machines

MDPI

*Article*

# Tool Condition Monitoring Model Based on DAE–SVR

**Xiaoning Sun [1], Zhifeng Yang [2], Maojin Xia [3], Min Xia [4], Changfu Liu [1,*], Yang Zhou [1] and Yuquan Guo [1]**

[1] School of Mechanical Engineering, LiaoNing Petrochemical University, Fushun 113001, China; sunxn817@163.com (X.S.); zhouyang970515@163.com (Y.Z.); guoyq1208@163.com (Y.G.)
[2] School of Mechanical Engineering, Shenyang Polytechnic College, Shenyang 110045, China; yzfjb@126.com
[3] China Petroleum Engineering & Construction Corp. Beijing Branch, Beijing 100101, China; xiamaojin@cnpc.com.cn
[4] School of Engineering, Lancaster University, Lancaster LA1 4YW, UK; xia3@lancaster.ac.uk
* Correspondence: lnpu_lcf1@163.com

**Abstract:** Cutting tools are executive components in metal processing, and tool wear directly affects the quality of the workpiece and processing efficiency; monitoring the change in its state is crucial to avoid accidents and ensure the safety of workers. The traditional monitoring model cannot compress a large amount of cutting data effectively, failing to obtain reliable feature data, and there are some defects in generalization ability and monitoring accuracy. For this purpose, this article takes milling cutters as the research object, and it integrates signals from force sensors, vibration sensors, and acoustic emission sensors, combining the advantages of the denoising autoencoder (DAE) model in data compression and the high monitoring accuracy of the support vector regression (SVR) model, to establish a tool wear monitoring model based on DAE–SVR. The results show that compared with traditional DAE and SVR models in multiple datasets, the maximum improvement in monitoring performance (MAE) is 43.58%.

**Keywords:** multi-sensor signal fusion; monitoring of tool wear status; DAE–SVR; monitoring accuracy

Academic Editor: Antonio J. Marques Cardoso

Received: 31 December 2024
Revised: 24 January 2025
Accepted: 30 January 2025
Published: 1 February 2025

**Citation:** Sun, X.; Yang, Z.; Xia, M.; Xia, M.; Liu, C.; Zhou, Y.; Guo, Y. Tool Condition Monitoring Model Based on DAE–SVR. *Machines* **2025**, *13*, 115. https://doi.org/10.3390/machines13020115



## 1. Introduce

Since the 21st century, global industry has been experiencing a wave of reform centered around intelligent manufacturing. In the context of the new era, CNC machine tools are accelerating towards the goal of intelligence. Cutting tools are one of the key components of CNC machine tools, and they are the most vulnerable and wasteful components. Intelligent monitoring technology for tool wear status is a modern intelligent diagnostic technology that has developed with high-tech advancements, such as computers and intelligent manufacturing. Therefore, the research on tool wear monitoring technology will promote the progress and development of the intelligent manufacturing industry [1]. Tool wear monitoring technology can monitor the usage of tools and predict their service life. Maximizing the utilization of cutting tools is significant for improving machine tool processing efficiency and reducing processing costs. Research has shown that the service life of cutting tools is usually only 50% to 80% of their effective life; the downtime of machine tools caused by tool failure accounts for 7% to 20% of the total downtime. According to the statistics, a reasonable and strategic tool replacement can effectively reduce downtime by 75%; in terms of production efficiency, it can increase by 10% to 60% and at the same time, it can save 10% to 40% of production costs [2–4].

The methods for monitoring tool wear status can be divided into two types: direct monitoring and indirect monitoring. The direct monitoring method includes the optical image method [5], radioactive element method [6], resistance measurement method [7], and contact method [8], etc. These methods can directly capture the geometric changes in worn tools using microscopes and charge-coupled device cameras [9,10] after the machine tool is stopped, thus having high measurement accuracy. However, the direct monitoring method is limited by environmental factors such as cutting fluid, chips, and light in practical applications [11], resulting in high costs, long time consumption, and significantly reduced machining efficiency [12]. Oguamanam, D. et al. [13] and Lanzetta, M. et al. [14] studied tool wear using optical imaging. Mustafa, K. et al. [15] pointed out that cutting force can also affect resistance changes, leading to significant uncertainty in measurement results and making it difficult to apply in practice. The indirect monitoring method establishes a model by collecting signals during the machining process, such as cutting force, vibration, and acoustic emission signals, to infer the health status of the cutting tool. In practical machining applications, indirect methods are simpler and more feasible than direct methods, and can monitor tool wear in real time without stopping the machine [16,17]. Cutting force signals are more susceptible to tool wear than other signals, according to Sobron Lubis et al. [18], who also discovered that tool wear can alter cutting force. Duan et al. [19] preprocessed the vibration signals in the tool dataset using a short-time Fourier transform and proposed an optimized model based on residual network (ResNet) for feature map layer-by-layer dimensionality reduction. This model has the best accuracy in classification tasks. Tool degradation estimation and slotting tool state identification may be accomplished with high accuracy using the monitoring and degradation estimation approach for tool state categorization and logistic regression based on acoustic emission data that Liu et al. [20] presented. Each type of signal acquisition has its advantages and disadvantages. Nowadays, acoustic emission signals, vibration signals, and cutting force signals are frequently employed in tool wear monitoring.

