water

MDPI

Article

Least Squares Support Vector Machine for Ranking Solutions of Multi-Objective Water Resources Allocation Optimization Models

Weilin Liu [1,*], Lina Liu [1] and Fang Tong [2]

[1] Nanchang Institute of Technology, Jiangxi Engineering Research Center of Water Engineering Safety and Resources Efficient Utilization, Nanchang 330099, Jiangxi, China; happylina617@163.com
[2] School of Civil Engineering, Hefei University of Technology, Hefei 230009, Anhui, China; tongfang824@163.com
* Correspondence: weilin_liu@nit.edu.cn; Tel.: +86-791-8215-3183

Academic Editor: Ashok K. Chapagain
Received: 23 January 2017; Accepted: 31 March 2017; Published: 5 April 2017

Abstract: There is an increasing trend in the use of multi-objective evolutionary algorithms (MOEAs) to solve multi-objective optimization problems of the allocation of water resources. However, typically the outcome is a set of Pareto optimal solutions which make up a trade-off surface between the objective functions. For decision makers to choose a satisfactory alternative from a set of Pareto-optimal solutions, this paper suggests a new method based on least squares support vector machine (LSSVM) and *k*-means clustering for ranking the optimal solutions for the multi-objective allocation of water resources. First, the *k*-means clustering method was adopted to reduce the large set of solutions to a few representative solutions. Then, to capture and represent the decision maker's preferences as well as to select the most desirable alternative, the LSSVM method was applied to obtain the utility value for each representative solution. According to the magnitude of the utility values, the final priority orders of the representative solutions were determined. Finally, this methodology was applied to rank the Pareto optimal solution set obtained from the multi-objective optimization problems of water resources allocation for the water-receiving areas of the South-to-North Water Transfer Project in Hebei Province, China. Moreover, the comparisons of the proposed method with the information entropy method and the artificial neural network (ANN) model were given. The results of the comparison indicate that the proposed method has the ability to rank the non-dominated solutions of the multi-objective operation optimization model and that it can be employed for decision-making on water allocation and management in a river basin.

Keywords: least squares support vector machine (LSSVM); *k*-means clustering; artificial neural network (ANN); water resources allocation

1. Introduction

As a limited natural resource, water is increasingly demanded for various purposes, and how to allocate water from a river basin is among the most widely discussed issues in water resources management [1]. In principle, the allocation of water resources is a very complex management issue involving social, economic, environmental, and political factors. Therefore, attempting to optimize the allocation of water resources also requires solving multi-objective problems that require powerful optimization tools to fully characterize the existing trade-offs [2,3].

In recent decades, there have been different optimization methods introduced into the literature for optimal allocation of water resources. These methods have been applied with various degrees of success, based on mathematical programming such as linear and dynamic programming [4–6], and

more recently on evolutionary algorithms, such as strength Pareto EA (SPEA), non-dominated sorting genetic algorithm (NSGA-II) and multi-objective particle swarm optimization (MOPSO) [7,8]. However, considering the characteristics of water resources allocation, such as nonlinearity, discreteness, non-convexity and high dimension decisions, multi-objective evolutionary algorithms (MOEAs) are proven to be more suitable for discovering and exploiting the critical tradeoffs of multi-objective water resource allocation problems, due to their efficiency and ease in handling non-linear and non-convex relationships of real-world problems, compared to the capacity of traditional optimization techniques [1–3,9].

Broadly, multi-objective optimization problems consist of two phases: the search for Pareto optimal solutions, and the decision-making process to identify the preferred ones. The former is a problem known as multi-objective optimization, the goal of which is usually to find the set of Pareto optimal solutions. The latter phase is traditionally defined as the evaluation of non-dominated solution sets, which focuses on problems of ranking alternatives from a set of options using multiple evaluation attributes. However, most of the research on multi-objective water resource allocation has been focused on the development of multi-objective optimization methods that generate a set of Pareto optimal solutions [10–12]; few studies were performed on the choice of alternatives from a set of Pareto optimal solutions. Reddy and Kumar [13] employed a pseudo-weight vector approach to provide the relative weight of each alternative generated by EM-MOPSO. Through analyzing the weight combinations, depending on the preference of the reservoir operator, a suitable policy was implemented. Liu [14] applied MOPSO to multi-objective water resource allocation optimization and presented an information entropy method to rank the decision results according to the magnitude of the superiority degrees.

In application, the decision maker might prefer the least possible number of representative solutions for further analysis. Hence, it is necessary to evaluate different trade-off alternatives for water resource managers, after obtaining the set of broad Pareto-optimal solutions. Furthermore, the evaluation and ranking of alternatives is grounded on their attribute values, and on the preferences of the various decision makers. The key to similar decision-making issues is to obtain the preference structure information of the decision makers. One common approach is to use the multi-attribute utility functions to measure the satisfaction level of the decision-makers to each of the alternatives, and then to sort the solution sets by utility values, ultimately selecting the most satisfying solution. Multiple attribute utility functions are usually broken down into the additive and multiplicative form of a single attribute utility function. The requirements for these utility function forms are very strict, for example, the independence between the attributes and linear utility function for each attribute [15]. However, due to the complexity of these decision-making problems in the real world, these strict conditions are difficult to meet. Therefore, further research on alternatives selection for multi-objective water resource allocation is required.

It is evident that selecting the satisfactory solution from a number of potential alternatives in water resource allocation involves a complex decision-making process with conflicting quantitative and qualitative criteria and multiple decision-makers. Therefore, it is difficult to derive an accurate formula to represent all the decision-making processes involved. In recent years, computing techniques such as support vector machine (SVM) and artificial neural network (ANN), have been widely used to model nonlinear processes. They can capture the behavior of the underlying processes (physical or otherwise) [16]. ANNs, on the basis of the learning function of a human brain, have shown strong ability of nonlinear mapping and have been gradually applied in various fields, including water resources. However, there are some disadvantages of ANNs, such as over-fitting, slow convergence, and becoming easily trapped in local minimum, which make it difficult to perform satisfactorily when dealing with complex nonlinear processes [16,17]. The SVM proposed by Vapnik is an effective learning method and is considered as an alternative method to ANNs [18–22]. Based on statistical learning theory, the SVM has a high ability for generalization and is less prone to over-fitting [23]. SVMs seem to overcome some of the basic weaknesses associated with ANNs modeling while retaining all of ANNs strengths [16]. SVMs provide a better option for the modeling of nonlinear processes and

the decision-making process can benefit from the use of SVM techniques. The least squares support vector machine (LSSVM), originally proposed by Suykens et al. [24] is modified from SVM. With better performance than SVM, LSSVM has been successfully applied in diverse fields [25,26]. In the last decade, LSSVM has been extended to a wide range of water resources problems [27–29]. However, it has not been used for research on choosing alternatives from a set of Pareto optimal solutions for the multi-objective water resource allocation.

