Article

Intelligent Prediction of Aeroengine Wear Based on the SVR Optimized by GMP SO

Bo Zheng ¹ , Feng Gao ¹, Xin Ma ^{2,*} and Xiaoqiang Zhang ¹

¹ Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China, Guanghan 618307, China; bzheng@cafuc.edu.cn (B.Z.); gaofeng@cafuc.edu.cn (F.G.); xqzhang@cafuc.edu.cn (X.Z.)

² College of Air Traffic Management, Civil Aviation Flight University of China, Guanghan 618307, China

* Correspondence: maxin@cafuc.edu.cn; Tel.: +86-838-518-2516

Abstract: In order to predict aeroengine wear accurately and automatically, as a predictor, support vector regression (SVR) was optimized by means of particle swarm optimization (PSO). The guided mutation strategy of PSO (GMP SO) is presented herein to determine the proper structure parameters of an SVR, as well as the embedding dimensions of the training samples. The guided mutation strategy was able to increase the diversity of particles and improve the probability of finding the global extremum. Furthermore, single-step and multi-step prediction methods were designed to meet different accuracy requirements. A prediction comparison study on spectral analysis data was carried out, and the contrast experiments show that compared with SVR optimized by means of a traditional PSO, a neural network and an auto regressive moving average (ARMA) prediction model, the SVR optimized by means of the GMP SO approach produced prediction results not only with higher accuracy, but also with better consistency.

Keywords: support vector regression; particle swarm optimization; guided mutation; spectral analysis; wear prediction

Citation: Zheng, B.; Gao, F.; Ma, X.; Zhang, X. Intelligent Prediction of Aeroengine Wear Based on the SVR Optimized by GMP SO. *Appl. Sci.* **2021**, *11*, 10592. <https://doi.org/10.3390/app112210592>

Academic Editor: Jordi Cusido

Received: 13 September 2021

Accepted: 9 November 2021

Published: 10 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The aeroengine is the crucial component for guaranteeing flight safety. The bearings and gears, as wearing parts, are widely used in the accessory transmission system, and they easily cause wear faults in the harsh work environment, thus threatening flight safety. Oil spectral analysis is a relatively mature technology in aeroengine condition monitoring and fault diagnosis [1]. By analyzing and predicting the composition and concentration of abrasives, the wear status and the internal fault mechanism of the aeroengine can be judged effectively. Therefore, it is conducive to make maintenance plans in advance, which is important in order to improve the work reliability of aeroengines.

Several previous studies have focused on wear prediction. In 1940, R. Holm, a German scientist, deduced a wear prediction equation under dry sliding friction conditions on the basis of a single wear mechanism and characteristics [2]. British scholars Wallbridge and Dowson defined wear as a stochastic process, determining the wear parameter distributions of some metals based on probability and mathematical statistics, and thus opening up a new research direction in wear prediction [3]. The Grey model GM (1, 1) has also been used to predict the wear trends of engines [4]. With the development of artificial intelligence technology, the accuracy of wear prediction has also been improved. Neural networks have been proposed to predict tool wear, with results showing that their prediction accuracy is better than those of other methods [5]. Furthermore, a novel deep learning-based method was introduced to monitor wear status, which improved the ability to undertake wear data processing [6].

The wear data of aeroengines are usually small sample data, which have the characteristics of nonlinearity, complexity and uncertainty [7]. Thus, support vector regression (SVR) has been proposed as a method of predicting the wear status. As is well known,

SVR is a statistical learning theory built on the Vapnik–Chervonenkis (VC) dimension and structural risk minimum principle. When processing data, the SVR has good generalization capability and avoids the defects of falling into a local minimum, over-learning, and relying on sample size [8]. Furthermore, particle swarm optimization (PSO) has been used to optimize the structure parameters of SVRs and the embedding dimension of training samples, which can ensure the accuracy and consistency of the prediction results.