A single-sensor signal may have varying defects, resulting in less-than-ideal monitoring accuracy. Consequently, the technology of multi-sensor signal fusion has been suggested. By analyzing multi-channel data, Dimla, D. E. et al. [21] discovered that the measurement findings are more accurate than those from a single sensor. Multi-signal fusion technology can effectively improve the accuracy of tool wear monitoring. Therefore, this work investigates the multi-sensor signal fusion-based feature extraction of milling machining signals, extracting the frequency–domain and time–frequency characteristics of vibration, force, and acoustic emission signals, respectively.

The cutting process of machine tools generates a large amount of data, and traditional models cannot compress these data more effectively, resulting in the inability to obtain reliable feature data, low robustness to damaged data, and low monitoring accuracy. A tool condition monitoring model based on DAE–SVR was established to address the above issues, combining the advantages of the DAE model in data compression and the high monitoring accuracy of the SVR model. This article studies the following content:

(1) The background and significance of tool wear research and the methods for monitoring tool wear status, including direct and indirect methods, were introduced. The problems of traditional models were analyzed, and a condition monitoring model of tool wear based on DAE–SVR was proposed to address the corresponding problems;
(2) The DAE–SVR mathematical model was established, which explained the principles of the denoising autoencoder and support vector regression, and illustrated the monitoring process of the DAE–SVR model;
(3) The experiment's required equipment and signal processing and the superiority of the proposed DAE–SVR model was demonstrated through experiments comparing it with traditional DAE models, SVR models, CNN, and random forest models.

## 2. Mathematical Model

### *2.1. Denoising Autoencoder (DAE)*

On the basis of autoencoders (AEs), denoising autoencoders (DAEs) are developed to effectively compress data and obtain reliable feature data, remove specific features, and have robustness to damaged data. The training objective shifts from simple reconstruction to removing artificially damaged inputs, that is, to reconstruct clean inputs from damaged versions. When learning is insufficient, the autoencoder needs to capture the input distribution structure to counteract the impact of the damage and reconstruct high-density nearby points. The DAE uses the damaged data to reconstruct the original data after adding artificial damage (such as Gaussian noise) to the input data. The reconstruction error function of DAE is [22]:

$$J_{DAE} = \frac{1}{m}\sum_{i=1}^{m}\left(\frac{1}{2}\left\|h_{W,b}\left(\hat{x}^{(i)} - x^{(i)}\right)\right\|^2\right) + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_l+1}\left(W_{ji}^{(l)}\right)^2 \quad (1)$$

The main function of DAEs is to restore the original input without adding interference vectors. If the DAE can effectively restore the original data under interference, such as Gaussian white noise, it indicates that the deep network of DAEs is robust to the input data.

The process can be expressed as:

$$\hat{x} \sim q_D(\hat{x}|x)$$

$$h_1 = \sigma_e(W_1\hat{x} + b_1)$$

$$y = \sigma_d(W_2 h_1 + b_3)$$

The loss function is expressed as [22]:

$$J_{DAE}(W) = \sum\nolimits_{\hat{x}\sim q_D(\hat{x}|x)}[L(x,y)] \quad (2)$$

In the formula, $q_D(\hat{x}|x)$ adds noise, $\hat{x}$ is the new input vector after adding white noise, and $W_1$, $b_1$, $W_2$, $b_3$ are the weights and biases of the encoder and decoder.

### *2.2. Support Vector Regression (SVR)*

SVM mainly studies binary classification problems, and SVR is the embodiment of SVM in function regression. Similarly to the maximum interval method, a hard $\bar{\varepsilon}$- band hyperplane can be established to solve local minima. Corresponding to the regression problem on the n-dimensional space of SVM, the classification problem of SVR is on the n + 1-dimensional space, and the hyperplane can be represented as $(w \cdot x) + \eta y + b = 0$, and the normal vector can be represented as $(w^T, \eta)^T$.

