1. Introduction

The establishment of an accurate deformation prediction model based on prototype observation data is of great significance for determining the operation behavior of the dam and ensuring its long-term safe operation \[1\]. Conventional statistical models are difficult to adapt to the complex nonlinear relationship between multiple factors and effect sizes \[2\], which makes it difficult to guarantee the accuracy of model predictions. Support vector machines have obvious advantages in solving nonlinear problems with few samples and high dimensions. However, when establishing the SVM-based dam deformation prediction model, the optimization of the kernel parameters and penalty factors is a problem worthy of study.  

Grey Wolf Optimizer (GWO) is a new intelligent optimization algorithm proposed by Mirjalili et al. \[3\] with reference to the social hierarchy and hunting behavior of grey wolves. The algorithm realizes intelligent optimization through the process of tracking, encircling, hunting, and attacking the grey wolf population. Studies have shown that the GWO algorithm is superior to other evolutionary algorithms in terms of the quality, speed and stability of the final solution \[4\]. However, the possibility of premature convergence reduces the probability of the algorithm finding the global optimum.

This research proposes a dam deformation prediction model (SVM-IGWO) based on improved grey wolf algorithm optimization support vector machine. The initial population is enriched by the difference algorithm, and the crossover and mutation operators of the difference algorithm are introduced to
improve the development and search capabilities of the grey wolf algorithm. The improved grey wolf optimization algorithm is used to optimize the parameters of SVM, and the SVM-IGWO model is proposed and applied to dam deformation prediction. Taking the Xiluodu super high arch dam as an example, the prediction results of SVM-IGWO model, SVM and SVM-GWO model are compared and analyzed. Engineering practice shows that compared with the SVM-IGWO model and the SVM and SVM-GWO models, the SVM-IGWO model has stronger generalization ability and higher prediction accuracy.

2. Method

2.1 Grey Wolf Optimization Algorithm

The grey wolf optimizer (GWO) algorithm simulates the hierarchy and hunting behavior of grey Wolves. Grey wolves have a very strict social dominance hierarchy, which is mainly divided into four parts: $\alpha$, $\beta$, $\delta$ and $\omega$. $\alpha$ is the best solution, followed by $\beta$ and $\delta$, and the remaining solutions belong to $\omega$. The top three best wolves that are closest to their prey are $\alpha$, $\beta$ and $\omega$, and they guide $\omega$ to search for prey in promising search areas. During the hunting process, the wolf will update its position around $\alpha$, $\beta$ and $\omega$, as shown in Eq. (1) and Eq. (2).

$$
D = C \cdot \bar{X}_{p(t)} - \bar{X}(t) 
$$

$$
\bar{X}(t+1) = \bar{X}_{p(t)} - A \cdot D
$$

Where: $t$ is the current iteration number, $\bar{X}_{p(t)}$ is the current position of the prey, $\bar{X}(t)$ is the current position of the wolf, and $D$ is the distance between the wolf and the prey.

The mathematical expressions of the coefficient vectors $A$ and $C$ are as follows:

$$
\bar{A} = 2 \bar{a} \cdot \bar{r}_1 - \bar{a}
$$

$$
\bar{C} = 2 \bar{r}_2
$$

Where: $\bar{r}_1$ and $\bar{r}_2$ are two vectors randomly generated between [0, 1], and the convergence factor $\bar{a}$ linearly decreases from 2 to 0 in the iterative process.

The position of the grey wolf is updated according to the following equation:

$$
\begin{align*}
\bar{D}_\alpha &= \bar{C}_1 \cdot \bar{X}_\alpha - \bar{X} \\
\bar{D}_\beta &= \bar{C}_2 \cdot \bar{X}_\beta - \bar{X} \\
\bar{D}_\delta &= \bar{C}_3 \cdot \bar{X}_\delta - \bar{X} \\
\bar{X}_1 &= \bar{X}_\alpha - A_1 \cdot (\bar{D}_\alpha) \\
\bar{X}_2 &= \bar{X}_\beta - A_2 \cdot (\bar{D}_\beta) \\
\bar{X}_3 &= \bar{X}_\delta - A_3 \cdot (\bar{D}_\delta) \\
\bar{X}(t+1) &= \frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3}{3}
\end{align*}
$$

Where: $\bar{X}_\alpha$ represents the location of $\alpha$, $\bar{X}_\beta$ represents the location of $\beta$, $\bar{X}_\delta$ represents the location of $\delta$, $\bar{X}$ represents the location of the current solution, $\bar{C}_1$, $\bar{C}_2$ and $\bar{C}_3$ represent randomly generated vectors, and $\bar{D}_\alpha$, $\bar{D}_\beta$ and $\bar{D}_\delta$ represent the distances of $\alpha$, $\beta$ and $\delta$ from other grey wolves, respectively. According to equation (5) - (7), the final position of the current solution is calculated after
defining the distance, where $\tilde{\mathbf{A}}$, $\tilde{\mathbf{A}}$, and $\tilde{\mathbf{A}}$ are random vectors, $\mathbf{X}(t+1)$ represents the final position of $\mathbf{\omega}$, and $t$ represents the current iteration number.