The rest of the paper is organized as follows. In Section 2, the SVR model and the optimization calculation for relative parameters are introduced. Section 3 presents the improved GMPPO approach and verifies the performance of GMPPO. In Section 4, an SVR optimized by means of GMPPO was used to predict the wear of an aeroengine and achieved better results compared with other algorithms. Finally, in the last section, conclusions are drawn.

2. SVR Prediction Model

SVR is derived from a classification application; the classification decision can be converted into a regression estimation depending on the ε -insensitive loss function. Assume that $x_i \in \mathbf{R}^n$ denotes the i th input vector, $y_i \in \mathbf{R}$ denotes the observation value of x_i , and the ε -insensitive loss function can be defined as follows in Equation (1):

$$L(y_i, f(x_i)) = |y_i - f(x_i)|_\varepsilon = \max\{0, |y_i - f(x_i)| - \varepsilon\} \quad (1)$$

where the insensitive loss coefficient ε is a positive number with minimized empirical risk, and $f(x_i)$ is the predicted value of vector x_i . In the case of $|y_i - f(x_i)| \leq \varepsilon$, the predicted values are free of loss. Furthermore, the regression optimization problem can be transformed as follows in Equation (2):

$$\begin{cases} \min \frac{1}{2} \|w\|^2 + C \frac{1}{l} \sum_{i=1}^l |y_i - f(x_i)|_\varepsilon \\ \begin{cases} |y_i - f(x_i)| - \varepsilon, & |y_i - f(x_i)| \geq \varepsilon \\ 0, & \text{otherwise} \end{cases} \end{cases} \quad (2)$$

where w is the normal vector of the hyperplane; C is the penalty factor; and l is the number of input vectors. The insensitive loss coefficient ε can control the width of the insensitive region of the sample data influenced by the regression function and this can affect the number of support vectors. Furthermore, the penalty factor C reflects the penalty for these sample data exceeding the ε width, which can affect the complexity and stability of the SVR. A smaller C value can lead to a weaker penalty, which makes the training error larger. If C is too large, the learning accuracy can be improved, but the generalization ability of the SVR will become worse.

For SVR prediction, the method used to construct the prediction function is important. In general, using the Lagrange function to construct and solve the convex quadratic programming problem, the prediction functions can be described in Equation (3):

$$f(x) = \sum_{i=1}^n \alpha_i \Phi(x) \Phi(x_i^*) + b \quad (3)$$

where $\Phi(\cdot)$ is a mapping transform from the original space to the high-dimensional feature space, so that the nonlinear regression in the original space can be transformed into the linear regression in high-dimensional space. α_i is the weight connecting the feature space to the output space, $b \in \mathbf{R}$, and b is a bias. x is the unknown input vector, x_i^* is the support vector determined through calculation, and n is the number of support vectors.

Actually, the construction of $\Phi(\cdot)$ is a complex theoretical problem. Moreover, the computation cost of the high-dimensional inner product is very large. In practice, the kernel function $K(x, x_i^*)$, meeting the Mercer condition, can be used to combine the nonlinear mapping and inner-product of two vectors in feature space so that the nonlinear mapping

is carried out implicitly. So $K(x, x_i^*) = \Phi(x) \cdot \Phi(x_i^*)$, which is defined in the original space instead of the high-dimensional feature space. Furthermore, Equation (3) can be written as follows in Equation (4):

$$f(x) = \sum_{i=1}^n \alpha_i K(x \cdot x_i^*) + b \quad (4)$$

The Gauss radial basis function can approach the train samples better when dealing with nonlinear problems [9], which is conducive to data prediction. Therefore, the Gauss radial basis function is taken as a kernel function in Equation (5):

$$K(x, x_i) = \exp(-\gamma \cdot \|x - x_i^*\|^2) \quad (5)$$

where γ is the parameter of the kernel function, and it mainly reflects the correlation among support vectors. If γ is small, the support vectors will be relatively loose, and the generalization ability cannot be guaranteed; if γ is large, it is difficult for the regression model to achieve sufficient accuracy.

