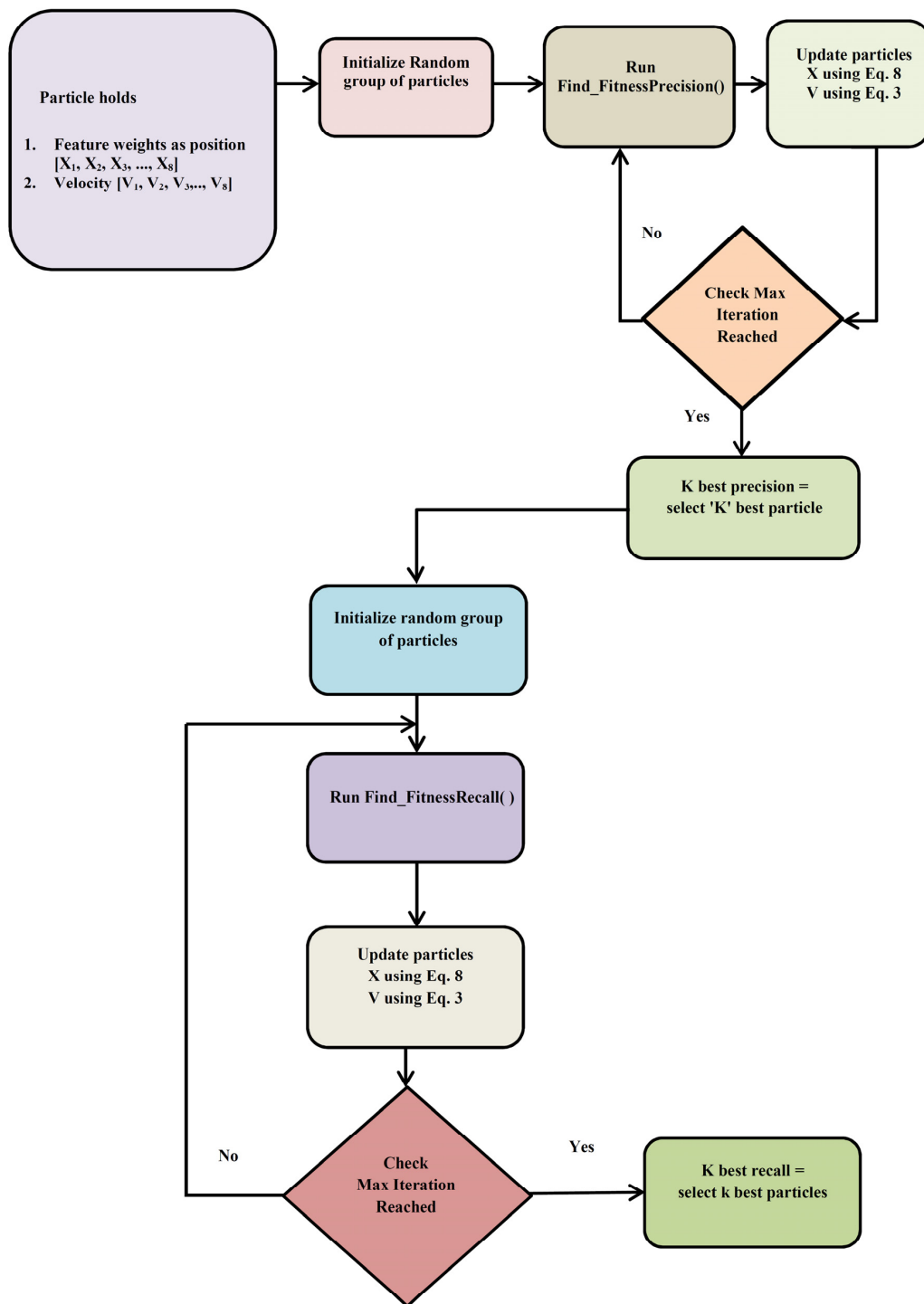


Figure 2. Stage 1 process flow diagram.