In the above equation, w is an n−dimensional vector, and η is a real number, corresponding to x and y. The convex optimization problem at this point can be expressed as follows [23]:

$$\begin{cases} min_{w,\eta,b} & \frac{1}{2}\|w\|^2 + \eta^2 \\ s.t. & (w \cdot x_i) + \eta(y_i \cdot \vec{\varepsilon})) + b \geq 1, i = 1, \cdots, l \\ & (w \cdot x_i) + \eta(y_i \cdot \vec{\varepsilon})) + b \geq -1, i = 1, \cdots, l \end{cases} \quad (3)$$

Solving the equation yields $(\overline{w}, \overline{\eta}, \overline{b})$, the differentiation hyperplane is as follows:

$$(\overline{w} \cdot x) + \overline{\eta} y + \overline{b} = 0 \quad (4)$$

After sorting, the regression function is obtained as follows:

$$y = (w^* \cdot x) + b^* \tag{5}$$

where in $w^* = -\frac{\bar{w}}{\bar{\eta}},\ b^* = -\frac{\bar{b}}{\bar{\eta}}$

The optimization problem is as follows [23]:

$$\begin{cases} min_{w,\mathrm{b}} & \frac{1}{2}\|w\|^2 \\ s.t. & (w \cdot x_i) + b - y_i \le \varepsilon, i = 1, \cdots, l \\ & y_i - (w \cdot x_i) - b \le \varepsilon, i = 1, \cdots, l \end{cases} \tag{6}$$

*2.3. Tool Wear Monitoring Based on DAE–SVR*

This article proposes a combined model based on Denoising Autoencoder–Support Vector Regression (DAE–SVR), which combines the superior feature extraction performance of denoising autoencoders with the superior generalization ability and higher prediction accuracy of support vector regression to achieve the monitoring of tool wear status. The model monitoring process is shown in Figure 1, and the monitoring steps of the model are as follows:

(a) Prepare tool wear data. The PHM2010 challenge dataset was chosen as the experimental data for the entire model, which has rich information on tool wear and provides abundant data support for subsequent feature extraction and model training;
(b) The DAE extracts features. Each group is set to 7000 sample data, which are sequentially input into the DAE layer. Choosing a quantity of 7000 samples can fully contain the feature information in the data without wasting computing resources and excessively prolonging model training time due to the large amount of data;
(c) Create labeled data. Use the wear values of each cutting step to label the features extracted by the DAE. The wear value of each cutting step is a key indicator reflecting the tool wear status. Using it as a marker can closely link the extracted features with the actual tool wear status, providing accurate label information for subsequent SVR model training;
(d) Import into the SVR model and output the results. These named labeled data serve as inputs for the SVR model, providing the necessary data foundation for model training and prediction. The trained SVR model processes the input-labeled data and outputs the final prediction result.

Due to the complex processing environment, it is inevitable to encounter some outliers. Using the median filtering method for processing, abnormal values are replaced with normal values, and other normal signal data are not affected. Input a concatenated cell array with a size of 630 × 1 into the DAE feature extraction layer, label the feature data after feature extraction, and the label data are the corresponding wear value after each cutting.

In this paper, the DAE feature extraction layer consists of three denoising autoencoders. The number of encoder iterations is set to 800, 200, and 100, respectively, taking into account loss reduction and the cost of computing resources and time. The encoder transfer function is “satlin”, which is more stable and works well with subsequent links. The L2 weight regularization coefficient is 0.001, which shows good generalization ability. Sparse regularization is 4 and the expected ratio is 0.5, which enables the model to better extract data features. The training loss function is the mean square error of L2 regularization and sparse regularization terms. The losses during training are shown in Figure 2.