The main objective of this study is to develop a new method for selecting alternatives for multi-objective water resource allocation, and to evaluate and rank alternatives from a set of Pareto optimal solutions. In this paper, aiming at the Pareto optimal set generated by MOEAs for a multi-objective water resources allocation problem, a new approach for ranking the optimal solutions was proposed based on LSSVM and *k*-means clustering. Then, a case study of the water-receiving areas of the South-to-North Water Transfer Project in Hebei Province of China was conducted. Results demonstrate that the proposed approach is able to offer plenty of alternative policies for water resource managers.

The rest of this paper is organized as follows: Section 2 provides a description of the study region and elaborates on the methodology proposed for ranking solutions of multi-objective water resource allocation optimization models. An illustrative case study is used to present the implementation of the methodology and its corresponding results and discussions of the application are presented in Section 3. Finally, conclusions are given in Section 4.

2. Study Area and Methodology

2.1. Study Area

The water-receiving area of the South-to-North Water Transfer Project in Hebei province of China is located in the south-central area of Hebei province, Northern China, between 114°23′ and 116°42′ longitude east and 36°18′ and 39°30′ latitude north, respectively. The region covers a total area of 62,100 km^2, and includes seven cities: Hengshui, Cangzhou, Langfang, Baoding, Shijiazhuang, Xingtai and Handan in Hebei province, China. As a semi-arid region, the annual precipitation is about 550 mm, 70%–80% of which is concentrated in the rainy seasons (July–September), and the average annual water resources is 5.48 billion m^3. As a result, water scarcity is common in this region. Water resources per capita of the region is less than 600 m^3, which is less than one-third of China's per capita amount.

The middle route of the South-to-North Water Transfer Project is a water diversion project built to alleviate the water shortage in northern areas, ranging from the southwest to the northeast of this region. With this water transfer project, the network system of water supply of this region has been formed, consisting of the two water source project, the six water diversion project and the nine large storage projects. Therefore, the South-to-North Water Transfer Project and its network system of water supply for the region can contribute to alleviating the shortage of water resources and the deterioration of the water environment and to promoting economic and social sustainable development. Thus, how to allocate water becomes an important issue in a water-receiving area. To achieve rational allocation of water resources in this region, a water allocation model with three objectives named square sum of relative water shortage, water supply cost, and the amount of organic pollutants in water, was built, and a set of non-dominated solutions was obtained [2]. However, in any application, the decision-maker might prefer the smallest possible number of well representative solutions for further analysis. Hence, it is necessary to evaluate the non-dominated solutions and make choices between different trade-offs for the water resource managers after obtaining a set of non-dominated solutions that are well spread and widely covered.

2.2. Methodology

In this study, a novel approach based on support vector machine and *k*-means clustering was proposed when choosing from a set of Pareto optimal solutions for multi-objective water resource allocation. Using the Pareto optimal set for the multi-objective water resource allocation problem

generated by MOEAs such as MOPSO and NSGA, the *k*-means clustering method was adopted to reduce the large set of solutions to a small number of representative solutions. Then, to capture and represent the decision maker's preferences and to select the most desirable alternative, the LSSVM method was employed to obtain the utility value for each representative solution.

2.2.1. *k*-Means Clustering Method

The *k*-means clustering method, proposed by MacQueen, is one of the most popular and effective clustering methods used to minimize the sum of squared distances between all points and the cluster centres [30]. The procedure follows a simple method of classifying a given data set into a certain number of clusters (assume *k* clusters) fixed a priori. The main procedure consists of the following steps.

Step 1: Normalize all the sample data. Because the *k*-means algorithm requires calculating the distance between the samples to the cluster center, some attributes with a large magnitude order may affect matching results, which makes the clustering inaccurate or even wrong [31]. In order to eliminate the influence of dimension and orders of magnitude, the data should be normalized before the implementation of the *k*-means algorithm.

Step 2: Choose *k* initial cluster centres $\{m_1, m_2, \cdots, m_k\}$. Considering different centres causes a different result, so the better choice is to place them as far away from each other as possible.

Step 3: Calculate the Euclidean distance from each point in the dataset to each cluster center. All points are assigned to their nearest cluster centers according to the Euclidean distance formula expressed as follows:

$$d(x_i, m_j) = \sqrt{\sum_{l=1}^{q} (x_{il} - m_{jl})^2},\ i = 1, 2, \cdots, N; j = 1, 2, \cdots k \tag{1}$$

where, $d(x_i, m_j)$ is the distance from the *i*-th vector data to the *j*-th cluster center; x_i is the *i*-th data and m_j is the mean for the points over cluster *j*.

Step 4: Recalculate the new *k* cluster centres as barycenters of the clusters resulting from the previous step when each point in the data set is assigned to a cluster. The new cluster centre is given by

$$m_{jl} = \frac{1}{n} \sum_{x_j \in C_j} x_{jl},\ m_j = \{m_{j1}, m_{j2}, \cdots, m_{jq}\} \tag{2}$$

where C_j is a data set in the *j*-th cluster.

Step 5: Repeat steps 3 and 4 until the center point is not changed. Once these *k* cluster centres change their location, the above steps should be carried on until no more changes arise, which indicates the convergence criterion for clustering. Through selecting a non-dominated solution from each of the final *k* clusters, a representative non-dominated set will be formed.

2.2.2. LSSVM Method

The LSSVM, originally proposed by Suykens et al., is modified from SVM. Compared with the standard SVM, the LSSVM adopts a least squares linear system as a loss function and requires only the solution to a convex linear problem, and not a quadratic problem (QP), as in the SVM. Thus, the computational process will be simplified immensely and computational costs might be reduced when large-scale QP problems are needed to be computed [32].