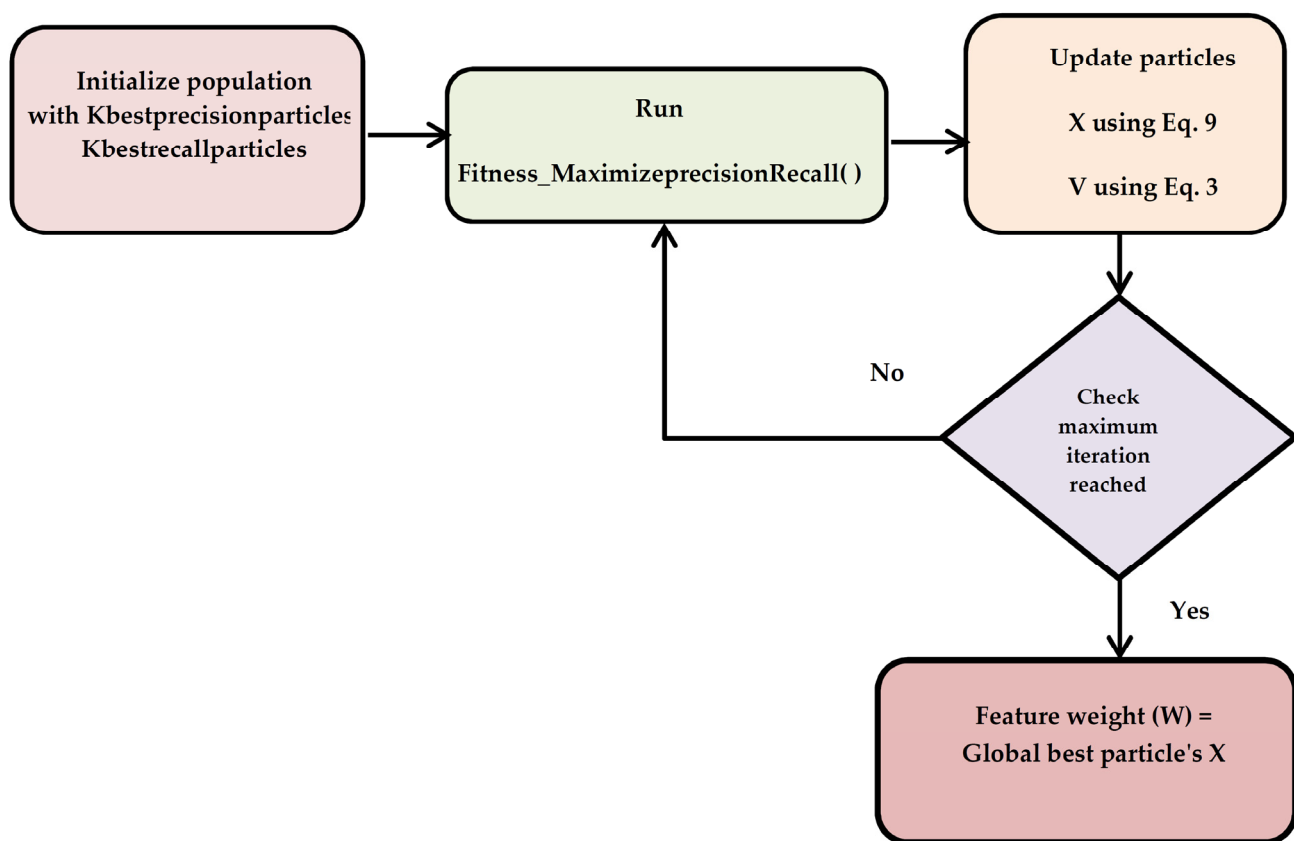


Figure 3. Stage 2 process flow diagram.

**Algorithm 3.** Pseudocode for PSO Stage 2

**Input:** Initial Population P, Velocity V, Position X

**Output:** Position  $X = [X_1, X_2 \dots X_D]$  //Position—Represents the weightage of the features

**Auxiliary Variables:** pBest\_Precision, pBest\_Recall, gBestparticle

**Initialization:** Initial Population  $P = [\{precision\_Above75\} \cup \{Recall\_Above1\}]$ , Iter = 0, Position  $X = [X_1, X_2 \dots X_D]$ , Velocity  $V = [V_1, V_2, V_3 \dots V_D]$

**Begin: PSO Stage 2 Algorithm**

1. **While** (Iter < Max\_Iter)
2.     **For** each particle  $P_i$  in P
3.         fval[] = Find\_fitness( $P_i$ )     //Returns Precision and Recall
4.         **If** fval[precision] > pBest\_precision\_or fval[recall] > pBest\_recall
5.             pBest\_precision = fval[precision]
6.             pBest\_recall = fval[recall]
7.         **End If**
8.     **End For**
9.     gBestparticle =  $P_i$  with Max (fval[precision] and fval[recall])
10.    **For** each particle  $P_i$  in P
11.       Find weight 'w' using Equation (7)
12.       Calculate particle velocity using Equation (3)
13.       Update particle position according to Equation (9)
14.    **End For**
15.    Increment Iter
16. **End While**
17. **Return** particle position X of gBestparticle

**End PSO Stage 2 Algorithm**

#### 4. Results and Discussion

For learning the impact of feature weighting techniques on drought occurrence prediction with two top drought indices at different timescales, three experiments were conducted on the dataset. The first experiment was done with the filter-based measure, Pearson correlation coefficient, as feature weight, the second experiment was with an information gain filter measure, and the third was with the wrapper-based measure.

To select the best wrapper algorithm for PSO-based feature weighting, the performance of the classifiers random forest, logistic regression, and naïve Bayesian were evaluated. The classification performance metrics of precision, accuracy, f1score, recall, and Matthew's correlation coefficient obtained for drought and non-drought prediction are given in Table 3. Out of the three classifiers, random forest produced very good results with a Matthew's score of 0.387 and a precision score of 0.88, so it was selected as the wrapper algorithm for two-stage PSO feature weighting. So, the third experiment was performed using random forest as a classifier with the two-stage PSO algorithm for producing the weight values.

**Table 3.** Comparison of performance measures of Wrapper Algorithms of PSO.

Wrapper Algorithms	Drought Precision	Drought Recall	Drought F1Score	Non Drought F1score	Non Drought Precision	Non Drought Recall	Matthews Correlation Coefficient
Naïve Bayesian	0.24	0.18	0.21	0.79	0.76	0.83	0.0046
Logistic Regression	0.50	0.02	0.04	0.86	0.76	0.99	0.0635
Random Forest	0.88	0.20	0.33	0.89	0.81	0.99	0.367

The performance metric precision, f1score, recall and Matthew's correlation coefficient were calculated from the confusion matrix measures—true positive (TP), true negative (TN), false positive (FP) and false negative (FN).

Precision measures place their focus on the effective measurement of false positive minimization. The formula is given in Equation (10).

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (10)$$

For measuring the effectiveness in minimization of false negatives the metric recall was used.