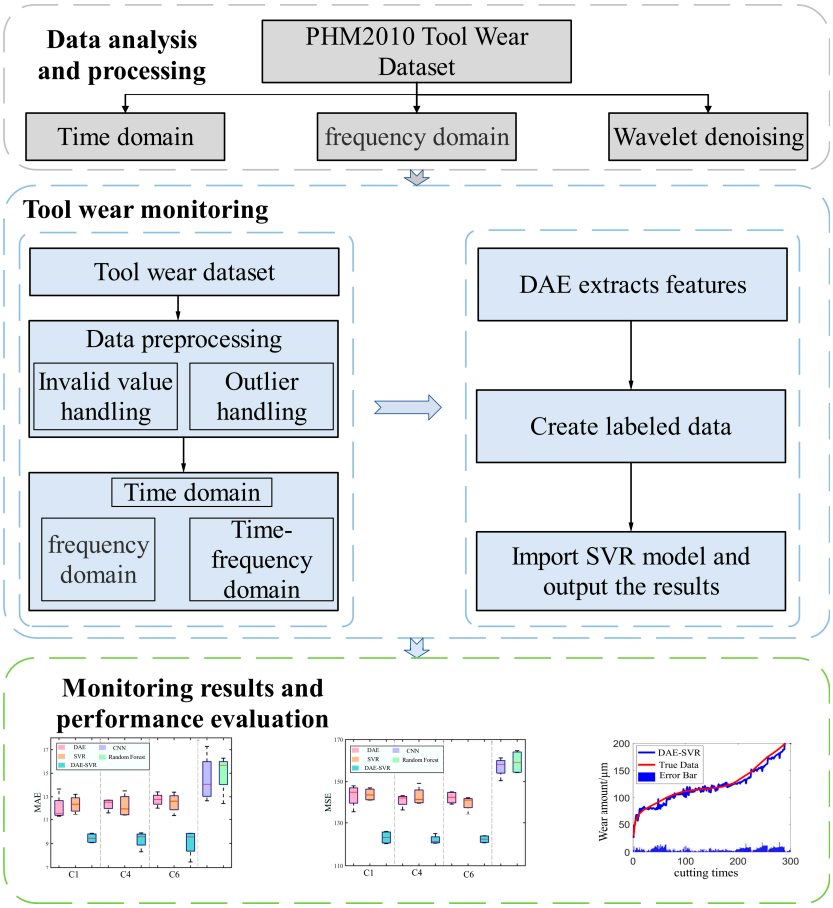


**Figure 1.** Framework of the DAE–SVR model.

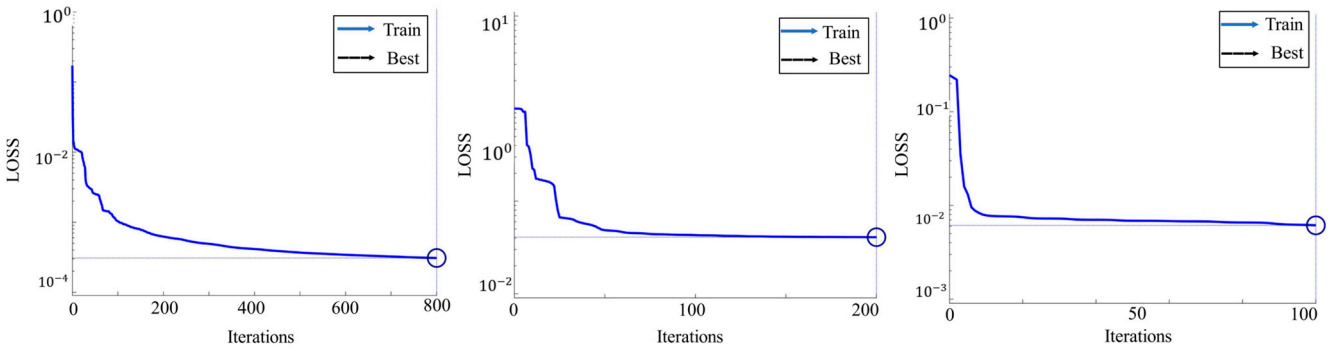


**Figure 2.** Training process loss.

Due to the large number of samples in the dataset, a set of 7000 data was input into the DAE feature extraction layer, resulting in a feature data size of 630 * 9000. Label the extracted feature data with the tool wear value for each cutting operation. Input the labeled data into the SVR model and output the results.

## 3. Experiment and Result Analysis

### *3.1. Experimental Equipment and Parameter Settings*

In order to objectively evaluate the performance of the proposed model, this study used the PHM2010 challenge datasets as experimental data. The datasets are c1, c2, c3, c4, c5, and c6. The training sets are c1, c4, and c6, while the test sets are c2, c3, and c5. With a total of 315 tool runs, each dataset relates to a tool life test. Every run measures the associated tool wear values by gathering force, vibration, and sound emission signals from the

X, Y, and Z axes. Figure 3 displays the structure of the experimental system. Tables 1 and 2 display the primary equipment and processing parameters.

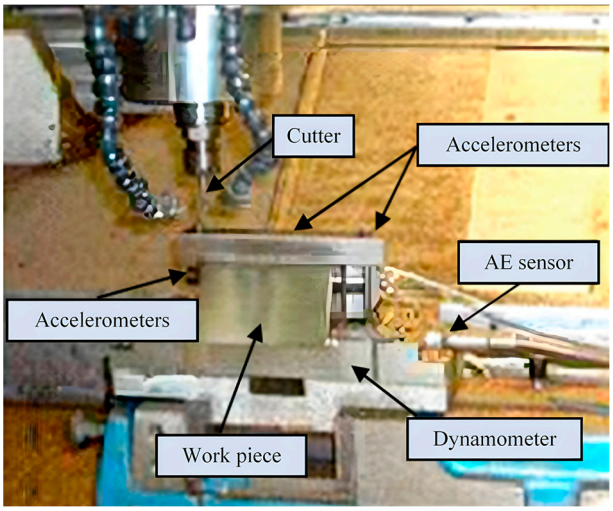


**Figure 3.** Experimental platform settings.

**Table 1.** Main experimental equipment.

| Hardware Conditions | Model and Main Parameters |
|---|---|
| CNC | Roders Tech RFM760 |
| Workpiece material | Inconel 718 |
| Tool | Ball-nose carbide tool |
| Data acquisition card | NI DAQ PCI 1200 |
| Wear measuring device | LEICA MZ12 microscope |

**Table 2.** Milling parameters.

| Spindle Speed | Feed Speed | Radia Cutting Depth | Axial Cutting Depth | Milling Method |
|---|---|---|---|---|
| 10,400 rpm | 1555 mm/min | 0.125 mm | 0.2 mm | Climb milling |

### *3.2. Signal Preprocessing*

Due to the high sampling frequency of the dataset in this article, a large amount of information about tool wear is obtained. The large amount of data affects the training speed and contains a lot of interference information, so it is necessary to process the signal and extract features. Time–domain signals can effectively obtain characteristic parameters reflecting the operating status of mechanical equipment, and time–domain features can better express the tool wear status. The time–domain expression of the signal is as follows:

$$x(\tau) = \int x(t)\delta(t-\tau)dt \quad (7)$$

In the extracted time–domain features, feature monotonicity analysis was conducted based on the Spearman coefficients, and four features were selected: the average value, standard deviation, peak, and root mean square. The median filtering method is used to smooth the four features in Figure 4.