2.2 Improved grey wolf optimization algorithm

The standard grey wolf algorithm randomly generates the initial population, which may fall into the dilemma of local optimization due to the inability to ensure the diversity of the population. DE algorithm generates group intelligence through mutual cooperation and competition between individuals, which guides the direction of optimization search [5]. Based on the respective advantages and disadvantages of the GWO and DE algorithms, the diversity of the random initial population is enriched by the mutation of differential evolution, the global optimal search ability of the GWO algorithm is used to improve the convergence of the combined algorithm, and a more efficient improved algorithm (IGWO) is proposed. Specific steps are as follows:

1. Set the relevant parameters of the IGWO optimization algorithm, such as the population size $N$, the maximum number of iterations $C_R$, the upper and lower bounds of the search range $u_b$ and $l_b$, etc.

2. Initialize the parameters $a$, $A$ and $C$. Generate intermediates (variant populations) through evolutionary mutation operations, and then generate initial population individuals through competitive selection operations, and set the iteration time $t = 1$.

3. Calculate the objective function value of a single grey wolf individual, sort according to the value of the objective function, and select the best three individuals as $\mathbf{X}_\alpha$, $\mathbf{X}_\beta$ and $\mathbf{X}_\delta$ respectively.

4. Calculate the distance between other grey wolf individuals and the optimal $\mathbf{X}_\alpha$, $\mathbf{X}_\beta$ and $\mathbf{X}_\delta$ according to Eq. (5), and update the position of each grey wolf according to Eq. (6).

5. Update the values of $a$, $A$ and $C$. The crossover operation is applied to individual positions in the population to retain better components, new individuals are generated through competitive selection operations, and the objective function values of all grey wolves are calculated.

6. The positions of $\mathbf{X}_\alpha$, $\mathbf{X}_\beta$ and $\mathbf{X}_\delta$ of the first three grey wolf individuals are updated.

7. Determine whether the maximum number of iterations $t_{max}$ has been reached. If yes, exit the algorithm and output the global optimal objective function value; otherwise, $t = t + 1$, and move to the third step to continue.

3. Construction process of SVM-IGWO-based model predicting dam deformation

The selection of SVM parameters for regression analysis has great effect on the approximation and generalization ability of SVM. Grid search method is usually applied to the optimization of penalty factor and kernel function parameter. Generally, the search for optimal parameter combination from a series of parameter combinations may lead to high computational burden and long operation time. So some intelligent optimization algorithms have been used to implement the optimization of SVM parameters. In this paper, IGWO is used to enhance the learning performance of the SVM model. A hybrid model is proposed, which uses an improved grey wolf optimization algorithm to select the optimal parameters of the support vector machine (SVM-IGWO) for concrete dam deformation analysis and prediction. The dam deformation predicting based on SVM-IGWO is described as follows and shown in Fig. 1.

1. Obtain dam deformation data.

2. Divide the data set, obtain the training data set and the test data set respectively, and normalize the data set.

3. Initialize the parameters of the grey wolf population. Set the parameters such as the population size and the maximum number of iterations to determine the scope of the search space, and at the same time encode the penalty coefficient $C$ and the kernel parameters $\gamma$ to become the two dimensions of the position in the wolf pack.

4. Use DE algorithm to initialize grey wolf population.

5. Initialize $a$, $A$ and $C$. Use equations (3) and (4) to initialize the coefficients $A$ and $C$, and the
convergence factor $a = 2$.

6. Calculate the fitness value of individual grey wolves. Determine the combination parameters $(C, \gamma)$ based on the grey wolf's position, and use the training data to train the SVM model based on the combination of these parameters, and determine the fitness value of the grey wolf individual using the calculated mean square error MSE.

