Research on Investment Risk Evaluation of Water Conservancy Project based on Support Vector Machine

Wu Fengping, Pan Wenwen
Department of management science and engineering, School of Business, Hohai University, Nanjing, China, cindypanwen@163.com

Abstract
The paper analyzes characteristics and investment risk of water conservancy project. Based on support vector machine (SVM), the paper explores water conservancy project investment risk evaluation method, establishes the SVM investment risk evaluation mode. The particle swarm optimization algorithm is employed to analyze data efficiently. Detailed empirical results shows effectiveness of the proposed scheme.

Keywords: Water Conservancy Project; Support Vector Machine; Particle Swarm Optimization; Investment Risk Evaluation

1. Introduction
Risk analysis as a technique has been widely used in industrial systems and other field of finance and insurance projects [1, 2]. In recent years, the risk analysis used in decision-making and management of the project has been made a lot of development [3, 4]. But the problem of cost overrun and the duration extension exist in the project due to the environment has become increasingly complex [5, 6]. The water conservancy project has the characteristic of the huge amount of investment, technical complexity and long duration [7, 8]. So many uncertainties exist in the project implementation process, which means that there is a huge risk in the project [9, 10]. The risk of water conservancy project has the following characteristics: firstly, the objectivity and universality; secondly, a specific risk contingency and a number of risks inevitability; thirdly, the diversity and multilevel [11, 12, 13].

One of the key steps of project risk management is the investment risk evolution. Risk evolution is generally operated in the prophase of engineering construction and lack of concrete data [14, 15]. The main basis of the evolution is expert experience and the intention of the decision-makers. The results of the evaluation are also kind of roughly to some extent [16, 17]. Common risk evolution methods usually consist of scoring method, fuzzy mathematics method, analytical hierarchy process and etc. However, the common deficiencies of those methods is that in the performance evaluation process the weight of factors mainly set by people and cause the inaccuracy of evolution results [18, 19, 20]. In order to solve this problem, this paper build investment risk evaluation of water conservancy project index system by adopting one risk evaluation method which based on support vector machine and establishes investment risk evaluation model using particle swarm optimization algorithm [21, 22]. The SVM was used in investment risk evaluation model and the improved SVM algorithm was put forward in this paper by using POS to looking optimal parameter. Finally, the model was verified by using data of 28 provincial towns water supply projects in Jiangsu province [23].

The paper analyzes characteristics and investment risk of water conservancy project. Based on support vector machine (SVM), the paper explores water conservancy project investment risk evaluation method, establishes the SVM investment risk evaluation mode. The particle swarm optimization algorithm is employed to analyze data efficiently. Detailed empirical results shows effectiveness of the proposed scheme.

2. Proposed Scheme

2.1. Investment risk
The pure public welfare infrastructure such as flood control and drainage, irrigation, flood control and water resources protection which belongs to public products. It is in the field of market mechanism completely out of order. So the government directly provides this type of product and service.

The social benefit infrastructure such as water delivery, hydroelectric power generation, reservoir, water tourism and water diversified mode which belongs to impure public goods or mixed public goods. The government should fully follow the market mechanism to consider the characteristic and function of impure public goods. The enterprises and social investors can act as the main body of investment construction, operation and maintenance. The government should make the stringent access market conditions to bid and choose the main body of investment construction, operation and maintenance. The government provides economic regulation such as financial subsidies, tax concessions and the regulation of related products and services prices or fees to investors and operators. The government also regulates profit and enhances the active participation of the investors and operators.

Investment can produce two effects: one is the demand effect; the other is the supply effect. Demand effect refers to the demand activity attendant on investment cycle. Supply effect is the formation of new production capacity that is as the result of investment. The supply effect achieves to reveal while fixed assets have a combination with floating capital. Supply effect is lagging behind the demand effect. Time lag effect of Water conservancy investment is that its benefit is demonstrated after the following years of the investment of irrigation works capital. The delay years depend on the specific engineering. It caused by characteristics of difficult project because of water conservancy industry investment cycle is long.