Equation (11) gives the recall formula.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (11)$$

Matthew's correlation coefficient (MCC) is a reliable measure since it is based on all four components of the confusion matrix. The formula is given in Equation (12)

$$\text{MCC} = \frac{\text{TN} \times \text{TP} - \text{FN} \times \text{FP}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \quad (12)$$

The F1score is the single metric that weighs precision and recall in a balanced way. It is the weighted average of precision and recall. The formula is given in Equation (13)

$$\text{F1score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

The threshold value, determining the positive or negative class in prediction plays a key role in increasing or decreasing the false positive and false negative values. Once the false positive increases the false negative decreases and so the precision decreases and recall increases.

The information gain value was calculated using Equation (5) and Pearson correlation coefficient value was calculated using Equation (6). These values were used as feature weights to construct the weighted dataset and were evaluated with random forest as classifier. In the third experiment, using the two-stage PSO algorithm, the weights were produced and the classification performance with this weighted dataset was evaluated with the random forest. The  $c_1$  value was kept at 1.2 and the  $c_2$  value was set at 1.9 to encourage more learning from the neighbours in the stage-1 PSO. In stage two,  $c_1$  was set to 1.9 and  $c_2$  to 1.2.

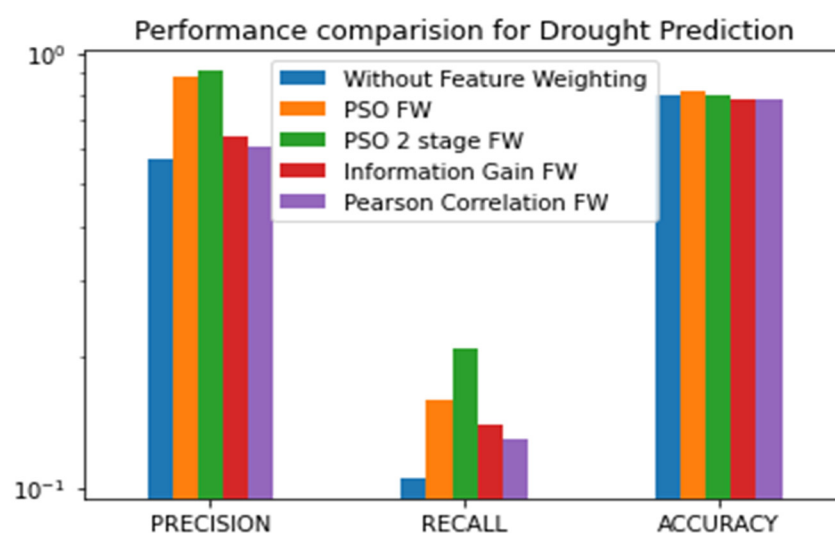
Totally 50 iterations with a population of 10 were used in the experiment. The weight value ( $w$ ) changed with respect to the iteration. The performance of the experimental results was compared with the classification performance without feature weighting and with the standard PSO feature weighting algorithm.

The precision and recall values for drought prediction are 0.55 and 0.1 only without feature weighting. The maximum performance metrics achieved in the 50 iterations are given in Table 4. When the feature weighting was carried out, the maximum precision value reached were 0.64, 0.61, 0.88, and 1.0 for feature weighting schemes with IG, PCC, standard PSO, and two-stage PSO, respectively. The recall values achieved were 0.14, 0.13, 0.16, and 0.2 for feature weighting schemes with IG, PCC, standard PSO, and two-stage PSO, respectively. The f1score and Mathew score is also favourable to two-stage PSO holding scores of 0.37 and 0.38.

**Table 4.** Performance Metrics of drought and Non-drought prediction Comparison.

Method	Drought Precision	Drought Recall	Drought F1Score	Non Drought F1score	Non Drought Precision	Non Drought Recall	Accuracy	Matthews Correlation Coefficient
Without FW	0.55	0.1	0.21	0.89	0.81	0.97	0.79	0.15
Information Gain FW	0.64	0.14	0.18	0.89	0.78	0.97	0.78	0.18
Pearson Correlation FW	0.61	0.13	0.26	0.92	0.82	0.98	0.78	0.32
PSO FW	0.88	0.16	0.33	0.89	0.81	0.99	0.82	0.36
PSO 2-Stage FW	1	0.21	0.37	0.88	0.82	0.99	0.8	0.38

Figure 4 shows the bar chart representation of the performance metrics in drought category prediction and Figure 5 shows the same for non-drought category prediction. Figures 6 and 7 are the precision–recall curves achieved for the classification with PSO FW and two-stage PSO FW methods. The AUC score with PSO weighting is 0.39, but with two-stage PSO feature weighting, the model was able to achieve a 0.5 AUC score.



**Figure 4.** Performance metrics comparison for drought prediction.

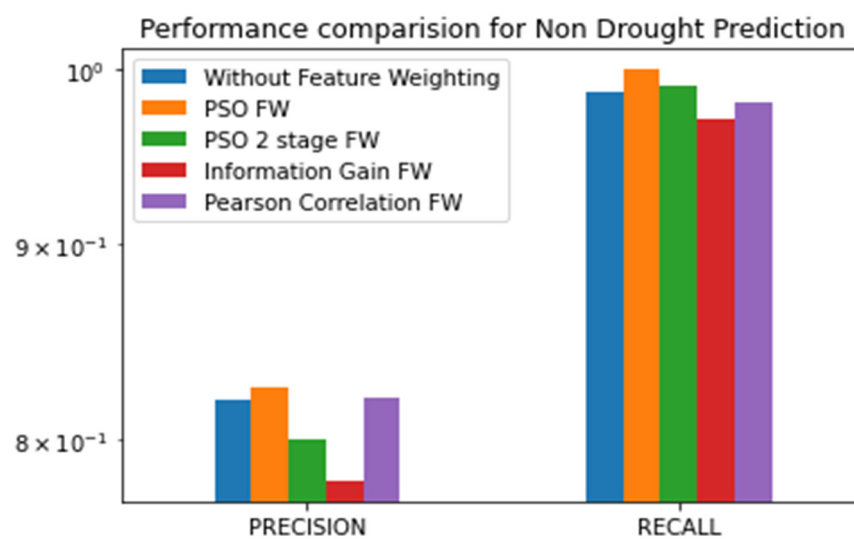


Figure 5. Performance metrics comparison for non-drought prediction.