Our research demonstrates that the factors affecting the prediction accuracy of the SVR include γ , C , and ε . The kernel parameter γ can reflect the correlation among support vectors. The penalty factor C can adjust the tolerance of error and the proportion of empirical risk. The insensitive loss coefficient ε can determine the number of support vectors and ensure the sparsity of solutions. Undoubtedly, the combination of different (C, γ, ε) can lead to obvious differences in the prediction performance of the SVR. Thus, the optimization calculation for (C, γ, ε) is necessary, which can ensure the accuracy and consistency of the prediction results.

3. Improved PSO Algorithm

PSO is a classic collaborative swarm heuristic search algorithm, which can provide solutions closer to the global optimal area. On the other hand, traditional PSO can easily be trapped in local suboptimal areas due to premature convergence. The reasons for premature convergence are discussed in detail in [10]. In this paper, a simple guided mutation strategy was designed to balance the exploitation and exploration, and guide the particles, jumping local suboptimal areas and increasing the probability of finding the global optimal area.

3.1. Guided Mutation Strategy

Detailed information about traditional PSO can be found in [10,11]. In order to improve the performance of traditional PSO, poor particles that cannot gain a better individual extremum will be guided to mutate in order to increase diversity. It is necessary to maintain the diversity at the different stages of iterations and improve the quality of the population; therefore, this strategy is termed the guided mutation strategy. Figure 1 shows the procedure involved in the guided mutation strategy.

The guided mutation strategy can be described as follows in Equation (6):

$$p^{t+1} = (1 - \beta) \cdot p^t + r \cdot \beta \cdot (p_{ge}^t - p^t) \quad (6)$$

where p denotes a particle, p_{ge} denotes the global extremum, t denotes the current iterative step, r denotes a random number set in $[0, 1]$, and β is a binary vector generated randomly in order to select random positions of particles.

Although the guided mutation strategy takes a simple mathematical form, it can make the population change. These poor particles can be guided to mutation under the influence of p_{ge} and obtain new positions around the p_{ge} . Moreover, p_{ge} is dynamically adjusted through the iterative process, which helps to exploit and explore the optimization space at the early and late iterations, respectively. Meanwhile, the population diversity can be maintained. Thus, the population can approach the optimal area from different dimensions and find the optimal solution with a higher probability.

Figure 1. The procedure of the guided mutation strategy.

3.2. PSO Performance Verification

In order to verify the performance of the improved PSO, some typical test functions can be used to compare the different well-known PSO variants. The typical test functions are listed in Table 1. In Table 1, x_i and d represent an element and the number of elements in each function, respectively.

Table 1. Parameters for radio communication.

Test Function	Search Range	Dimension	Optimal Solution	Optimal Extremum
Sphere: $f_{\text{Sph}} = \sum_{i=1}^d x_i^2$	$[-100, 100]$	50	$(0, 0, \dots, 0)50$	0
Schaffer: $f_{\text{Sch}} = 0.5 + \frac{\left(\sin \sqrt{\sum_{i=1}^d x_i^2} - 0.5\right)^2}{\left[1 + 0.001 \left(\sum_{i=1}^d x_i^2\right)\right]^2}$	$[-100, 100]$	50	$(0, 0, \dots, 0)50$	0
Griewank: $f_{\text{Gri}} = \sum_{i=1}^d \frac{x_i^2}{4000} - \prod_{i=1}^d \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	$[-100, 100]$	50	$(0, 0, \dots, 0)50$	0
Ackley: $f_{\text{Ack}} = -20 \cdot e^{-0.2 \sqrt{\frac{1}{d} \sum_{i=1}^d x_i^2}} - e^{\frac{1}{d} \sum_{i=1}^d \cos(2\pi x_i)} + 20 + e$	$[-100, 100]$	50	$(0, 0, \dots, 0)50$	0
Rastrigin: $f_{\text{Ras}} = \sum_{i=1}^d [x_i^2 - 10 \cdot \cos(2\pi x_i) + 10]$	$[-100, 100]$	50	$(0, 0, \dots, 0)50$	0
Rosenbrock: $f_{\text{Ros}} = \sum_{i=1}^{d-1} [100(x_i^2 - x_{i+1})^2 + (x_i - 1)^2]$	$[-100, 100]$	50	$(1, 1, \dots, 1)50$	0
SDPF: $f_{\text{SDPF}} = \sum_{i=1}^d x_i ^{i+1}$	$[-100, 100]$	50	$(0, 0, \dots, 0)50$	0
RHEF: $f_{\text{RHEF}} = \sum_{i=1}^d \left(\sum_{j=1}^i x_j\right)^2$	$[-100, 100]$	50	$(0, 0, \dots, 0)50$	0