The regression estimation with the LSSVM is to estimate a function according to a given data set $\{r_i, u_i\}_{i=1}^{n}$ with input data $r_i \in R^m$ and output $u_i \in R$, where *n* is the number of samples. In feature space, the LSSVM regression function is given by Equation (3), where φ (.) is the high dimensional feature space that is non-linearly mapped from the input space, ω is the adjustable weight vector, and *b* is the scalar threshold. Then, for function estimation, the following optimization problem can be described by Equation (4), where ξ_i is the random error, *c* is the regularization parameter. Finally, by

introducing Lagrange multipliers a_i, the regression estimation of the LSSVM can be formulated by Equation (5), where $K(r_i, r_j)$ is the kernel function to meet Mercer's conditions. Such frequently used kernel functions are linear, polynomial, sigmoid and radius basis function (RBF). Among them, the RBF is an effective option applied most frequently with a stronger nonlinear approximation ability and fewer parameters than others. It can be expressed as Equation (6), where σ is the width of the RBF. A more detailed review of LSSVM algorithm for nonlinear function estimation can be found in [28]. In this study, the LSSVM-base regression was performed using a LSSVM toolbox for MATLAB [33].

$$u = f(r) = \omega^T \varphi(r) + b \tag{3}$$

$$\begin{cases} \min \frac{1}{2}\|\omega\|^2 + \frac{1}{2}c\sum_{i=1}^{n}\xi_i^2 \\ s.t. \quad u_i = \omega^T \varphi(r_i) + b + \xi_i \\ i = 1, 2, \cdots, n \end{cases} \tag{4}$$

$$f(r) = \sum_{i=1}^{n} a_i K(r, r_i) + b \tag{5}$$

$$K(r_i, r_j) = \exp(\|r_i - r_j\|^2 / 2\sigma^2) \tag{6}$$

Making a satisfactory choice from a representative set of non-dominated solutions in water resource allocation is a complex decision-making process that involves evaluating and ranking alternatives. The key to solving this problem is obtaining the preference structure information of decision makers. One common approach is to use the multi-attribute utility function to measure the degree of satisfaction of decision-makers to each alternatives, and then to sort the solution set according to the value of the utility function in order to select the most satisfactory solution, which can be expressed as follows:

$$u_i = F(r_i) \tag{7}$$

In general, multiple attribute utility functions are usually broken down into the additive and multiplicative form of a single attribute utility function. However, the requirements for these utility function forms are quite strict, such as the independence between the attributes and linear utility function for each attribute. Due to the complexity of the actual decision-making problems, such conditions are difficult to meet. In essence, the above utility function can be regarded as a mathematical mapping problem, in which the input is the objective function value vector $r_i = \{r_{i1}, r_{i2}, \ldots r_{im}\}$, and the output is the corresponding utility value u_i made by the decision makers. Therefore, this study adopts the above LSSVM algorithm for nonlinear function estimation to estimate such mapping F of $R^m \rightarrow R$ (Figure 1) between the objective function value vector r_i and the utility value u_i, which reflects the preference structure of decision makers.

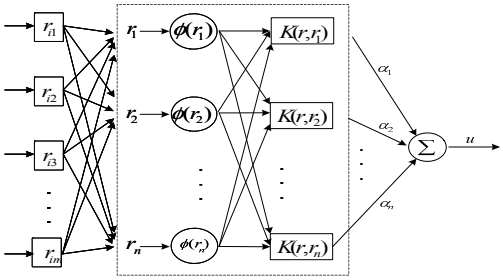


Figure 1. Schematic diagram of support vector machine for multi attribute decision-making.

2.2.3. Implementation of LSSVM Model

In order to apply the LSSVM to estimate the above mapping F, there are three main problems to be solved, including the selection of kernel function and its parameters, and the construction of

the training sample dataset. For selecting kernel function, no systematic methodology is available; it depends on the distribution of the data and generally can be selected through a trial and error approach [22]. Due to the good generalizing properties of the RBF, it is employed in most of the applications [16–28], so the RBF was chosen as kernel function for the function estimation of the mapping F in this study. As for the selection of parameters, there are only two parameters, c and σ, to tune into LSSVM model with RBF kernel, which have great impact on the performance and generalization ability of the LSSVM. In previous studies, the parameters are usually determined by a trial-and-error process which is less efficient and more difficult to reach a better set of parameters. To find optimal parameters, many optimization methods have been applied, such as particle swarm optimization (PSO), simple genetic algorithm, a grid search technique and k-folds cross-validation. Due to the simple concept, easy implementation and quick convergence, we used PSO combined with 10-fold cross validation procedure, by which these parameters were automatically tuned into the training phase. A more detailed review of the kernel function and the parameters selection for LSSVM can be found in [34]. Since there are plenty of studies on kernel function and parameters selection, construction of the training samples dataset is the main focus of this research.

Suppose m is the number of objectives, k is the number of representative Pareto solutions, and there is a set of training samples S = $\{r_i, u_i\}_{i=1}^{n}$ where r_i is the i-th input vector and u_i is its corresponding observed result. Here, r_i denotes the objective function value, and u_i is the assessment result for multi-objective problems. It is evident that a set of training samples for LSSVM is the foundation of conducting an overall assessment of the Pareto solutions, but due to the complexity of the multi-attribute utility function structure, it is usually difficult to obtain utility values. In this study, an approach to construct learning samples for the MCDM problem based on estimating the utility functions is adopted. In general, two possible choices for utility functions are suggested, namely linear and sigmoid functions. In this study, the sigmoid function was used which can be expressed as follows:

$$u = 1/(1 + e^{-a(r-b)}) \tag{8}$$

where a and b are constants.

In order to make the learning samples reflect the overall characteristics of the decision-making objects, three types of data samples: ideal point samples, random samples and scheme samples, are constructed as follows:

(1) Ideal point samples

Let I_1 and I_2 respectively represent the index set of benefit-typed indices and cost-typed indices. Then, r_j^* is the ideal point for the jth indice r_j if $r_j^* = \max\limits_i r_{ij}, j \in I_1, r_j^* = \min\limits_i r_{ij}, j \in I_2$, while r_j^- is the negative ideal point for the j-th indice r_j if $r_j^* = \min\limits_i r_{ij}, j \in I_1, r_j^* = \max\limits_i r_{ij}, j \in I_2$. Thus, ideal point samples can be formed by the attribute values with the same utility value. For example, the ideal solution r^* = $\{r_1^*, r_2^*, \ldots\ r_m^*\}$ and its corresponding utility value u^* can compose the ideal point sample, denoted as $\{r_1^*, r_2^*, \ldots\ r_m^*, u^*\}$, while the negative ideal solution r^- = $\{r_1^-, r_2^-, \ldots\ r_m^-\}$ and its corresponding utility value u^- can form the negative ideal sample denoted as $\{r_1^-, r_2^-, \ldots\ r_m^-, u^-\}$.

(2) Random samples

A random sample can be formed by the attribute values $\widehat{r}_{dj}$ $(d = 1, 2, \cdots, p, j = 1, 2, \cdots, m)$ and its corresponding utility value $\widehat{u}_d$ $(d = 1, 2, \cdots, p)$, where the utility values $\widehat{u}_d$ $(d = 1, 2, \cdots, p)$ are valued according to a certain step size (h) in the interval 0.05–0.95, and the corresponding attribute values $\widehat{r}_{dj}$ $(d = 1, 2, \cdots, p,\ j = 1, 2, \cdots, m)$ can be calculated at different utility values $\widehat{u}_d$ by the inverse of the utility function. Here, p is the size of random samples depending on the step size (h), which can be obtained by calculating p = (0.95—0.05)/h − 1. Then, the attribute values with the same utility value and its utility value can compose the random sample, denoted as $\{\hat{r}_{d1}, \hat{r}_{d2}, \cdots, \hat{r}_{dm}, \hat{u}_d\}$.