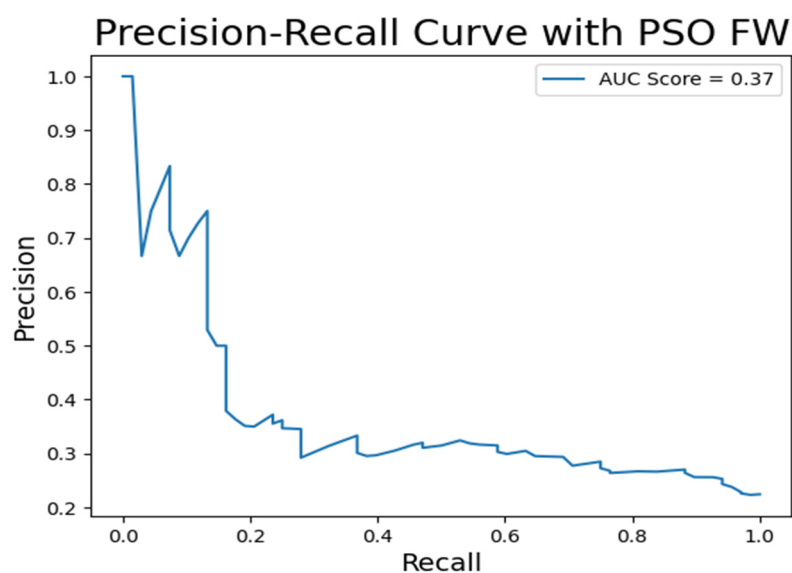


Figure 6. Precision—Recall Curve with PSO Feature weighting.

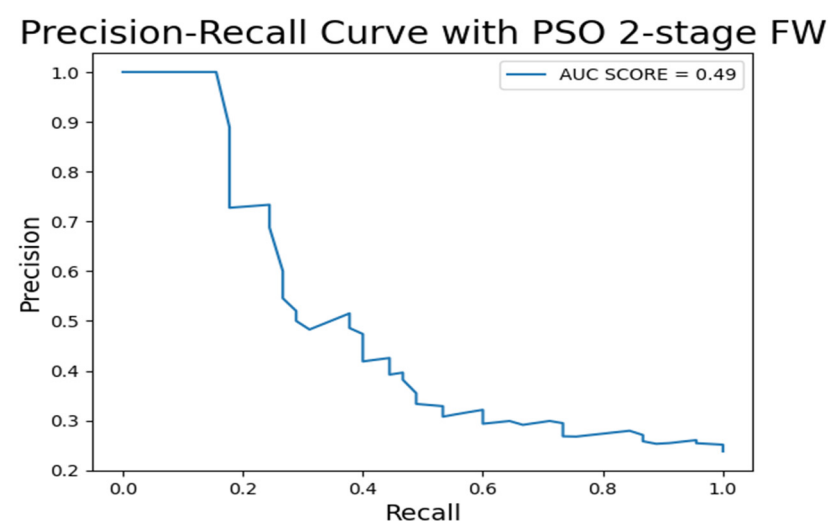


Figure 7. Precision—Recall Curve with PSO 2-stage Feature weighting.

The iterations and the respective precision and recall values of the exploration stage are shown in Figures 8 and 9. The precision curve starts at 0.825 and steadily increases and reaches a maximum of 0.99. The recall curve faces a lot of ups and downs and reaches above 0.2. The multi-objective optimization curve of the two-stage PSO is in Figure 10, and for the exploitation stage it clearly shows when precision increases, the recall decreases, and vice versa. A balance between these two values is achieved near iteration number 30 with a precision value of 1.0 and recall of 0.21. The multi-objective optimization curve of the standard PSO algorithm in Figure 11 clearly depicts the struggle it faces in achieving the balance between precision and recall. In all the iterations below 40, when precision increases recall decreases and vice versa.

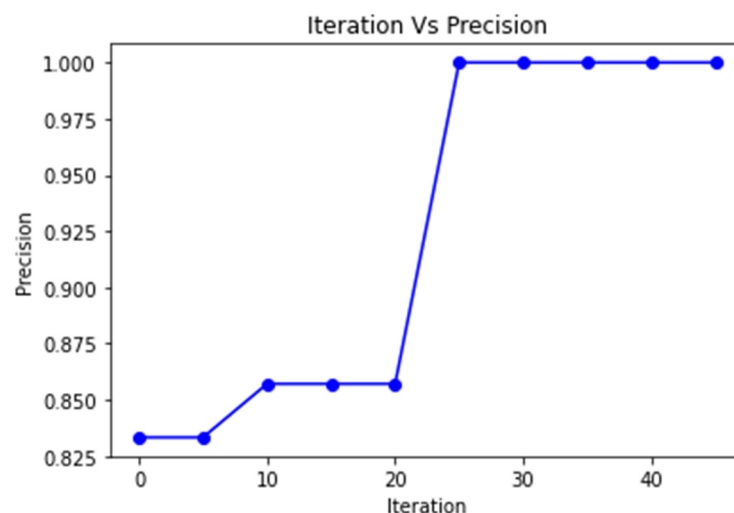


Figure 8. Iteration Vs Precision at the Exploration stage.