The follow PSO variants were improved based on different strategies, so they were used for the comparison discussed in this paper. PSO variants include SRPSO (self-regulating PSO) [12], ALCPSO (aging leader and challengers PSO) [13], SLPSO (social learning PSO) [14], IWPSO (inertia weights PSO) [11], SFPSO (shrinkage factor PSO) [15], DNPSO (dynamic neighborhood PSO) [16], SAPSO (simulated annealing PSO) [17], and

MAPSO (multiple agents PSO) [18]. Furthermore, the population size was 60, and the total iteration number was 400, and all the algorithms were operated in the same computational environment. Furthermore, the calculations of each PSO variant were repeated 100 times to obtain the mean of the fitness values. Table 2 shows the mean changes, which can prove the global optimal performance of PSO variants. Table 2 proves the advantage of GMPSO (guided mutation strategy PSO). Obviously, it was able to solve all of the test functions in Table 1, and its better global optimization performance is conducive to engineering applications.

Table 2. Comparison of the optimization results of PSO variants.

Function	SRPSO	ALCPSO	SLPSO	IWPSO	SFPSO	DNPSO	SAPSO	MAPSO	GMPSO
f_{Sph}	0.00025	8.4617	4174.53	474.278	1956.71	0.1176	132.214	0.0042	3.98×10^{-68}
f_{Sch}	0.4481	0.4548	0.4928	0.4451	0.4886	0.4995	0.4455	0.3816	6.67×10^{-11}
f_{Gri}	0.0055	0.3550	2.0718	1.1099	1.4751	0.0105	0.9869	0.0041	0
f_{Ack}	21.1626	19.2621	18.6374	19.7884	20.6655	20.0057	19.4486	18.8497	0
f_{Ras}	193.353	950.846	4564.84	1124.34	2613.14	471.213	1176.23	204.646	0
f_{Ros}	155.083	5219.32	8819.17	5891.12	2059.34	168.554	2634.42	148.853	47.3102
f_{SDPF}	3.15×10^8	1.22×10^{24}	6.12×10^{52}	1.48×10^{38}	1.43×10^{39}	0.1566	9.66×10^{35}	7.84×10^7	1.77×10^{-101}
f_{RHEF}	917.416	1033.42	9612.11	10436.4	6855.24	675.486	2506.73	1258.76	4.37×10^{-56}

4. Aeroengine Wear Prediction Case

The wear situation of a certain civil aviation engine can be reflected by the Fe element in the oil, and the content of the Fe element in the oil can be obtained periodically through spectroscopic analysis. Figure 2 demonstrates the change in the Fe content with the increasing of the flight cycle. There are total 88 data; in this paper, the last five data were selected as the test data to verify the optimized SVR.

Figure 2. The time sequence of Fe in the oil.