7. Keep the three grey wolves with the best fitness value, namely $\alpha$ wolf, $\beta$ wolf and $\delta$ wolf.

8. With reference to the positions of $\alpha$ wolf, $\beta$ wolf and $\delta$ wolf, use equations (5) to (7) to update the positions of individual grey wolves.

9. Calculate the fitness values of all grey wolves. Determine the combination parameters $(C, \gamma)$ based on the grey wolf's position, and use the training data to train the SVM model based on the combination of these parameters, and determine the fitness value of the grey wolf individual using the calculated mean square error MSE.

10. Update the positions and fitness values of $\alpha$ wolves, $\beta$ wolves and $\delta$ wolves.

11. Update $a$, $A$ and $C$. Use formulas (3) ~ (4) to initialize the coefficients $A$ and $C$, and the convergence factor $a$ here linearly decreases from the beginning $a = 2$ to $a = 0$.

12. Determine whether the maximum number of iterations has been reached. If it is reached, output the best parameter combination $(C, \gamma)$; if it is not reached, return to step (8), and increase the number of iterations by 1.

13. Output the optimal parameter combination $(C, \gamma)$, and retrain the SVM based on the training set.

14. Test the sample to be tested through the SVM-IGWO model, and evaluate the performance of the model's prediction results.

![Diagram](image)

Fig. 1. Construction process of SVM-based model predicting dam deformation

In order to measure the performance of the SVM-IGWO algorithm, three performance indicators, mean square error (MSE), mean absolute percentage error (MAPE) and square correlation coefficient ($R^2$) are used to evaluate the performance of the prediction model [2]. The closer the square correlation coefficient ($R^2$) is to 1, the better the prediction effect of the model; the smaller the mean square error (MSE) and the mean absolute percentage error (MAPE), the better the prediction effect of the model.
4. Case study

4.1 Description of the gravity dam

The data studied in this paper correspond to Xiluodu concrete arch dam located in southwest China [6]. The dam crest elevation is 610.0m, and the maximum dam height is 285.5m. The normal storage level and dead water level of the reservoir are 600.0m and 540.0m, respectively. The reservoir started impoundment on May 4, 2013, and the reservoir first stored water to a normal water level on September 28, 2014. The plumb line system was installed to monitor the horizontal displacement of the dam, as shown in Fig. 2.

The monitoring point PL15-1 of the No. 15 dam section is taken as an example. From June 2014 to June 2018, the radial displacement and water level changes of the measuring points are shown in Fig. 3 and Fig. 4. Take the displacement observation data from June 2014 to December 2015 as the training set, and the remaining time period data as the prediction set. In Fig. 3, the symbol (−) indicates the displacement to the downstream, and the symbol (+) indicates the displacement to the upstream.

![Fig. 2. Vertical arrangement of the dam body](image)

![Fig. 3. Time curve on observed horizontal displacement of PL15-1](image)

![Fig. 4. Time curve on observed water level of PL15-1](image)

4.2 Effective factors

This paper takes the dam deformation factors as the input variables of the model [2]. The dam's influencing factors mainly include water level factor, temperature factor and time-effect factor. Therefore, it is determined that the input variable of the dam deformation prediction model based on the intelligent algorithm is
\[ \begin{align*}
&\{H - H_0, (H - H_0)^2, (H - H_0)^3, (H - H_0)^4, \\
&\sin \frac{2\pi it}{365} - \sin \frac{2\pi t_0}{365}, \cos \frac{2\pi it}{365} - \cos \frac{2\pi t_0}{365}, \\
&\sin \frac{4\pi it}{365} - \sin \frac{4\pi t_0}{365}, \cos \frac{4\pi it}{365} - \cos \frac{4\pi t_0}{365}, \\
&\theta - \theta_0, \ln \theta - \ln \theta_0 \} 
\end{align*} \]

(8)

Where, \( H \) is the current water level. \( H_0 \) is the upstream water level at the reference time. \( t \) is the cumulative number of days between the current monitoring time and the start of the monitoring time. \( t_0 \) is the cumulative number of days between the start time of the data sequence used to construct the prediction model and the current time, \( \theta = t / 100, \theta_0 = t_0 / 100 \). In order to match the consistency of the model and avoid the situation where big data information overwhelms small data information, all source data are normalized in the range of \([0,1] \) \([2]\).