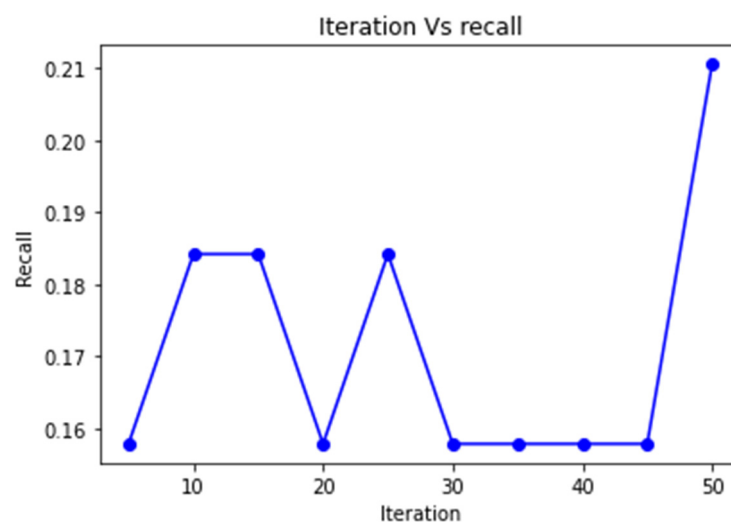


Figure 9. Iteration Vs Recall at Exploration stage.

To depict how the population has explored the solution space, the level of diversification of the particle population space was measured using the population space diversity (PSD) metric. The PSD metric was based on the Euclidean measure of the distance between the particle position ( $P_{id}$ ) and the population's average position ( $P_d^{avg}$ ). The population's average position ( $P_d^{avg}$ ) is the average of the particle position values and is calculated using Equation (14). The Euclidean distance is calculated using Equation (15) and finally, the PSD is found using Equation (16).

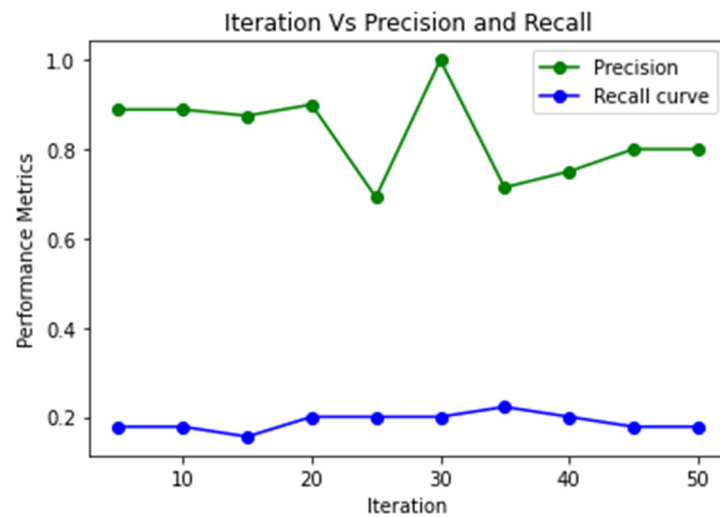
$$P_d^{avg} = \frac{1}{S} \sum_{i=1}^S P_{id} \quad (14)$$

$$Euclipsd = \sqrt{\sum_{d=1}^D (P_{id} - P_d^{avg})^2} \quad (15)$$

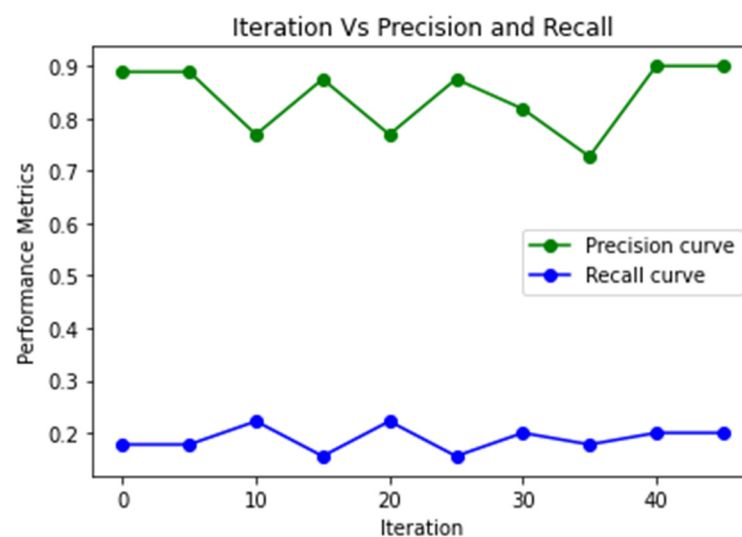
$$PSD = \frac{P_{id} - Popmin}{Popmax - Popmin} \quad (16)$$

*popmin*—Minimum Position value in Population

*popmax*—Maximum Position value in Population



**Figure 10.** Iteration Vs Precision and Recall (Multi-objective) at the Exploitation stage.



**Figure 11.** Iteration Vs Precision and Recall (Multi-objective) by standard PSO algorithm.

Figure 12 shows the population diversity achieved over the iterations in stage one, so there are many steep rises and falls visible. However, in Figure 13 the population diversity of exploitation shows small ups and downs.