To improve the prediction accuracy and convergence speed, reducing the difference between the data, the data in Figure 2 can be scaled using Equation (7):

$$s_i = (1 + \exp(-q_i/q_1)) \quad (7)$$

where q_i denotes the i th original data in the Fe data series, and s_i denotes the i th scaled data. Assume that p_i denotes the i th prediction data, which can be restored using Equation (8):

$$q_i = -\lg(-1/p_i - 1) \cdot q_1 \quad (8)$$

The fitness function outputs the measurement value of the PSO's optimization effect. In order to improve the prediction performance, k -fold cross validation (k -CV) is used to design the fitness function. Generally, k -CV is used to determine the model parameters, so that the generalization performance of the model is better. The principle of k -CV is that the original data are divided into k groups, each group is taken as the test data successively, and then the rest of the groups are taken as the training data, so there are k test results, and the mean of k results will be the output of the fitness function. For example, a known data set is divided into k groups. First, group 1 is selected as the test data, group 2 to k are used to train the algorithm. The trained algorithm will be tested using group 1 and output a test result for group 1. Then, group 2 is selected as the test data and groups 1 and 3 to k are used to train the algorithm, then output a test result for group 2. Successively, we can obtain k test results; the mean of k test results is the output of the fitness function. the suitable parameters can be found through cross training. Obviously, the advantage of the fitness function based on k -CV is the maintenance of the ergodicity of the data due to non-repeated sampling.

On the other hand, [19] discusses the organizational form of time series data; therefore, we will organize the time series data according to [19]. Furthermore, the embedding dimension D of the training samples is a key parameter in organizing the data. Every D of continuous data can form a training sample and the value of D has an obvious influence on the prediction performance. Meanwhile, as previously mentioned, the combination of different (C, γ, ε) parameters has a marked impact on the prediction performance of the SVR. In order to obtain good prediction results, these parameters must be optimized so that the embedding dimension D of the time series data and the parameters (C, γ, ε) of the SVR can match well. Accordingly, the parameters $(C, \gamma, \varepsilon, D)$ should be optimized by means of PSO. The position vector of each particle is defined as $(C, \gamma, \varepsilon, D)$, which represents a feasible solution.

There are two forms of prediction, including single-step and multi-step prediction. Single-step prediction means that the trained SVR only predicts one data every time, then the model is retrained for the next prediction. Multi-step prediction means that the trained SVR can predict multiple data each time without being retrained.

4.1. Multi-Step Prediction

In this study, the initial space of the particle position vector was set to $C \in [0.1, 100]$, $\gamma \in [0.01, 1000]$, $\varepsilon \in [0.01, 100]$, and $D \in [4, 20]$. Other parameters were set as follows: the population size was 40, and the total iteration number was 400, k -CV was 10. The last five wear data were used as the testing data, so we were able to obtain five testing samples. Multi-step prediction was used to verify the performance of the proposed method.

After the iterative operation of PSO, the optimization result outputted by GMPSO was $(C, \gamma, \varepsilon, D)_{\text{best}} = (0.1, 13.5812, 0.001, 13)$. This means the wear time serial can be organized for training samples with an optimal embedding dimension $D = 13$, so we can obtain 70 trained samples, then the SVR with the optimal parameters $(C, \gamma, \varepsilon)_{\text{best}} = (0.1, 13.5812, 0.001)$ can be trained using the 70 trained samples. Finally, we can obtain 38 support vectors—the support vector number $n = 38$, the weight vector $\alpha = [0.1000, \dots, -0.0574, \dots, -0.0228]_{38 \times 1}$, and the bias $b = 0.8594$. Therefore, the prediction function can be constructed as follows in Equation (9):

$$f(x) = \sum_{i=1}^{38} \alpha_i \exp\left(13.5887 \cdot \|x - x_i^*\|^2\right) + 0.8594 \quad (9)$$

where x is the samples to be predicted, and x_i is the i th support vector. Thus, the five testing samples can be brought into Equation (9). Meanwhile, the influence of randomness on the PSO is discussed in details in [20]. GMPSO and traditional PSO were operated several times, and we have verified the influence of the optimization results on the prediction

accuracy. Table 3 shows the comparison of the multi-step prediction results of any two operations.

Table 3. Comparison of multi-step prediction results.