4.3 Parameter settings
The initial parameters of the population are set as follows. The population size is 20, the maximum number of iterations is 200, \( \varepsilon \) is 0.01, the range of penalty factor \( C \) is \([0.01, 100]\), and the range of the kernel parameter \( \gamma \) is \([0.01, 1000]\). The calculation is terminated when the number of iterations reaches 200. The relationship between fitness value and iteration number of the GWO and IGWO algorithms is shown in Fig. 5. For the measured data series of Xiluodu high arch dam, the optimal parameters \([C, \gamma]\) of SVM obtained by the IGWO algorithm are \([0.1088, 24.2515]\).

It can be seen from Fig. 5, the IGWO algorithm reduces the number of iterations and can find the solution closest to the best goal faster. It is mainly because the IGWO algorithm improves the diversity of the initial population and improves the global search capability, thereby accelerating the convergence speed and convergence accuracy.

4.4 Calculation results and performance comparison
The prediction performance of the SVM, SVM-GWO and SVM-IGWO models are shown in Table 1 and Fig. 6-7. From Table 1, we can see that the square correlation coefficient (R\(^2\)) is ranked from large to small as SVM-IGWO model> SVM-GWO model> SVM model, mean absolute percentage error (MAPE) is ranked from small to large SVM-IGWO model< SVM model< SVM-GWO model, and mean square error (MSE) is ranked from small to large SVM-IGWO model< SVM model< SVM-GWO model. The SVM-IGWO model has the largest R\(^2\) of 0.9747, and the smallest MAPE and MSE are 0.0569 and 6.1547, respectively. As can be seen from Fig. 6, SVM, SVM-GWO and SVM-IGWO models can effectively predict the trend of dam displacement, but as time goes by, the deviation of the specific prediction value from the actual observation value becomes larger and larger. The calculation results show: (1) The applicability of the GWO optimized SVM algorithm in dam deformation prediction; (2)
The IGWO algorithm proposed in this paper has more outstanding optimization ability in optimizing SVM parameters than the conventional GWO.

### Table 1. Predictive performance of SVM, SVM-GWO and SVM-IGWO models

| Prediction model | MSE    | MAPE  | R²    |
|------------------|--------|-------|-------|
| SVM              | 33.1295| 0.1148| 0.9517|
| SVM-GWO          | 21.0568| 0.1395| 0.9616|
| SVM-IGWO         | 6.1547 | 0.0569| 0.9747|

![Fig. 6. Time curves on predicted results of SVM, SVM-GWO and SVM-IGWO models](#)

![Fig. 7. Time curves on prediction errors of SVM, SVM-GWO and SVM-IGWO models](#)

### 5. Conclusions

In this paper, the crossover and mutation operators in the difference algorithm are introduced to enrich the initial population of the conventional GWO algorithm, and then combined with SVM, a dam deformation prediction model based on IGWO-SVM is proposed. Based on the measured displacement data of the Xiluodu super high arch dam, the prediction performance of the SVM, GWO-SVM and IGWO-SVM models are compared and analyzed. The main conclusions are as follows:

1. The optimization performance of the IGWO algorithm proposed in this paper is superior to that of the GWO algorithm. Using the DE algorithm to ensure the initial population diversity can effectively improve the grey wolf optimization algorithm's ability to find high-quality solutions. Example calculation results show that the IGWO algorithm has faster convergence speed and higher accuracy.

2. It is verified by the example of Xiluodu super high arch dam that SVM, SVM-GWO and SVM-IGWO models can effectively predict the dam deformation trend, but the prediction accuracy of the SVM-IGWO model is higher and the model is more stable.

### References

[1] Su, H., Li, X., Yang, B. (2018) Wavelet support vector machine-based prediction model of dam deformation. Mechanical Systems and Signal Processing, 110: 412-427.

[2] Li, M., Wang, J. (2019) An empirical comparison of multiple linear regression and artificial neural network for concrete dam deformation modelling. Mathematical Problems in Engineering, 1-13.
[3] Sm, A., Smm, B., Al, A. (2014) Grey wolf optimizer. Advances in Engineering Software, 46–61.
[4] Zhang S., Zhou Y., Li Z. (2016) Grey wolf optimizer for unmanned combat aerial vehicle path planning. Advances in Engineering Software, 99: 121-136.
[5] Sarker A., Elsayed M., Tapabrata R. (2014) Differential evolution with dynamic parameters selection for optimization problems. IEEE Transactions on Evolutionary Computation. 18(5): 689–707.
[6] Wang S., Xu Y., Gu C. (2020) Two spatial association-considered mathematical models for diagnosing the long-term balanced relationship and short-term fluctuation of the deformation behaviour of high concrete arch dams. Structural Health Monitoring, 19(5): 1421-1439.