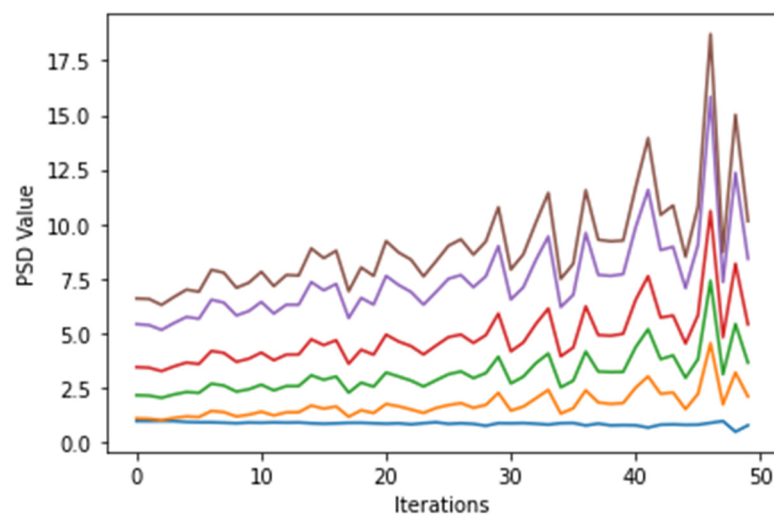


Figure 12. Iterations vs. PSD at Exploration stage.

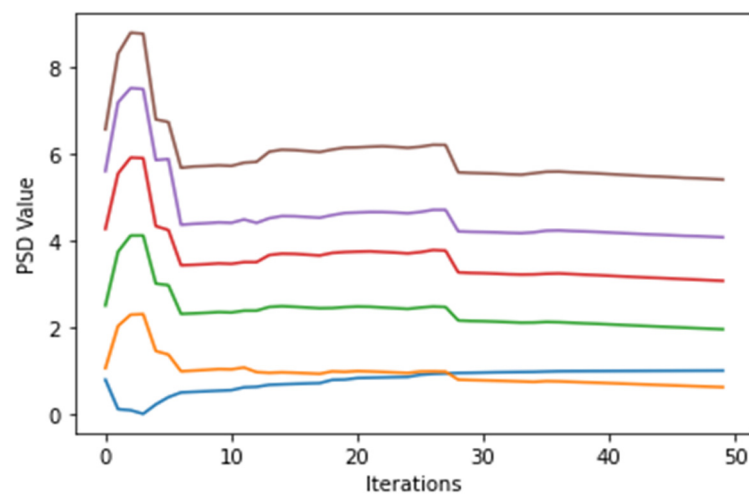


Figure 13. Iterations vs. PSD at the Exploitation stage.

The feature weights are given in Table 5 and its graphical representation in Figure 14. The information gain measure gives more weightage to SPI 12 and SPEI 12 than the shorter timescales. The Pearson correlation gives more importance to SPI 6 and SPEI 6, a medium timescale. The PSO gives weightage to SPI 3, SPEI 12, SPEI 3, and SPEI 6. Two-stage PSO gives more importance to SPI3, SPI6, and SPEI12. The experimental results show that the SPI 3, SPI 6, and SPEI 12 have higher weights by a majority of methods, which means more importance need to be given to them.

Table 5. Feature weights obtained from various methods.

S. No.	Feature Weighting Methods	SPI 3	SPI 6	SPI 12	SPEI 3	SPEI 6	SPEI 12
1.	Information Gain FW	0.468	0.466	0.488	0.463	0.472	0.489
2.	Pearson Correlation Coefficient FW	0.132	0.16	0.113	0.132	0.157	0.099
3.	PSO FW	0.77	0.31	0.53	0.56	0.58	0.7
4.	PSO 2-stage FW	0.97	0.98	0.68	0.876	0.77	0.96

## Feature Weights Comparison

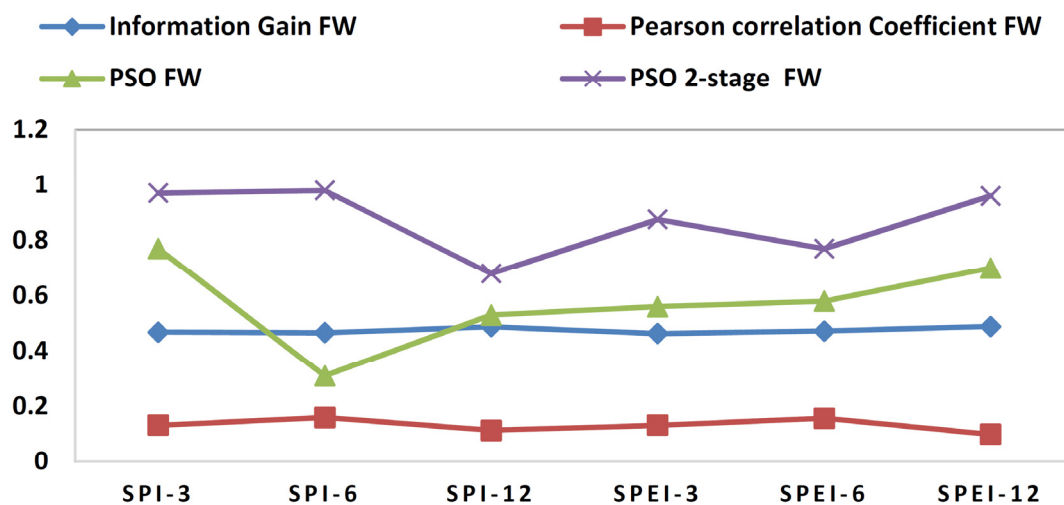


Figure 14. Comparison of feature weights given by various methods.