True Values	GMPSO-SVR				PSO-SVR			
	Exp.No1 Prediction	RE	Exp.No2 Prediction	RE	Exp.No1 Prediction	RE	Exp.No2 Prediction	RE
5.98	5.9842	0.07%	5.9842	0.07%	6.1838	3.41%	6.1387	2.65%
6.02	6.0338	0.23%	6.0338	0.23%	5.9301	1.49%	6.0604	0.67%
5.75	5.9962	4.28%	5.9962	4.28%	5.7172	0.57%	5.9905	4.18%
6.23	5.8659	5.85%	5.8659	5.85%	5.6559	9.21%	5.8062	6.8%
5.88	5.7425	2.34%	5.7425	2.34%	5.4062	8.06%	5.7914	0.15%
$(C, \gamma, \varepsilon, D)_{\text{best}}$	(0.1, 13.5887, 0.01, 13)		(0.1, 13.5887, 0.01, 13)		(100, 0.01, 0.01, 11)		(0.1, 11.8511, 0.01, 11)	
Fitness value	0.049%		0.049%		0.053%		0.051%	
Average RE	2.55%		2.55%		4.55%		2.89%	

In Table 3, RE denotes the parameter of relative errors. Table 3 proves that the optimization performance has an obvious impact on the prediction accuracy, SVR optimized by GMPSO shows higher accuracy and a more consistent output, which can be used in the application of aeroengine wear prediction. Figure 3 shows the influence of D on the average relative errors. The average relative error is still obtained through several calculations under the identical conditions. When $D = 13$, the average relative error is the lowest. Figure 3 thus demonstrates the necessity of optimizing the embedding dimension and the optimization capability of GMPSO.

Figure 3. The influence of D on prediction accuracy.

4.2. Single-Step Prediction

The parameter setting process of GMPSO is the same as that described in Section 4.1. Nevertheless, according to the definition of single-step prediction, the SVR will be trained five times to calculate five test samples. Table 4 shows the comparison between the multi-step and single-step predictions. The accuracy of single-step prediction is higher than that of multi-step prediction, but the elapsed time of multi-step prediction is better than that of single-step prediction. Thus, the proper prediction form can be determined by considering the performance and efficiency comprehensively.

Table 4. Comparison between multi-step and single-step predictions.

True Values	Multi-Step GMP SO-SVR		$(C, \gamma, \varepsilon, D)_{\text{best}}$	Single-Step GMP SO-SVR		
	Prediction	RE		Fitness Value	Prediction	RE
5.98	5.9842	0.07%	(0.1, 13.5887, 0.01, 13)	0.049%	5.9842	0.07%
6.02	6.0338	0.23%	(0.1, 11.6259, 0.01, 13)	0.052%	6.0213	0.022%
5.75	5.9962	4.28%	(0.1, 4.9941, 0.01, 16)	0.056%	5.7554	0.094%
6.23	5.8659	5.85%	(0.1, 54.7909, 0.01, 26)	0.0489%	6.2236	0.103%
5.88	5.7425	2.34%	(0.1, 16.4312, 0.01, 3)	0.0442%	5.8819	0.032%
Average RE	2.55%			0.064%		
Elapsed time	70.32 s			358.35 s		

4.3. Comparison with Other Methods

In order to verify the accuracy and consistency of the GMSO-SVR prediction method proposed in this paper, it was compared with the commonly used back propagation (BP) network [21] and autoregressive moving average (ARMA) [22]. Because the weights of the BP network are initialized randomly, the BP network performs calculations several times, and the best result was used for the comparison. Specifically, based on the data structure, BP network is a three-layer network with 13 neurons in the input layer and one neuron in the output layer. As for the hidden layer, it can be calculated using an empirical formula $\sqrt{1 + 13} + \eta$, in which 1 and 13 represent the number of neurons in the output and input layers, respectively, and η is a positive integer generated randomly between 1 to 10. The error goal for BP is set as 0.001 and the iteration number is 500. ARMA is a classical method for processing time series; after calculation, the order of the model is (4, 8), then the autoregressive coefficient is calculated as $[0.5909 \ 0.5777 \ 0.3926 \ 0.5616]_{1 \times 4}$, and the moving average coefficient is calculated as $[0.2464, 0.4266, \dots, -0.7759, -0.4764]_{1 \times 8}$. Table 5 shows the comparison of the different prediction results. Table 5 demonstrates that the SVR optimized by means of GMP SO possessed more stable performance and better prediction accuracy.