(3) Scheme samples

In scheme samples, the utility values $\widehat{u}_s$ ($s = 1, 2, \cdots, q$) are derived from the real attribute data by utility function. For different attribute values $\widehat{r}_{sj}$ ($s = 1, 2, \cdots, q$, $j = 1, 2, \cdots, m$), the corresponding utility values $\widehat{u}_s$ ($s = 1, 2, \cdots, q$) can be determined by the utility functions, where q is the size of scheme samples determined according to the number of actual samples, and, in general, is computed by $q = k \cdot m$. Then, a scheme sample, denoted as $\{\hat{r}_{s1}, \hat{r}_{s2}, \cdots, \hat{r}_{sm}, \hat{u}_s\}$, can be formed by the attribute values with the same utility value.

2.2.4. Steps of the Methodology

The specific steps of the proposed method are described as follows:

Step 1: Obtaining the objective function value matrix. Suppose $X = \{x_1, x_2, \cdots, x_n\}$ is a set of non-dominated solutions obtained from MOEAs. Then the corresponding objective function value matrix $F = (f_{ij})_{n \times m}$ can be calculated for multi-objective problems with m objectives, where f_{ij} is the value of the j-th objective of the i-th non-dominated solution, and n is the number of non-dominated solutions.

Step 2: Standardizing the objective function value matrix F. In order to eliminate the effects of different dimensions between indicators, the original data of indicators in F should be dimensionless by data standardization. The normalized formula is defined as follows:

$$r_{ij} = (f_{ij} - \mu_j)/\sigma_j \tag{9}$$

where r_{ij} is the standardized value of the j-th objective of the i-th non-dominated solution; μ_j, σ_j are average value and standard deviation of the j-th objective, respectively.

Step 3: Forming representative objective function value matrix. Aiming at the large set of non-dominated solutions for a multi-objective water resource allocation problem generated by MOEAs, a representative objective function value matrix $R = (r_{ij})_{k \times m}$ can be selected from each of the k clusters by the k-means clustering method.

Step 4: Establishing learning samples dataset for LSSVM. According to the above-mentioned method of constructing sample dataset in the previous section of the paper, ideal point samples, random samples and scheme samples can be obtained, which form a learning sample dataset S = $\{\hat{r}_i, \hat{u}_i\}_{i=1}^{p}$, with the input $\hat{r}_i \in R^m$, and the output $\hat{u}_i \in R$, where R^m is the m-dimensional vector space and R is the one-dimensional vector space. Then, for the training LSSVM model, the constructed dataset is divided into two subsets, namely training dataset and testing dataset.

Step 5: Establishing the LSSVM model. For the LSSVM model, RBF is used as kernel function. Thus, there are only two parameters, c and σ to be tuned. Then, PSO combined with 10-fold cross validation procedure is adopted to choose two optimal parameters for the LSSVM model. In 10-fold cross validation, the training dataset is randomly divided into ten subsets of equal size. Nine of the subsets are used as the training set, and the remaining one is used as the validation set. Then, the cross validation process is repeated ten times with each subset used once as the validation set for each parameter setting. The mean square error (MSE) of the entire training set for each parameter setting is calculated based on the regression model, which is expressed by the following equation:

$$MSE = \frac{\sum_{i=1}^{10} \sum_{j=1}^{l} (u_{ij} - \widehat{u}_{ij})^2)/l}{10} \tag{10}$$

where MSE is the mean square error, u_{ij}, $\widehat{u}_{ij}$ are the j-th prediction and actual value of i-th validation set, respectively, l is the size of the validation subset. The optimal parameters (c, σ) can be selected as parameters of LSSVM, where MSE is minimum. By parameters optimization, the optimum values of c and σ can be determined, and the LSSVM model can be established, which can capture the input and

output relations for the training dataset. Then, the performance of the developed LSSVM model is examined based on the testing dataset.

Step 6: Calculating the utility values of the objective function values with trained LSSVM model. To carry out the assessment, after establishing the LSSVM model, using the objective function values corresponding to a representative set of non-dominated solutions as the input of the developed LSSVM model, the corresponding utility values can be obtained. According to the magnitude of the utility values, the final priority order of the i representative non-dominated solution can be determined.

3. Results and Discussion

As stated above, a water allocation model with three objectives was built in [2], involving minimization of square sum of relative water shortage, water supply cost, and the amount of organic pollutants in water. Moreover, a set of non-dominated solutions that were well spread and widely covered for this multi-objective water resource allocation was generated based on MOPSO technique, and the Pareto optimal front was shown in Figure 2. However, in any application, the decision-maker might prefer the smallest possible number of representative solutions for further analysis, and choose the most optimal solution among them. So, it is necessary to evaluate different alternatives for water resource management after obtaining a set of well-spread non-dominated solutions. In this study, based on the obtained alternatives, the proposed method is applied to carry out the decision-making analysis for the scheme choices for the water-receiving area of the South-to-North Water Transfer Project in Hebei province.