## 5. Conclusions

Water is a crucial resource required for the sustainable growth of any nation. Drought impacts our economy, social harmony, environment, and health, so its early detection is very much necessary to meet sustainable development. The purpose of this study is to determine the significance and efficiency of the SPI and SPEI drought indices at 3, 6, and 12 time scales used in metrological drought forecasting. The challenging issue is that the dataset used to predict drought occurrence is imbalanced, and most research focuses on forecasting drought index values rather than actual drought incidence. The feature weighting strategy significantly enhanced classification performance. In evaluating an imbalanced dataset, the metrics precision, recall, f1score, and Matthew correlation coefficient are used instead of the performance metric accuracy alone. The precision and recall are contradicting goals hence there is a need for multi-objective optimization, so the two-stage PSO algorithm was designed, and feature weights were found. The two-stage PSO feature weighting approach gave the precision and recall value of 1.0 and 0.21 whereas without feature weighting it was only 0.55 and 0.1. On average there was nearly a 45% improvement in precision and around 12% improvement in the recall was achieved with the two-stage PSO algorithm. Most weighted drought indices by various weighting methods are SPI 3, SPI 6, SPEI 6, and SPEI 12. The experimental results show that the use of weighted drought indices improves meteorological drought prediction. In the future, the local feature-weighting approach can be applied on the imbalanced dataset and their performance is to be studied. The neighbourhood learning strategies of the PSO algorithm can be improved in order to achieve a good balance between the precision and recall values.

**Author Contributions:** Conceptualization, K.S. (Kathiravan Srinivasan); Methodology, K.S. (Kathiravan Srinivasan) and K.S. (Karpagam Sundararajan); Software, K.S. (Karpagam Sundararajan); validation, K.S. (Kathiravan Srinivasan); Writing—Original draft preparation, review, and editing, visualization, K.S. (Kathiravan Srinivasan) and K.S. (Karpagam Sundararajan). All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** All data used in the research is mentioned in the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Mishra, A.K.; Singh, V.P. A review of drought concepts. *J. Hydrol.* **2010**, *391*, 202–216.
- Amin, Z.; Rehan, S.; Bahman, N.; Faisal, K. A review of drought indices. *Environ. Rev.* **2011**, *19*, 333–349.
- Edwards, D.C. *Characteristics of 20th Century Drought in the United States at Multiple Time Scales*; Air Force Inst of Tech Wright-Patterson: Hobson Way, OH, USA, 1997; No. AFIT-97-051.
- Beguería, S.; Vicente-Serrano, S.M.; Reig, F.; Latorre, B. Standardized precipitation evapotranspiration index (SPEI) revisited: Parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *Int. J. Climatol.* **2014**, *34*, 3001–3023.
- Bouaguel, W.; Mufti, G.B.; Limam, M. A fusion approach based on wrapper and filter feature selection methods using majority vote and feature weighting. In Proceedings of the International Conference on Computer Applications Technology (ICCAT), Sousse, Tunisia, 20–22 January 2013.
- Niño-Adán, I.; Manjarres, D.; Landa-Torres, I.; Portillo, E. Feature Weighting Methods: A Review. *Expert Syst. Appl.* **2021**, *184*, 115424.
- Jankowski, N.; Usowicz, K. Analysis of feature weighting methods based on feature ranking methods for classification. In Proceedings of the International Conference on Neural Information Processing, Shanghai, China, 13–17 November 2011.
- Wu, X.; Xu, X.; Liu, J.; Wang, H.; Hu, B.; Nie, F. Supervised feature selection with orthogonal regression and feature weighting. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *32*, 1831–1838.
- Agrawal, P.; Abutarboush, H.F.; Ganesh, T.; Mohamed, A.W. Metaheuristic algorithms on feature selection: A survey of one decade of research (2009–2019). *IEEE Access* **2021**, *9*, 26766–26791.
- Valdez, F.; Castillo, O.; Melin, P. Bio-Inspired Algorithms and Its Applications for Optimization in Fuzzy Clustering. *Algorithms* **2021**, *14*, 122.
- Kennedy, J.; Eberhart, R. Particle Swarm Optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995.
- Shami, T.M.; El-Saleh, A.A.; Alswaiti, M.; Al-Tashi, Q.; Summakieh, M.A.; Mirjalili, S. Particle swarm optimization: A comprehensive survey. *IEEE Access* **2022**, *10*, 10031–10061.
- Eberhart, R.C.; Shi, Y. Particle Swarm Optimization: Developments, Applications and Resources. In Proceedings of the IEEE Congress on Evolutionary Computation, Seoul, Republic of Korea, 27–30 May 2001; Volume 1, pp. 27–30.
- Jiang, L.; Zhang, L.; Li, C.; Wu, J. A Correlation-based Feature Weighting Filter for Naive Bayes. *IEEE Trans. Knowl. Data Eng.* **2018**, *31*, 201–213.
- Khadr, M. Forecasting of meteorological drought using Hidden Markov Model (case study: The upper Blue Nile river basin, Ethiopia). *Ain. Shams. Eng. J.* **2016**, *7*, 47–56.
- 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.
- Zhang, R.; Chen, Z.Y.; Xu, L.J.; Ou, C.Q. Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China. *Sci. Total Environ.* **2019**, *665*, 338–346.
- Mathivha, F.; Sigauke, C.; Chikoore, H.; Odiyo, J. Short-Term and Medium-Term Drought Forecasting Using Generalized Additive Models. *Sustainability* **2020**, *12*, 4006.
- Pande, C.B.; Al-Ansari, N.; Kushwaha, N.L.; Srivastava, A.; Noor, R.; Kumar, M.; Moharir, K.N.; Elbeltagi, A. Forecasting of SPI and Meteorological Drought Based on the Artificial Neural Network and M5P Model Tree. *Land* **2022**, *11*, 2040. [[CrossRef](#)]
- Ndayiragije, J.M.; Li, F.; Nkunuzimana, A. Assessment of Two Drought Indices to Quantify and Characterize Drought Incidents: A Case Study of the Northern Part of Burundi. *Atmosphere* **2022**, *13*, 1882. [[CrossRef](#)]
- Yang, W.; Zhou, X.; Luo, Y. Simultaneously Optimizing Inertia Weight and Acceleration Coefficients via Introducing New Functions into PSO Algorithm. *J. Phys. Conf. Ser.* **2021**, *1754*, 012195.
- Zhang, L.; Tang, Y.; Hua, C.; Guan, X. A new particle swarm optimization algorithm with adaptive inertia weight based on Bayesian techniques. *Appl. Soft Comput.* **2015**, *28*, 138–149.
- Liang, H.; Kang, F. Adaptive mutation particle swarm algorithm with dynamic nonlinear changed inertia weight. *Optik* **2016**, *127*, 8036–8042.
- Kadirkamanathan, V.; Selvarajah, K.; Fleming, P.J. Stability analysis of the particle dynamics in particle swarm optimizer. *IEEE Trans. Evol. Comput.* **2006**, *10*, 245–255.
- Pluhacek, M.; Senkerik, R.; Davendra, D.; Oplatkova, Z.K.; Zelinka, I. On the behavior and performance of chaos driven PSO algorithm with inertia weight. *Comput. Math. Appl.* **2013**, *66*, 122–134.
- Li, W.; Meng, X.; Huang, Y.; Fu, Z.H. Multipopulation cooperative particle swarm optimization with a mixed mutation strategy. *Inform. Sci.* **2020**, *529*, 179–196.
- Ye, W.; Feng, W.; Fan, S. A novel multi-swarm particle swarm optimization with dynamic learning strategy. *Appl. Soft Comput. J.* **2017**, *61*, 832–843.
- Tripathy, B.K.; Reddy Maddikunta, P.K.; Pham, Q.V.; Gadekallu, T.R.; Dev, K.; Pandya, S.; ElHalawany, B.M. Harris Hawk Optimization: A Survey on Variants and Applications. *Comput. Intell. Neurosci.* **2022**, *2022*, 2218594.
- Reddy, G.T.; Reddy, M.P.K.; Lakshmana, K.; Kaluri, R.; Rajput, D.S.; Srivastava, G.; Baker, T. Analysis of Dimensionality Reduction Techniques on Big Data. *IEEE Access* **2020**, *8*, 54776–54788.