Table 5. Comparison of different prediction results.

True Values	Multi-Step GMP SO-SVR		Single-Step GMP SO-SVR		BP Network		ARMA	
	Prediction	RE	Prediction	RE	Prediction	RE	Prediction	RE
5.98	5.9842	0.07%	5.9842	0.07%	5.6074	6.23%	5.5091	7.87%
6.02	6.0338	0.23%	6.0213	0.022%	6.3361	5.25%	5.5980	7.01%
5.75	5.9962	4.28%	5.7554	0.094%	5.8454	1.66%	6.2631	8.92%
6.23	5.8659	5.85%	6.2236	0.103%	5.9702	4.17%	4.7603	23.59%
5.88	5.7425	2.34%	5.8819	0.032%	5.8382	0.71%	3.8466	34.58%
Average RE	2.55%		0.064%		3.60%		16.39%	

5. Conclusions

PSO is designed to optimize the structure parameters of SVRs and the embedding dimensions of training samples. Based on contrast experiments, the following conclusions can be drawn:

- (1) The guided mutation strategy can increase the diversity of particles and improve the global optimal performance. Therefore, GMP SO is basically better than other PSO variants.
- (2) The structure parameters and the embedding dimension have an obvious impact on the prediction accuracy; it is necessary to optimize them to improve the prediction accuracy of SVR.

- (3) Multi-step prediction has relatively lower accuracy than that of single-step prediction due to the error cumulative effect, but the time costs of the two forms exhibit the opposite trend.

The SVR optimized using GMPSO had better accuracy and higher consistency, and can thus provide reliable predictions of aeroengine wear and ensure the safe running of aeroengines.