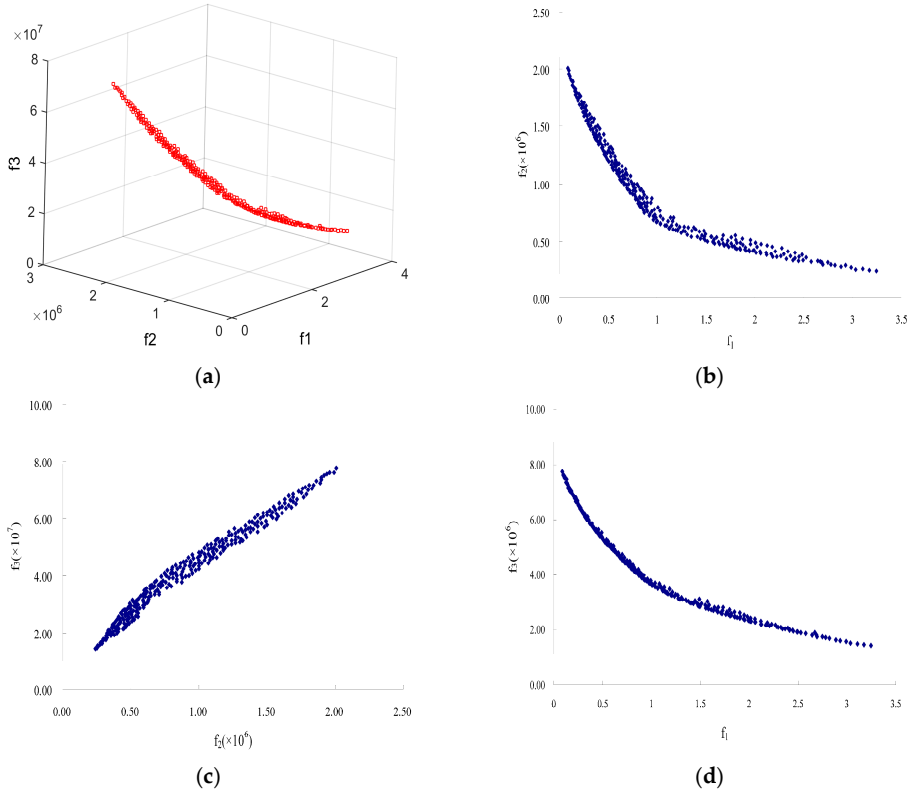


Figure 2. (**a**) The Pareto frontier obtained by multi-objective particle swarm optimization (MOPSO) for the three-objective water resources allocation model; Relation between objectives of the Pareto frontier for the water allocation model: (**b**) f_1 and f_2; (**c**) f_2 and f_3; (**d**) f_1 and f_3. (f_1—square sum of relative water shortages; f_2—water supply cost; f_3—the amount of organic pollutants (COD) in water).

Firstly, in order to eliminate the influence of dimension difference, the objective function value matrix $F = (f_{ij})_{300\times3}$ for all alternatives generated by MOPSO were standardized according to

Equation (9). Then, through *k*-means cluster, the large number of alternatives was reduced to a small number of representative alternatives, within which the number of clusters *k* was chosen as 20. Thus, according to the *k*-means cluster algorithm described above, 20 representative alternatives were obtained for the three-objective water resources allocation model. Figure 3 shows the 20 representative clustered alternatives obtained by *k*-means cluster method.

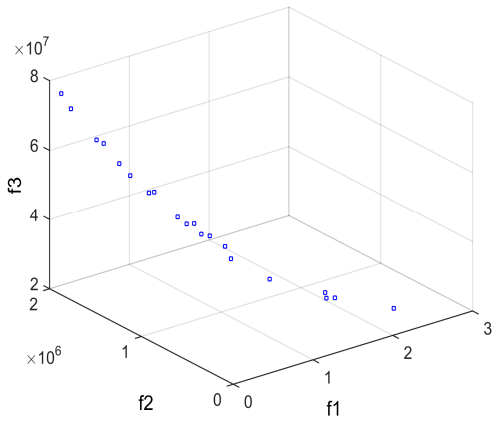


Figure 3. 20 representative alternatives obtained by *k*-means clustering. (f_1—square sum of relative water shortages; f_2—water supply cost; f_3—the amount of organic pollutants (COD) in water).

Then, LSSVM was used to evaluate the 20 representative alternatives obtained from *k*-means cluster method. The square sum of relative water shortages (f_1), water supply cost (f_2) and the amount of organic pollutants in water (f_3) were selected as the input of the model, the corresponding utility was taken as the model output. For *j*-th objective function, the maximum value was selected as its ideal point r_j^* from standardized alternatives, and its utility value u^* was set to 0.95. At the same time, the minimum value of j-th objective function was selected as its negative ideal point r_j^- and its utility value u^- was set to 0.05. According to the definition of ideal point samples, $\{r_1^*, r_2^*, r_3^*, u^*\}$ and $\{r_1^-, r_2^-, r_3^-, u^-\}$ composed two ideal point samples (Table 1).

Table 1. List of the ideal point and negative ideal point.

Sample Name	f_1	f_2	f_3	Utility Value
Ideal point	1.1491	1.5433	1.7070	0.95
Negative ideal point	−2.4907	−1.7455	−1.7995	0.05

Note: The values of f_1, f_2 and f_3 in table 1 are all normalized.

To construct random samples and scheme samples, the sigmoid function was selected as utility function expressed by Equation (8), in which a_j and b_j were determined by $\{r_j^*, u^*\}$ and $\{r_j^-, u^-\}$, respectively; where a_j and b_j denote the parameters of utility function for *j*-th objective. For instance, for f_1, a_1 and b_1 were computed by Equation (8) using the ideal point r_1^*, the negative ideal point r_1^- and their utility values, where $r_1^* = 1.1491$, $u^* = 0.95$, $r_1^- = -2.4907$ and $u^- = 0.05$ obtained from Table 1. Thus, the two parameters, a_1 and b_1, could be obtained, which were equal to 1.6179 and −0.6708, respectively. Then, based on the formation of Equation (8), the utility function for f_1, u_1, can be presented as:

$$u_1 = 1/(1 + e^{-1.6179(r+0.6708)}) \quad (11)$$

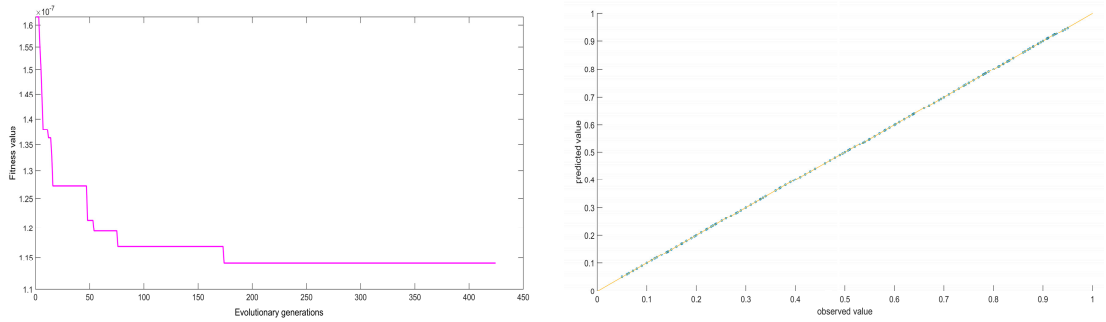
Similarly, the utility functions for f_1 and f_2, can also be given as:

$$u_2 = 1/(1 + e^{-1.7906(r+0.1011)}) \quad (12)$$

$$u_3 = 1/(1 + e^{-1.6794(r+0.0463)}) \quad (13)$$

According to the step size of 0.01, the utility values were valued in the interval [0.05, 0.95], and utility values $\widehat{u}_d$ ($d = 1, 2, \cdots, 89$) can be obtained. Then, by the inverse of the determined utility functions, the corresponding attribute values $\widehat{r}_{dj}$ ($d = 1, 2, \cdots, 89$, $j = 1, 2, 3$) were calculated at different utility values $\widehat{u}_d$. For example, for $\widehat{u}_2 = 0.06$when d takes 2, the values of the corresponding attributes can be obtained by introducing $\widehat{u}_2 = 0.06$ into the inverse of Equation (8), respectively, with $\widehat{r}_{21} = -2.3715$, $\widehat{r}_{22} = -1.6378$ and $\widehat{r}_{23} = -1.6846$. In the same way, the remaining 88 values of the corresponding attributes can be computed. Thus, 89 random samples, which consisted of the attribute values $\widehat{r}_{dj}$ ($d = 1, 2, \cdots, 89$, $j = 1, 2, 3$) and its corresponding utility value $\widehat{u}_d (d = 1, 2, \cdots, 89)$, were formed, denoted as $\{\hat{r}_{d1}, \hat{r}_{d2}, \hat{r}_{d3}, \hat{u}_d\}_{d=1}^{89}$. Similarly, 54 scheme samples, $\{\hat{r}_{s1}, \hat{r}_{s2}, \hat{r}_{s3}, \hat{u}_s\}_{i=1}^{54}$, were obtained from 20 representative alternatives (except for the ideal point samples). Moreover, for the training LSSVM model, the constructed sample dataset (144 samples × 3 variables) was divided into the training dataset and the testing dataset. Here, 10 sample data were chosen as a testing dataset and the rest were used as a training dataset.

To apply the LSSVM model, the radial basis function (RBF) was used as the kernel function of LSSVM. Further, to obtain optimum parameters c and σ, PSO was employed in which the evaluation was provided by a 10-fold cross-validation. For parameters used in PSO, the initial population size was set to 50; the cognitive parameter c1 and social parameter c2 were set to 2.0; the inertial weight w was set to 0.9; the number of iterative step was set to 1000. The termination criterion is whether it reaches the maximum number of iterations or if there is no significant improvement in the solution in 250 consecutive iterations. Figure 4a shows the details of the improvement of fitness values over the iterations. From Figure 4a it can be observed that for up to 175 iterations there was no significant improvement in the fitness values which was 1.136×10^{-7}. Based on the optimization, the values of the optimal parameters were determined as c = 1000 and σ^2 = 5.5548.



(**a**) (**b**)

Figure 4. (**a**) The curve of fitness values over evolutionary generation; (**b**) Scatter plot of training dataset for observed value and predicted value.

Then LSSVM model with the optimal parameters was applied to calculate the utility values. The performance of the LSSVM model was assessed by the root mean square error (RMSE) and the relative mean error (RME) in this study. The performance statistics of the LSSVM model is given in Table 2. From Table 2, it can be seen that RME and RMSE for the training dataset were 0.06731% and 2.262×10^{-4}, respectively, indicating that the developed LSSVM model had good performance. Furthermore, as illustrated by the scatter plot of the observed value and the predicted value for the training set shown in Figure 4b, it also shows that the developed model performed well, because the value of R^2 for training set was 1. Therefore, the developed LSSVM model can successfully capture the input and output relationship for the training dataset. Then, the performance of the developed

LSSVM model was examined for the 10 testing data. The simulation results are shown in Table 3 and the performance statistics for testing data is given in Table 2. As shown in Table 2, the LSSVM model with RMSE = 1.025×10^{-4}, and RME = 0.0416%, predicted the utility value with high accuracy for testing data, and all was less than 0.2%, indicating that the simulating effect of the model was good and that the developed LSSVM model had the capability of predicting the utility values. The results demonstrate that the developed LSSVM model performed similarly in training periods and testing periods, which all had high accuracy.

Table 2. The performance statistics of least squares support vector machine (LSSVM) model and radial basis function (RBF) network model during training and testing periods.

Model	Training		Testing	
	RMSE	RME	RMSE	RME
LSSVM	2.262×10^{-4}	0.06731%	1.025×10^{-4}	0.0416%
RBF network	3.211×10^{-4}	0.1448%	3.159×10^{-4}	0.1053%

Table 3. Predicting results for testing dataset based on the developed LSSVM model.

Testing Sample NO.	1	2	3	4	5	6	7	8	9	10
Expected value	0.07	0.15	0.25	0.35	0.45	0.55	0.65	0.75	0.85	0.93
Simulated value	0.0699	0.1499	0.2501	0.3502	0.4501	0.5499	0.6499	0.7499	0.8501	0.9301
Relative error (%)	−0.1742	−0.0910	0.0335	0.0454	0.0171	−0.0130	−0.0207	−0.0078	0.0075	0.0055

Finally, the developed LSSVM model was used to rank representative alternatives. Taking the objective function values corresponding to 20 representative alternatives as the input of the developed LSSVM model, the corresponding utility values were determined by the developed LSSVM model (Table 4). According to the magnitude of the utility values, the alternatives were ranked, and the final priority order of each representative alternative was obtained. The sequences for 20 representative clustered alternatives were given in Table 4. Based on the ranking results, the final decision can be made.

Table 4. The ranking results of representative alternatives based on LSSVM, RBF networks and information entropy.

Alternatives NO.	f_1	f_2	f_3	LSSVM		RBF Networks		Information Entropy
		(10^6 Million Yuan/Year)	(10^4 Tons/Year)	Results	Priority Order	Results	Priority Order	Priority Order
1	0.2545	1.71	65.2	0.441780	18	0.397256	18	18
2	0.6819	1.1	46.3	0.643288	10	0.637061	9	11
3	0.9351	0.84	38	0.732777	2	0.732997	2	6
4	0.7518	1	44.2	0.676973	7	0.673814	5	9
5	0.6397	1.16	47.8	0.620609	11	0.611983	10	12
6	1.1993	0.65	32.9	0.749499	1	0.742739	1	5
7	1.6858	0.46	28.2	0.689679	5	0.652028	7	3
8	0.8246	0.81	42.8	0.723348	3	0.721588	3	7
9	0.0994	1.96	76.2	0.416740	20	0.351093	20	20
10	0.1415	1.89	72.5	0.419814	19	0.360474	19	19
11	0.4118	1.48	57	0.502330	15	0.475338	15	15
12	0.7492	0.91	45	0.694340	4	0.690562	4	8
13	0.3446	1.54	60.1	0.481052	16	0.447618	16	16
14	1.7138	0.47	26.2	0.688663	6	0.647925	8	1
15	2.4148	0.35	20.1	0.584945	12	0.49719	14	2
16	0.4964	1.35	53.3	0.546941	14	0.527872	13	14
17	0.2535	1.63	65.3	0.452988	17	0.409534	17	17
18	0.4717	1.27	54.7	0.561099	13	0.540233	12	13
19	1.8766	0.53	24.5	0.655651	9	0.604544	11	4
20	0.662	1	47.9	0.660599	8	0.652785	6	10