30. Available online: <https://rda.ucar.edu/datasets/ds298.0/index.html#!cgi-bin/datasets/getWebList?dsnum=298.0> (accessed on 23 September 2022).
31. Available online: [https://spei.csic.es/spei\\_database/](https://spei.csic.es/spei_database/) (accessed on 30 September 2022).
32. Available online: <https://digitalcommons.unl.edu/droughtnetnews/57/> (accessed on 15 September 2022).
33. Elhariri, E.; El-Bendary, N.; Taie, S.A. Using Hybrid Filter-Wrapper Feature Selection With Multi-Objective Improved-Salp Optimization for Crack Severity Recognition. *IEEE Access* **2020**, *8*, 84290–84315.
34. Lei, S. A feature selection method based on information gain and genetic algorithm. In Proceedings of the 2012 International Conference on Computer Science and Electronics Engineering, Hangzhou, China, 23–25 March 2012; Volume 2, pp. 355–358.
35. Yang, M.; Liu, Y.; Yang, J. A hybrid multi-objective particle swarm optimization with central control strategy. *Comput. Intell. Neurosci.* **2022**, *2022*, 1522096.
36. Ekbalm, A.; Saha, S.; Garbe, C. Feature selection using multiobjective optimization for named entity recognition. In Proceedings of the 20th International Conference on Pattern Recognition (ICPR), Istanbul, Turkey, 23–26 August 2010; pp. 1937–1940.
37. Wei, B.; Xia, X.; Yu, F.; Zhang, Y.; Xu, X.; Wu, H.; Gui, L.; He, G. Multiple adaptive strategies based particle swarm optimization algorithm. *Swarm Evolut. Comput.* **2020**, *57*, 100731.
38. Lim, W.H.; Isa, N.A.M. An adaptive two-layer particle swarm optimization with elitist learning strategy. *Inf. Sci.* **2014**, *273*, 49–72.
39. Taherkhani, M.; Safabakhsh, R. A novel stability-based adaptive inertia weight for particle swarm optimization. *Appl. Soft Comput.* **2016**, *38*, 281–295.
40. Liu, H.; Zhang, X.-W.; Tu, L.-P. A Modified Particle Swarm Optimization Using Adaptive Strategy. *Expert Syst. Appl.* **2020**, *152*, 113353.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.