Our country is in shunt period of the planned economy to market economy. The market mechanism has not yet formed, price system is not perfect, and the enterprise has not become the main body of the market. Input and output is non-equilibrium and externality due to two reasons. one is the separation of investors and beneficiaries, another is that benefit is not immediately reflected because of long construction period. It is also an important cause of unbalanced industry of irrigation works.

The three phases of long term Investment cycle are longer decision phase, the appropriate construction phase as well as the recovery phase which should be as short as possible. In order to guarantee the yield of investment, investors hope to recover investment as soon as quickly. A short recovery period is based on the correct scientific investment decision and construction quality. Water conservancy investment cycle consists of three phases, namely decision-making period, construction period, recovery period.

Investment risk of water conservancy project may exist in the various stages of project investment and may also exist in various parts of the project investment. The situation is a concept of thinking form. It refers to the work in progress, and the results cannot be expected in a situation. Water conservancy project investment risk can be defined as followed. At various stages of the of water conservancy project investment, main body of water conservancy project investment get specific benefits as the goal. But the uncertainty water conservancy project itself may not be required to achieve to the target, which can cause loss of things called water project investment risk.

2.2. Risk evaluation model

The basic idea of the investment risk analysis method of water conservancy project based on support vector machine regression as followed: indicators reflected the investment risk of the project regard as a support vector machine input vector; the evaluation criteria value of the investment risk regard as output of support vector machine. Define a ε loss of function to the train support vector machine and make the different input vector to get a different output values. After adaptive learning, the series of right values and coefficient values that connect different support vector can be gotten. If the situation meets the training requirements, support vector machine can be used as an effective tool for investment risk evaluation of water conservancy projects.

Investment risk evaluation of water conservancy project is a complex systems engineering. It involves social, economic, technical, environment and other aspects. The choice of investment risk evaluation indexes is the premise and key of risk evaluation. The indexes should follow four principles according to the concept of risk framework. They are objective principles, systematic principles, scientific principles and maneuverability principles respectively. The investment risk evaluation index system can be divided into the target layer, rule layer and index layer. The paper establishes the
multi-criteria, multi-level index system. The 10 indexes are listed to describe the investment risk as Figure 1.

![Figure 1. Investment Risk Evaluation Index System](image)

### 2.3. Model construction

**HYPOTHESIS 1:** set the factors of investment risk evaluation of water conservancy project as \( m \), evaluation levels as \( k \) and study sample as \( l \). Then the sample set is \( s \) \( \{x_i, y_i\}, x \in \mathbb{R}^n; y \in \mathbb{R}; i = 1, 2, \cdots, l \).

This paper introduced the improved \( \varepsilon-SVR \) model to evaluate investment risk. The \( \varepsilon-SVR \) model is the solution of an optimal regression hyper-plane and can be deducted to a two convex programming problem:

\[
\min \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{l} (\xi_i^+ + \xi_i^-) .
\]

Subject To:

\[
\begin{align*}
&y_i - \omega \cdot \varphi(x_i) - b \leq \varepsilon + \xi_i^+, \\
&\omega \cdot \varphi(x_i) + b - y_i \leq \varepsilon + \xi_i^-, \\
&\xi_i^+, \xi_i^- \geq 0, i = 1, 2, \cdots, l
\end{align*}
\]

\( \xi \) corresponding to the sample points above the regression hyper-plane and \( \xi^- \) corresponding to the sample points below the regression hyper-plane. \( C \) is a given parameter. Since the objective function and constraints are convex sets, based on the optimization theory, the global optimization solution is uniquely existed. Using the Lagrange multiplier method and applying conditions of KKT (Karush-Kuhn-Tucker), the optimal regression hyper-plane can be obtained. Firstly the definitions of Lagrange function can be described as follows:

Firstly the definitions of Lagrange function as follows:

\[
L(\omega, b, \xi, \xi^-) = \frac{1}{2} \| \omega \|^2 + C \sum_{i=1}^{l} (\xi_i^+ + \xi_i^-) - \sum_{i=1}^{l} \alpha_i (\varepsilon + \xi_i - y_i + \omega \varphi(x_i) + b) - \sum_{i=1}^{l} \eta_i (\xi_i^+ - \varepsilon^+) 
\]