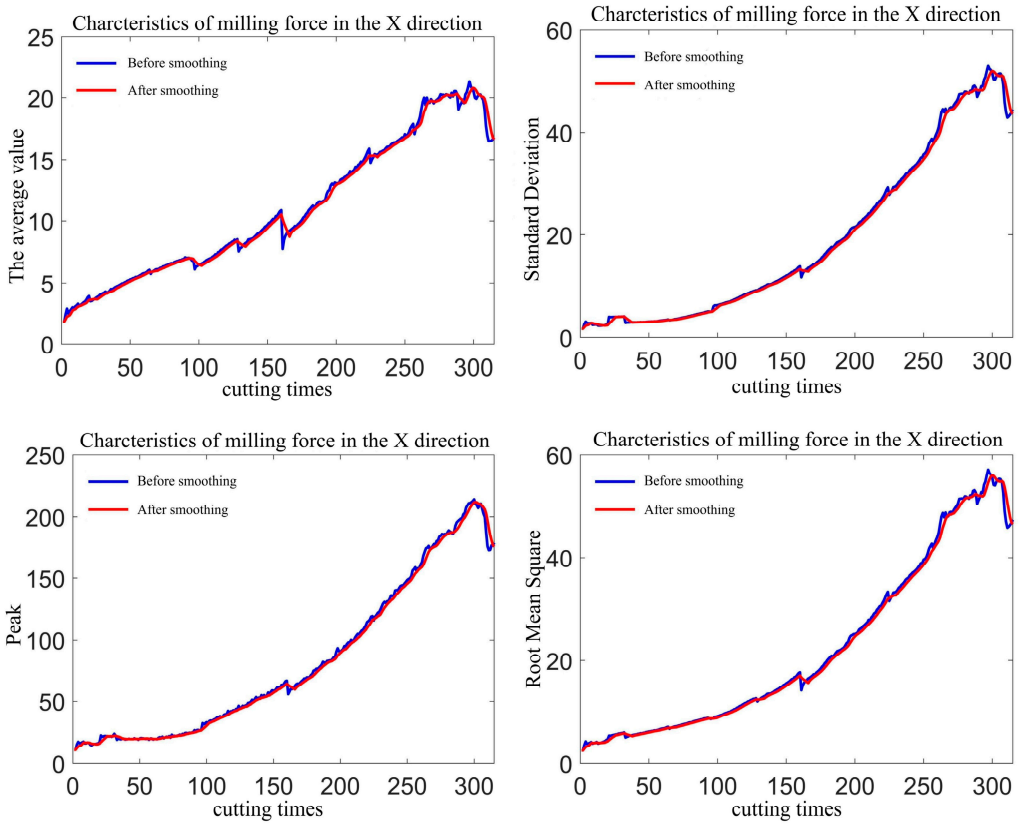


**Figure 4.** Comparison before and after feature smoothing.

### *3.3. Model Comparison and Result Analysis*

In this paper, the average amplitude of errors represents the reliability of tool wear prediction, and is measured using the Mean Square Error (MSE), which provides an objective description of the predictive performance of the proposed model. To determine whether the tool wear monitoring model is applicable, use the Mean Absolute Error (MAE) to reflect the size of the average error value. Within the same dataset, the model performs better the smaller the error value and the smaller the MSE and MAE indicators. The formula for the two indicators is as follows:

MSE represents the mean of the sum of squared differences between the predicted and test values [10]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\tilde{y}_i - y_i)^2 \tag{8}$$

MAE represents the mean absolute value of the difference between the predicted value and the test value [10]:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\tilde{y}_i - y_i| \tag{9}$$

In the above equation, $\tilde{y}_i$ represents the predicted value, $y_i$ represents the test value, and $n$ represents the total number of samples.

The tool wear dataset has three sets of data: C1, C4, and C6. Two sets are selected each time as the training set, and the third set is the test set, as shown in Table 3. Using the MSE and MAE as evaluation criteria, the DAE–SVR model will be compared and evaluated with traditional DAE and SVR models. In order to better verify the accuracy of the

proposed model, this article also compares and analyzes it with two methods: the random forest and convolutional neural network (CNN) methods.