Author Contributions: Conceptualization, B.Z. and F.G.; methodology, B.Z. and X.M.; software, X.Z.; validation, B.Z., F.G. and X.Z.; writing—original draft preparation, F.G.; writing—review and editing, B.Z.; visualization, X.Z.; project administration, B.Z.; funding acquisition, B.Z., X.Z. and X.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded using funds from the Project of Sichuan Province Science and Technology Program, grant number 2021YJ0519; China Civil Aviation Administration Development Foundation Educational Talents Program, grant number 14002600100018J034; General Foundation of Civil Aviation Flight University of China, grant number 2019-53.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Li, H.; Diaz, H.; Guedes Soares, C. A Failure Analysis of Floating Offshore Wind Turbines using AHP-FMEA Methodology. *Ocean Eng.* **2021**, *234*, 109261. [\[CrossRef\]](#)
- Schmidt, W.; Groche, P. Wear prediction for oscillating gear forming processes using numerical methods. *Key Eng. Mater.* **2018**, *767*, 283–289. [\[CrossRef\]](#)
- Zheng, B.; Ma, X. Application of improved Kohonen network in aeroengine classification fault diagnosis. *Aeroengine* **2020**, *46*, 23–29.
- Li, Y.; Chen, Y.; Zhong, J.; Huang, Z. Niching particle swarm optimization with equilibrium factor for multi-modal optimization. *Inf. Sci.* **2019**, *494*, 233–246. [\[CrossRef\]](#)
- Huang, C.G.; Huang, H.Z.; Li, Y.F. A bi-directional LSTM prognostics method under multiple operational conditions. *IEEE Trans. Ind. Electron.* **2019**, *66*, 8792–8802. [\[CrossRef\]](#)
- Lei, Y.G.; Jia, F.; Zhou, X.; Jing, L. A deep learning-based method for machinery health monitoring with big data. *J. Mech. Eng.* **2015**, *51*, 49–56. [\[CrossRef\]](#)
- Li, H.; Guedes Soares, C.; Huang, H.Z. Reliability analysis of a floating offshore wind turbine using Bayesian networks. *Ocean Eng.* **2020**, *217*, 107827. [\[CrossRef\]](#)
- Hu, J.; Zhang, H. Support vector machine method for developing ground motion models for earthquakes in western part of China. *J. Earthq. Eng.* **2021**, *2021*, 3086. [\[CrossRef\]](#)
- Zheng, B.; Gao, H.Y.; Ma, X.; Zhang, X. Graph Partition Based on Dimensionless Similarity and Its Application to Fault Diagnosis. *IEEE Access* **2021**, *9*, 35573–35583. [\[CrossRef\]](#)
- Li, H.; Diaz, H.; Guedes Soares, C. A developed failure mode and effect analysis for floating offshore wind turbine support structures. *Renew. Energy* **2021**, *164*, 133–145. [\[CrossRef\]](#)
- Shailaja, K.; Anuradha, B. Improved face recognition using a modified PSO based self-weighted linear collaborative discriminant regression classification. *J. Eng. Appl. Sci.* **2017**, *12*, 7234–7241.
- Tanweer, M.R.; Suresh, S.; Sundararajan, N. Self regulating particle swarm optimization algorithm. *Inf. Sci.* **2015**, *294*, 182–202.
- Singh, R.P.; Mukherjee, V.; Ghoshal, S.P. Particle swarm optimization with an aging leader and challengers algorithm for optimal power flow problem with FACTS devices. *Electr. Power Energy Syst.* **2015**, *64*, 1185–1196. [\[CrossRef\]](#)
- Cheng, R.; Jin, Y.C. A social learning particle swarm optimization algorithm for scalable optimization. *Inf. Sci.* **2015**, *291*, 43–60. [\[CrossRef\]](#)
- Zeng, B.; Liu, S.; Cuevas, C. A self-adaptive intelligence gray prediction model with the optimal fractional order accumulating operator and its application. *Math. Methods Appl. Sci.* **2017**, *40*, 7843–7857. [\[CrossRef\]](#)
- Zheng, B.; Gao, F. Fault diagnosis method based on S-PSO classification algorithm. *Acta Aeronaut. Astronaut. Sin.* **2015**, *36*, 3640–3651.
- Zeng, R.; Wang, Y.Y. A chaotic simulated annealing and particle swarm improved artificial immune algorithm for flexible job shop scheduling problem. *EURASIP J. Wirel. Commun. Netw.* **2018**, *2018*, 101–123. [\[CrossRef\]](#)

18. Kalidindi, K.R.; Gottumukkala, P.S.V.; Davuluri, R. Derivative-based band clustering and multi-agent PSO optimization for optimal band selection of hyper-spectral images. *J. Supercomput.* **2020**, *76*, 5873–5898. [[CrossRef](#)]
19. Li, H.; Diaz, H.; Guedes Soares, C. A two-stage failure mode and effect analysis of offshore wind turbines. *Renew. Energy* **2020**, *162*, 1438–1461. [[CrossRef](#)]
20. Ziani, R.; Felkaoui, A.; Zegadi, R. Bearing fault diagnosis using multiclass support vector machines with binary particle swarm optimization and regularized Fisher's criterion. *J. Intell. Manuf.* **2017**, *28*, 405–417. [[CrossRef](#)]
21. Khorram, B.; Yazdi, M. A new optimized thresholding method using ant colony algorithm for MR brain image segmentation. *J. Digit. Imaging* **2019**, *32*, 162–174. [[CrossRef](#)] [[PubMed](#)]
22. Ugalde, A.; Egozcue, J.J.; Ranero, C.R. A new autoregressive moving average modeling of H/V spectral ratios to estimate the ground resonance frequency. *Eng. Geol.* **2021**, *280*, 105957. [[CrossRef](#)]