For (2), \( \omega, b, \xi, \) and \( \xi^- \) are calculated partial derivative and the partial derivative is zero.

\[
\frac{\partial L}{\partial \omega} = 0 \Rightarrow \omega - \sum_{i=1}^{l} (\alpha_i - \alpha_i^-) \varphi(x_i) = 0 ,
\]

\[
\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^{l} (\alpha_i - \alpha_i^-) = 0 ,
\]
\[
\frac{\partial L}{\partial \xi_i} = 0 \Rightarrow C - \alpha_i - \eta_i = 0, \quad (6)
\]
\[
\frac{\partial L}{\partial \xi_i^*} = 0 \Rightarrow C - \alpha_i^* - \eta_i^* = 0, \quad (7)
\]
where \(\alpha, b, \xi, \xi^*, \eta, \eta^*\) can be eliminated by the above four formulas into (2), dual optimization problem (1):
\[
\max \sum_{i=1}^{l} y_i (\alpha_i - \alpha_i^*) - \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) K(x_i, x_j) - \frac{1}{4C} \sum_{i=1}^{l} (\alpha_i^* + \alpha_i^*) , \quad (8)
\]
Subject to:
\[
\begin{aligned}
\sum_{i=1}^{l} (\alpha_i - \alpha_i^*) = 0, \\
[\alpha_i, \alpha_i^*] \subseteq [0, C]
\end{aligned}
\quad (9)
\]
Get optimal linear regression hyper plane analytic formula in the feature space.
\[
\begin{align*}
f(x) &= \sum_{i=\text{SV}} (\alpha_i^* - \alpha_i) K(x_i, x) + b^* , \\
b^* &= \frac{1}{N_{\text{SV}}} \left\{ \sum_{b, \alpha_i^* \in \mathcal{C}} (y_i - \sum_{j \in \text{SV}} (\alpha_j^* - \alpha_j) K(x_i, x_j) - \frac{\alpha_i}{2C}) + \sum_{b, \alpha_i \in \mathcal{C}} (y_i - \sum_{j \in \text{SV}} (\alpha_j^* - \alpha_j) K(x_i, x_j) + \frac{\alpha_i}{2C}) \right\} ,
\end{align*}
\]
where \(N_{\text{SV}}\) is the standard support vector quantity.

This is the optimal regression function of support vector machine in investment risk evaluation of water conservancy project. The values of support vector corresponding to the coefficients \(\alpha\) and \(b\) should be calculated by nuclear function calculation and then the original sample space of nonlinear regression function can be received.

Step 1: Evaluate the investment risk level and grading standards of single index, use stochastic simulation technology to generate a sufficient number of evaluation index sequences, and construct modeling sequence with generated sequence of these indicators and its rating value. The lower limitation of index values of the \(K\) evaluation grade is \(b_j\), upper is \(a_j\), and the corresponding evaluation grade value is \(y_i\), \(i = 1, 2, \cdots, N\); \(j = 1, 2, \cdots, m\); \(k = 1, 2, \cdots, K\).

\(n_i\) is the \(K\) evaluation grade generate index sequence capacity, \(m\) is Index number, \(K\) is evaluation grade number. Evaluation index of random simulation formula:
\[
x_{ij}^k = \text{RAND}(n_i) (b_j - a_j) + a_j , \quad (12)
\]
The \(K\) evaluation grade will generate \(n_i\) group of \((x_{ij}^k, y_i^k)\) by (3). For all \((x_{ij}^k, y_i^k)\), \(k = 1, 2, \cdots, K\). If it rearranges the subscript, we will get available sequence \(j, i = 1, 2, \cdots, N; j = 1, 2, \cdots, m\).

Step 2: The investment risk evaluation samples are normalized. In order to eliminate sample data dimension of the risk evaluation and simplify operation, we can use extreme value type normalization.
\[
x_{ij} = \frac{x_{ij} - m_j}{M_j - m_j} , \quad (13)
\]
Step 3: Use the \(\varepsilon - SVR\) model to construct nonlinear mapping relationship. Based on the basic theory of support vector machine, the corresponding mapping model is as followed:
\[
f(x) = \sum_{i=\text{SV}} (\alpha_i^* - \alpha_i) K(x_i, x) + b^* , \quad (14)
\]
where \(x_i\) is support vector sample after study training samples. \(x\) is test sample. \(\alpha_i^*\), \(b^*\) are knowledge of investment risk evaluation after study the training samples. Through analysis and
comparison, the paper select RBF kernel function to establish the risk evaluation model and use the particle swarm optimization algorithm to search the optimal kernel function parameter $\sigma^2$ and penalty factor $C$.