Furthermore, comparative research was carried out between the LSSVM and the radial basis function (RBF) neural networks. The RBF network is a kind of feedforward neural network that is used

for classification problems, function approximation, noisy interpolation, and regularization. Due to its advantages over traditional multilayer perceptrons, namely faster convergence, smaller extrapolation errors, and higher reliability, it has increasingly attracted interest for engineering applications. Here, the RBF network is used as a regression tool to approximate function with a radial basis function as activation function. A more detailed review of the RBF network algorithm for nonlinear function estimation can be found in [35]. The RBF network model was developed using the same training dataset, testing dataset and normalization technique employed for the LSSVM. Based on trial and error, the final RBF network architecture was determined, which consisted of one input layer, one hidden layer with six neurons and one output layer. Its performance statistics were shown in Table 2. For the training data, RME and RMSE were 0.1448% and 3.211×10^{-4}, respectively, while RME = 0.1053% and RMSE = 3.159×10^{-4} for the testing data. Compared with the performance between the RBF network model and LSSVM model shown in Table 2, it is clear that although the RBF network and LSSVM have close performance during the training and testing periods, the LSSVM model, overall, has performed better than the RBF network model. The results obtained confirmed the capability of the LSSVM model for utility function estimates. Then, the RBF network model was applied to estimate utility values of representative alternatives, and according to the magnitude of the utility values, the alternatives were ranked, which are given in Table 4. Table 4 shows that the utility values of the RBF network were very close to those of the LSSVM with the maximum difference of 0.08775. As for ranking results, though there were ten ranking results that were not quite consistent, it should be pointed out that the differences of ranking results between the LSSVM and the RBF network were within the two levels. In general, ranking results obtained by the LSSVM and the RBF network were relatively consistent with each other.

Another comparison was conducted between the developed LSSVM and the information entropy approach. The same dataset and normalization technique were adopted as used in the above problem. Based on the information entropy approach, the weights for each objective can be determined, which was $w = (0.224, 0.405, 0.371)$. The superiority degrees were obtained as a linear weighted function of each objective standardized. According to the magnitude of the superiority degrees, the priority orders of representative clustered alternatives were determined, which are shown in Table 4. Table 4 shows that ranking results of the LSSVM model were essentially the same as those of the information entropy model. This is because both LSSVM and the information entropy model are essentially an objective weighting method, which directly applies distribution characteristics of sample data to determine the weights [36,37]. The difference is that in the information entropy method, the weight values of indicators are determined by calculating the entropy according to the degree of variation of each indicator, while in LSSVM, support vectors, which can be regarded as weights, are calculated by solving an optimization problem. From the angle of mathematical transformation, evaluating and ranking alternatives in water resource allocation is a kind of mapping from the high dimension space to the low. For the information entropy method, this process is regarded as a linear process, and the priority order is calculated as a linear weighted function of each objective; LSSVM uses a nonlinear process, which leads to slightly different result obtained by the two methods. The underlying reason for this requires further explanation.

In general, the comparison results with the information entropy method show that the proposed method used to evaluate the choice of scheme for the multi-objective water resource allocation problem is feasible and useful. Accordingly, it can be concluded that the LSSVM model based on the *k*-means cluster method can be used to carry out the decision analysis for the scheme choices for a multi-objective water resource allocation problem.

4. Conclusions

In this paper, a new method based on LSSVM and *k*-means clustering was proposed for ranking the optimal solutions generated by MOEAs for multi-objective water resources allocation. By using the proposed method, the utility value for each representative alternative can be obtained. Meanwhile,

according to the magnitude of the utility values, the final priority order of the representative solutions can be determined, and a suitable policy can be implemented. Then, the proposed approach was applied to rank the Pareto optimal solution set obtained from the multi-objective optimal allocation of water resources in the water-receiving areas of the South-to-North Water Transfer Project in Hebei Province, China. The results can be used for more effective decision-making processes regarding water supply, environment and economic cost. Moreover, the application in water-receiving areas demonstrates the effectiveness in supporting relevant decision-making processes for water resource allocation, compared to the information entropy method and the RBF network. Accordingly, the proposed approach can provide a useful tool for water resource managers during decision making processes regarding rational allocation of water resources. It is also a viable alternative method for relevant decision making processes for multi-objective water resource and hydrology problems.

Nevertheless, the proposed method is based on a deterministic evaluation. Considering the uncertainty in the decision making processes of water resource allocation due to a lack of data and changing environmental factors, further work can focus on defining uncertainties associated with the decision making process and ultimately improve the quality of decisions. Moreover, since the ranking results of different methods may be inconsistent in practical applications, some aggregation methods such as Borda and Copeland should be used to integrate the ranking results, and to enhance applicability and accuracy.

Acknowledgments: This work was financially supported by CRSRI Open Research Program (CKWV2013218/KY), National Natural Science Foundation of China (No. 51309130), the Natural Science Foundation of Jiangxi Province (20132BAB213025) and Open Foundation of Jiangxi Engineering Research Center of Water Engineering Safety and Resources Efficient Utilization (OF201610). Thanks to Kuolin Hsu from Center for Hydrometeorology and Remote Sensing, University of California, Irvine, for helpful discussion.

Author Contributions: Weilin Liu and Lina Liu had the original idea, developed the models and carried out the model simulations. All authors analyzed the results and wrote the paper.

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

References

1. Liu, D.; Chen, X.; Lou, Z. A model for the optimal allocation of water resources in a saltwater intrusion area: A case study in Pearl River Delta in China. *Water Resour. Manag.* **2010**, *24*, 63–81. [CrossRef]
2. Liu, W.; Liu, L.; Dong, Z. On the use of multi-objective particle swarm optimization for allocation of water resources. In Proceedings of the 2013 Ninth International Conference on Natural Computation (ICNC), Shenyang, China, 23 July 2013; pp. 612–617.
3. Baltar, A.M.; Fontane, D.G. Use of multiobjective particle swarm optimization in water resources management. *J. Water Resour. Plan. Manag.* **2008**, *134*, 257–265. [CrossRef]
4. Yeh, W.W.G.; Becker, L. Multiobjective analysis of multireservoir operations. *Water Resour. Res.* **1982**, *18*, 1326–1336. [CrossRef]
5. Salman, A.Z.; Al-Karablieh, E.K.; Fisher, F.M. An inter-seasonal agricultural water allocation system (SAWAS). *Agric. Syst.* **2001**, *68*, 233–252. [CrossRef]
6. Shangguan, Z.P.; Shao, M.A.; Horton, R; Lei, T.; Qin, L.; Ma, J. A model for regional optimal allocation of irrigation water resources under deficit irrigation and its applications. *Agric. Water Manag.* **2002**, *52*, 139–154. [CrossRef]
7. Deb, K. *Multi-Objective Optimization Using Evolutionary Algorithms*; Wiley: New York, NY, USA, 2001.
8. Coello, C.C.; Lamont, G.B.; Van Veldhuizen, D.A. *Evolutionary Algorithms for Solving Multi-Objective Problems*, 2nd ed.; Springer: New York, NY, USA, 2007.
9. Reed, P.M.; Hadka, D.; Herman, J.D.; Kasprzyk, J.R.; Kollat, J.B. Evolutionary multiobjective optimization in water resources: The past, present, and future. *Adv. Water Resour.* **2013**, *51*, 438–456. [CrossRef]
10. Liu, D.; Guo, S.; Chen, X.; Shao, Q.; Ran, Q.; Song, X.; Wang, Z. A macro-evolutionary multi-objective immune algorithm with application to optimal allocation of water resources in Dongjiang River basins, South China. *Stoch. Environ. Res. Risk Assess.* **2012**, *26*, 491–507. [CrossRef]