**Table 3.** Training set and test set configuration.

| Training Set | Test Set |
|---|---|
| C4 + C6 | C1 |
| C1 + C6 | C4 |
| C1 + C4 | C6 |

From Figures 5–7, the early stages of the C1, C4, and C6 datasets can be seen. In Figure 5a, the monitoring curve of the DAE–SVR model has a good fit with the real data curve. This indicates that DAE–SVR can quickly and accurately capture the relationship between tool wear and cutting times, providing reliable monitoring results in the initial stage. In Figure 5b, 5c, for the DAE and SVR models, they were able to maintain a certain level of accuracy during the initial 150 tool runs, indicating that these two traditional models can monitor tool wear to some extent. However, during 150–270 runs, they showed significant deviations. This means that traditional DAE and SVR models have an insufficient generalization ability when facing the middle part of the dataset, and cannot adapt well to changes in the data, resulting in significant errors between monitoring results and true values. During the 270–315 runs, the DAE model gradually stabilized, indicating its ability to adjust and adapt to subsequent data to some extent. However, the SVR model experienced fluctuations due to the influence of erroneous data. This reflects that the SVR model has poor robustness to abnormal data, and the reliability of its monitoring results will be greatly affected when encountering erroneous data. Overall, DAE–SVR significantly outperforms traditional DAE and SVR models in terms of accuracy. Although the accuracy difference between DAE and SVR is relatively small, their performance in dealing with complex data is not as good as DAE–SVR, indicating that DAE–SVR has better performance in tool wear monitoring tasks by combining the advantages of DAE and SVR.