Step 4: Use particle swarm optimization algorithm (PSO) to search the optimal kernel function parameter $\sigma^2$ and penalty factor $C$.

Step 5: Calculate parameter values by the PSO algorithm optimization and set up support vector machine model. If the accuracy meets the requirements, the model proceeded. If the precision does not meet requirements, return to step 4.

Step 6: Use the support vector machine model to evaluate to determine quota samples of investment risk evaluation.

3. Empirical results

The evaluation objects are 28 water reservoir projects in Jiangsu province. The paper uses the above model to evaluate the investment risk of the projects for the sake of provision the scientific theories and method for the future water conservancy projects.

3.1. Risk evaluation degree of investment

Investment risk evaluation of water conservancy project plays an important role in the construction and management of the project. At present, many countries and regions have established a relatively mature evolution system. The paper refers to the large number of literatures to summarize the investment risk degree of project and divides the investment risk degree into five grades. Table 1 shows each standard and its feature. The each standard respectively is higher risk, high risk, moderate risk, low risk and lower risk.

<table>
<thead>
<tr>
<th>Risk degree</th>
<th>Evaluation grade</th>
<th>Feature description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher risk</td>
<td>I</td>
<td>Investment return rate is extremely low, debt paying ability is poor.</td>
</tr>
<tr>
<td>High risk</td>
<td>II</td>
<td>Investment return rate is low, debt paying is poor.</td>
</tr>
<tr>
<td>Moderate risk</td>
<td>III</td>
<td>Investment return rate is general, debt paying ability is general.</td>
</tr>
<tr>
<td>Low risk</td>
<td>IV</td>
<td>Investment return rate is high, debt paying ability.</td>
</tr>
<tr>
<td>Lower risk</td>
<td>V</td>
<td>Investment return rate is higher, debt paying ability is strong.</td>
</tr>
</tbody>
</table>

Corresponding to five levels above, the paper selects the appropriate division of indexes. There is not significant and authoritative research on the division of project risk. So the Delphi method is used to get the division of indexes by expert scoring.

3.2. Evaluation samples of the investment risk

The investment risk evaluation indexes value are usually based on the project proposal, project feasibility report, project evaluation reports and other information collected. If each index dimension is not the same, the large difference should be between the indexes. Indicators values directly are entered into a support vector machine learning that is likely to reduce the performance of support vector machine learning. Therefore, the training samples should be normalized. The indicators should be linearly stretched to the interval [0, 1].

According to the established evaluation index system, 10 indexes were selected as the inputs of the model and evaluated level as the output, and the 5 ratings randomly generate a total of 125 sample values. The paper choose a total of 110 sample values as SVM training sample and the remaining samples as test samples. Kernel function use radial basis function. After being trained, test samples will be verified this model, parameter optimization selects particle swarm optimization algorithm, fit the training samples, see Figure 2 and 3.
3.3. Parameter calculation

The paper applies particle swarm optimization algorithm (PSO) and the GRID parameter optimization method to seek a better parameter of kernel function $\sigma^2$ and penalty factor $C$. The choice of GRID parameter optimization shows the superiority of PSO to find the parameters. See figure 4, $c$ is on behalf of the penalty parameter, and $g$ is on behalf of the kernel function parameter. Compared figure 4, regression process shows that using PSO find parameter optimization more efficient and shorter time than using GRID parameter optimization method.

The paper take the population number is 100, the evolution of termination algebra is 100, and mean squared error is used as fitness function. According to the step 4 and step 5 ultimately determine the optimal penalty factor is $75.8346$, the parameter of kernel function is $14.1754$, the loss function parameters is $0.025594$. 