11. Kim, T.; Heo, J.-H.; Jeong, C.S. Multi-reservoir system optimization in the Han River basin using multi-objective genetic algorithms. *Hydrol. Process.* **2006**, *20*, 2057–2075. [CrossRef]
12. Guo, J.; Zhou, J.; Zou, Q.; Liu, Y.; Song, L. A novel multi-objective shuffled complex differential evolution algorithm with application to hydrological model parameter optimization. *Water Resour. Manag.* **2013**, *27*, 2923–2946. [CrossRef]
13. Reddy, M.J.; Nagesh Kumar, D. Multi-objective particle swarm optimization for generating optimal trade-offs in reservoir operation. *Hydrol. Process.* **2007**, *21*, 2897–2909. [CrossRef]
14. Liu, W.; Liu, L. Optimal Allocation of Water Resources Based on Multi-Objective Particle Swarm Algorithm and Information Entropy. *Appl. Mech. Mater.* **2014**, *641*, 75–79. [CrossRef]
15. Wang, Q.; Shen, Y.P.; Chen, Y.W. The Support vector machine method for multiple attribute decision making. *Syst. Eng. Theory Pract.* **2006**, *6*, 54–58.
16. Behzad, M.; Asghari, K.; Eazi, M.; Palhang, M. Generalization performance of support vector machines and neural networks in runoff modeling. *Expert Syst. Appl.* **2009**, *36*, 7624–7629. [CrossRef]
17. Guo, J.; Zhou, J.; Qin, H.; Zou, Q.; Li, Q. Monthly streamflow forecasting based on improved support vector machine model. *Expert Syst. Appl.* **2011**, *38*, 13073–13081. [CrossRef]
18. Vapnik, V. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, 1995.
19. Cortes, C.; Vapnik, V. Support vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
20. Wen, X.; Si, J.; He, Z.; Wu, J.; Shao, H.; Yu, H. Support-vector-machine-based models for modeling daily reference evapotranspiration with limited climatic data in extreme arid regions. *Water Resour. Manag.* **2015**, *29*, 3195–3209. [CrossRef]
21. Su, C.; Wang, L.; Wang, X.; Huang, Z.; Zhang, X. Mapping of rainfall-induced landslide susceptibility in Wencheng, China, using support vector machine. *Nat. Hazards* **2015**, *76*, 1759–1779. [CrossRef]
22. Widodo, A.; Yang, B.S. Support vector machine in machine condition monitoring and fault diagnosis. *Mech. Syst. Signal Process.* **2007**, *21*, 2560–2574. [CrossRef]
23. Mohammadpour, R.; Shaharuddin, S.; Chang, C.K.; Zakaria, N.A.; Ab Ghani, A.; Chan, N.W. Prediction of water quality index in constructed wetlands using support vector machine. *Environ. Sci. Pollut. Res.* **2015**, *22*, 6208–6219. [CrossRef] [PubMed]
24. Suykens, J.A.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [CrossRef]
25. Samui, P.; Kothari, D.P. Utilization of a least square support vector machine (LSSVM) for slope stability analysis. *Sci. Iran.* **2011**, *18*, 53–58. [CrossRef]
26. Mellit, A.; Pavan, A.M.; Benghanem, M. Least squares support vector machine for short-term prediction of meteorological time series. *Theor. Appl. Climatol.* **2013**, *111*, 297–307. [CrossRef]
27. Ismail, S.; Samsudin, R.; Shabri, A. River flow forecasting: A hybrid model of self organizing maps and least square support vector machine. *Hydrol. Earth Syst. Sci. Discuss* **2010**, *7*, 8179–8212. [CrossRef]
28. Shabri, A. Suhartono. Stremflow forecasting using least-squares support vector machines. *Hydrol. Sci. J.* **2012**, *57*, 1275–1293. [CrossRef]
29. Deka, P.C. Support vector machine applications in the field of hydrology: A review. *Appl. Soft Comput.* **2014**, *19*, 372–386.
30. Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666. [CrossRef]
31. Cao, M.; Zhang, S.; Yin, Y.; Shao, L. Classification and the case matching algorithm of the blast furnace burden surface. In Proceedings of the AIP (American Institute of Physics) Conference, Wuhan, China, 25 February 2017; AIP Publishing: Melville, NY, USA, 2017; p. 080009.
32. Yu, L.; Yao, X.; Wang, S.; Lai, K.K. Credit risk evaluation using a weighted least squares SVM classifier with design of experiment for parameter selection. *Expert Syst. Appl.* **2011**, *38*, 15392–15399. [CrossRef]
33. De Brabanter, K.; Karsmakers, P.; Ojeda, F.; Alzate, C.; De Brabanter, J.; Pelckmans, K.; Suykens, J.A.K. LS-SVMlab toolbox user's guide. *ESAT-SISTA Technol. Rep.* **2011**, *10*, 146.
34. Chamkalani, A.; Zendehboudi, S.; Bahadori, A.; Kharrat, R.; Chamkalani, R.; James, L.; Chatzis, I. Integration of LSSVM technique with PSO to determine asphaltene deposition. *J. Pet. Sci. Eng.* **2014**, *124*, 243–253. [CrossRef]
35. Moradkhani, H.; Hsu, K.L.; Gupta, H.V.; Sorooshian, S. Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. *J. Hydrol.* **2004**, *295*, 246–262. [CrossRef]

36. Zhao, J.; Jin, J.; Zhu, J.; Xu, J.; Hang, Q.; Chen, Y.; Han, D. Water Resources Risk Assessment Model based on the Subjective and Objective Combination Weighting Methods. *Water Resour. Manag.* **2016**, *30*, 3027–3042. [CrossRef]
37. Jia, X.; Li, C.; Cai, Y.; Wang, X.; Sun, L. An improved method for integrated water security assessment in the Yellow River basin, China. *Stoch. Environ. Res. Risk Assess.* **2015**, *29*, 2213–2227. [CrossRef]