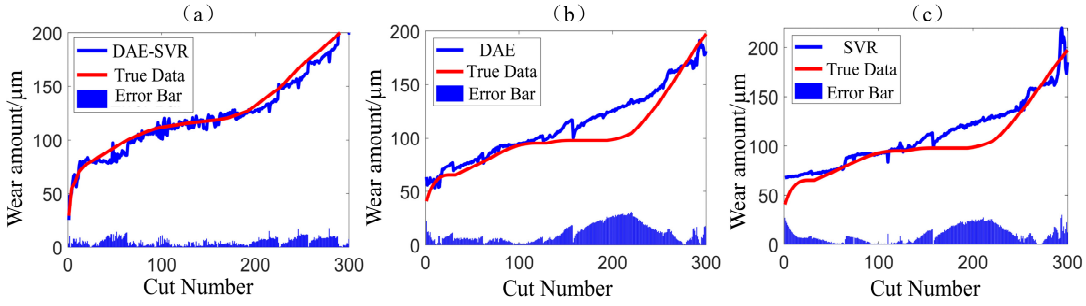


**Figure 5.** Model comparison when C1 is used as test set. (**a**) DAE–SVR. (**b**) DAE. (**c**) SVR.

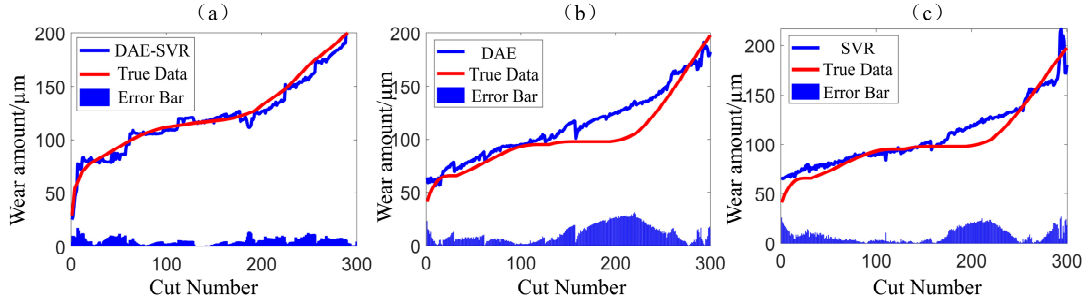


**Figure 6.** Model comparison when C4 is used as test set. (**a**) DAE–SVR. (**b**) DAE. (**c**) SVR.

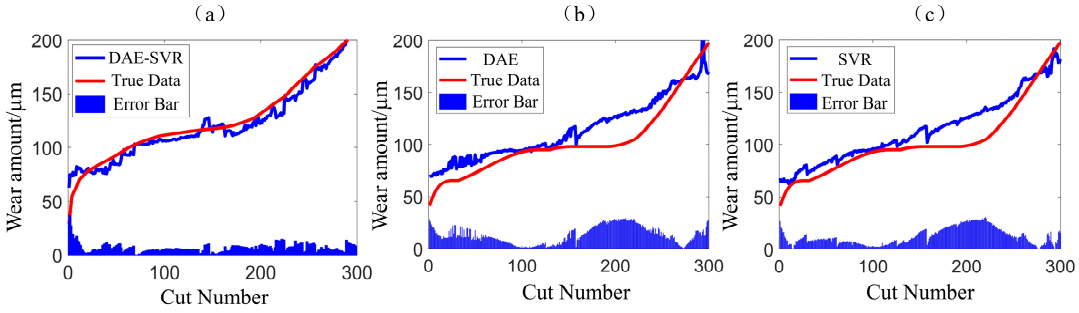


**Figure 7.** Model comparison when C6 is used as test set. (**a**) DAE–SVR. (**b**) DAE. (**c**) SVR.

From Figure 8a, it can be seen that the DAE–SVR model curve closely matches the real data curve throughout the entire cutting frequency range, accurately capturing the trend of tool wear changes, with small and stable prediction errors, high reliability, and a strong generalization ability. It can be effectively used for tool wear monitoring. In Figure 8b, the CNN model curve deviates from the real data in some stages. The predicted values in the early stage of wear are close to the real values, but the deviation increases in the middle and later stages. When processing complex tool wear data, the grasp of the changing trend is not accurate enough, the prediction accuracy decreases, and the performance is not as good as DAE–SVR. In Figure 8c, the curve of the random forest model deviates significantly from the real data and does not fit well throughout the entire range of cutting times. There are shortcomings in feature extraction and model construction, making it difficult to monitor accurately and meaning it performs worse than DAE–SVR. By comparison, DAE–SVR has outstanding advantages in tool wear monitoring, while the CNN and random forest models have significant room for improvement in prediction accuracy. In actual industrial production, DAE–SVR can provide a reliable decision-making basis for tool replacement timing, help improve production efficiency, reduce costs, and have a higher application value.

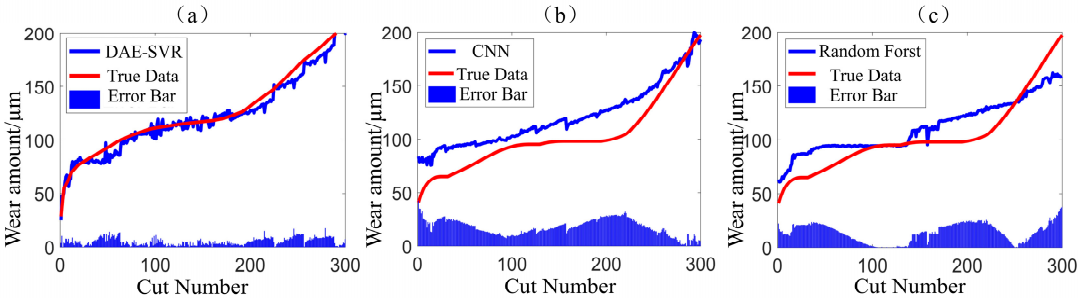


**Figure 8.** Comparison of DAE–SVR with CNN and random forest models. (**a**) DAE–SVR. (**b**) CNN. (**c**) RF.

Table 4 shows the performance comparison between the DAE–SVR model and several other models. In terms of the MAE evaluation metric, the DAE–SVR model exhibits significant advantages. Compared with the DAE and SVR models, it achieved a performance improvement of 43.58% and 41.29%, respectively. Its MAE value remained below 10, indicating that DAE–SVR can highly accurately predict tool wear and is therefore very suitable for tool wear monitoring tasks. For MSE, DAE–SVR also outperforms the other models. Compared to DAE and SVR, it shows improvements of 23.84% and 25.16%, respectively. A lower MSE value means that the tool wear values predicted by DAE–SVR are on average closer to the actual values. In direct comparison with CNN and random forest models, the superiority of the DAE–SVR model is once again demonstrated. The latter two show relatively high MAE and MSE values on each dataset, indicating their

lower accuracy in predicting tool wear. The data are further supported by the MAE and MSE box plots in Figures 9 and 10. The values of DAE–SVR are significantly more concentrated and lower than other models, which directly proves the accuracy of DAE–SVR monitoring at the data level.

**Table 4.** Model performance comparison.

| Group | Model | MAE | MSE | Run Time (s) | Performance Improvement (MAE) | Performance Improvement (MSE) |
|---|---|---|---|---|---|---|
| C1 | DAE | 12.2662 | 147.4307 | 45 | 23.35% | 15.37% |
| | SVR | 11.6603 | 149.8447 | 47 | 19.36% | 16.74% |
| | DAE–SVR | 9.4023 | 124.7664 | 32 | | |
| C4 | DAE | 12.7213 | 137.6456 | 48 | 36.26% | 8.06% |
| | SVR | 13.8156 | 133.7265 | 46 | 41.29% | 5.36% |
| | DAE–SVR | 8.1101 | 126.5564 | 31 | | |
| C6 | DAE | 14.8916 | 159.1643 | 49 | 43.58% | 23.84% |
| | SVR | 11.2535 | 161.9679 | 44 | 25.34% | 25.16% |
| | DAE–SVR | 8.4019 | 121.2238 | 32 | | |
| | CNN | 15.0072 | 161.4438 | 48 | 45.96% | 24.91% |
| | Random forest | 17.6286 | 165.9262 | 46 | 53.99% | 26.94% |