Figure 2. Training Sample Fitting

Figure 3. Test Sample Prediction

Figure 4. GRID Optimization Curve
3.4. Investment risk evaluation results

This paper uses the LIBSVM to evaluate 28 water reservoir projects in Jiangsu province. The results of Evaluation show Table 2. According to SVM output values in table 2 and the present situation investment risk of water conservancy projects in Jiangsu; this paper determines the investment risk evaluation benchmark, see Table 3.

<table>
<thead>
<tr>
<th>Project number</th>
<th>SVM value</th>
<th>Risk grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.753</td>
<td>II</td>
</tr>
<tr>
<td>2</td>
<td>2.375</td>
<td>II</td>
</tr>
<tr>
<td>3</td>
<td>2.174</td>
<td>II</td>
</tr>
<tr>
<td>4</td>
<td>4.058</td>
<td>IV</td>
</tr>
<tr>
<td>5</td>
<td>3.819</td>
<td>IV</td>
</tr>
<tr>
<td>6</td>
<td>2.894</td>
<td>III</td>
</tr>
<tr>
<td>7</td>
<td>2.367</td>
<td>II</td>
</tr>
<tr>
<td>8</td>
<td>4.152</td>
<td>IV</td>
</tr>
<tr>
<td>9</td>
<td>2.847</td>
<td>III</td>
</tr>
<tr>
<td>10</td>
<td>3.057</td>
<td>III</td>
</tr>
<tr>
<td>11</td>
<td>3.386</td>
<td>III</td>
</tr>
<tr>
<td>12</td>
<td>3.165</td>
<td>III</td>
</tr>
<tr>
<td>13</td>
<td>3.247</td>
<td>III</td>
</tr>
<tr>
<td>14</td>
<td>2.828</td>
<td>III</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Project number</th>
<th>SVM value</th>
<th>Risk grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>15.065</td>
<td>IV</td>
</tr>
<tr>
<td>16</td>
<td>3.068</td>
<td>III</td>
</tr>
<tr>
<td>17</td>
<td>3.272</td>
<td>III</td>
</tr>
<tr>
<td>18</td>
<td>3.867</td>
<td>IV</td>
</tr>
<tr>
<td>19</td>
<td>4.352</td>
<td>IV</td>
</tr>
<tr>
<td>20</td>
<td>4.281</td>
<td>IV</td>
</tr>
<tr>
<td>21</td>
<td>4.096</td>
<td>IV</td>
</tr>
<tr>
<td>22</td>
<td>4.269</td>
<td>IV</td>
</tr>
<tr>
<td>23</td>
<td>4.056</td>
<td>IV</td>
</tr>
<tr>
<td>24</td>
<td>3.521</td>
<td>IV</td>
</tr>
<tr>
<td>25</td>
<td>3.731</td>
<td>IV</td>
</tr>
<tr>
<td>26</td>
<td>3.870</td>
<td>IV</td>
</tr>
<tr>
<td>27</td>
<td>4.071</td>
<td>IV</td>
</tr>
<tr>
<td>28</td>
<td>4.375</td>
<td>IV</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output value</th>
<th>0.5~1.5</th>
<th>1.5~2.5</th>
<th>2.5~3.5</th>
<th>3.5~4.5</th>
<th>4.5~5.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk degree</td>
<td>Higher</td>
<td>High</td>
<td>Moderate</td>
<td>Low</td>
<td>Lower</td>
</tr>
</tbody>
</table>

Table 3. Investment risk interval

4. Conclusion

We can draw the conclusion from table 2 and table 3 that he smaller SVM output value, the greater the investment risk of water conservancy project. Jiangsu province has 15 provincial towns water supply project and their investment risk that is relatively low. It shows that the projects have a higher rate of return on investment. The results can provide reference to the future water conservancy project investment and financing. There are many investment risk factors of water conservancy project. The aspects of investment risk must be objectively evaluated by effective methods to ensure the successful project. Analysis showed that the SVM has a weak dependence on the number of samples. Limited sample is learned to establish the model which still has very strong generalization ability. The model solves the inaccurate problem of human factors that make risk assessment value. Therefore, the application of SVM in investment risk prediction and evaluation of water conservancy project will have very broad prospects.

5. References