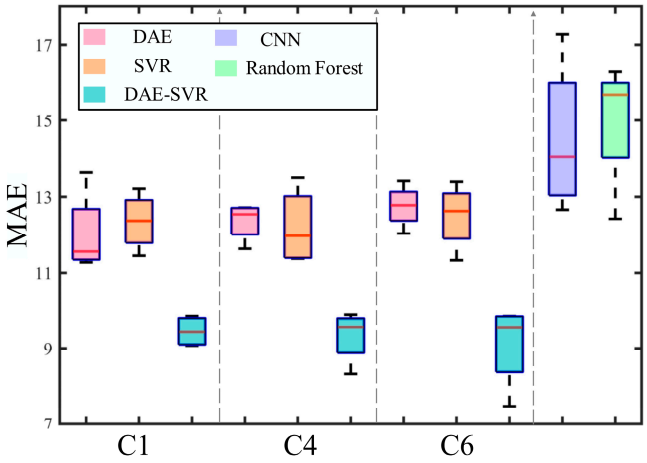


**Figure 9.** MAE box diagram of each model.

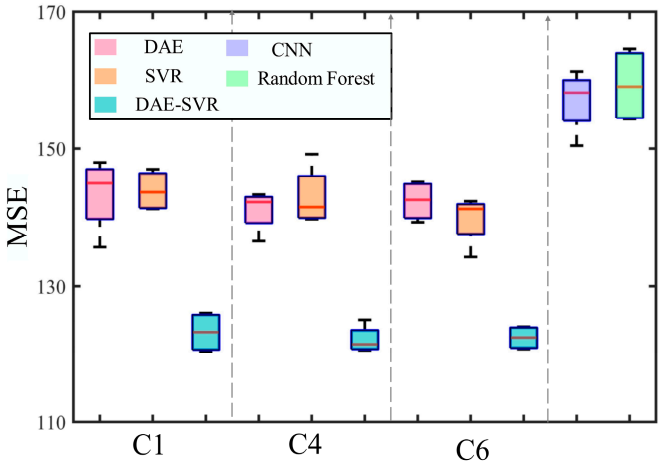


**Figure 10.** MSE box diagram of each model.

## 4. Conclusions and Future Work

In response to the problem of the ineffective compression of a large amount of cutting data generated during machine tool processing, which leads to the inability to obtain reliable feature data, as well as the low robustness and monitoring accuracy of traditional methods against damaged data, corresponding research has been conducted, and the following conclusions have been drawn:

(1) DAE–SVR has significantly better accuracy than DAE and SVR. In the early stages, DAE–SVR can quickly and accurately capture the relationship between tool wear and cutting frequency, with high fitting accuracy. However, DAE and SVR are only accurate in the initial 150 runs, with large deviations and insufficient generalization between 150 and 270 runs, and a poor anti-interference ability of SVR in the later stages.

(2) DAE–SVR closely matches real data throughout the entire cutting process, with small prediction errors and a strong generalization ability. CNN is initially accurate, but later has large deviations, making it difficult for it to grasp complex data patterns. RF deviates significantly throughout the entire process, and there are deficiencies in feature extraction and model construction, both of which have worse values than with DAE–SVR.

In the field of tool wear monitoring, the DAE–SVR model has excellent performance, far exceeding models such as DAE, SVR, CNN, and RF. The DAE–SVR model can better grasp the tool wear status in actual machining, provide a reliable basis for tool replacement timing, and effectively avoid problems such as reduced production efficiency and increased costs caused by excessive tool wear.

The model proposed in this article has achieved certain results in the field of tool wear monitoring and prediction, but there are still some shortcomings and many works that need further deepening and improvement. How to select and optimize features for complex working conditions in actual machining will become an important research topic for subsequent issues related to tool wear. In the training process of the model, there is currently no theoretical criterion for how to choose the best network parameters, and most of them rely on experience to determine whether the training results meet expectations. Therefore, the next research content can focus on how to choose the best network parameters. There are various types of cutting tools and materials used in processing, and further research is needed on the combination of other cutting tools and workpieces.

**Author Contributions:** Conceptualization, X.S., C.L.; methodology, X.S., Y.Z.; software, X.S.; validation, Z.Y., M.X. (Maojin Xia), M.X. (Min Xia); data curation, X.S., Y.Z.; writing—original draft preparation, X.S.; writing—review and editing, C.L.; supervision, C.L.; project administration, C.L., Y.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the talent scientific research fund of LIAONING PETROCHEMICAL UNIVERSITY (NO. 2021XJJL-005), Liaoning Provincial Natural Science Foundation Program (No. 2022-BS-293), Basic Scientific Research Project of Liaoning Provincial Department of Education (No. LJKMZ20220718) and Fushun Revitalization Talents Program (No. FSYC202207005), Science and Technology Research Project of Liaoning Provincial Department of Education (No. LJKZ0386).

**Data Availability Statement:** The data set was queried by the Prognostics and Health Management Society.

**Conflicts of Interest:** Author Maojin Xia was employed by the China Petroleum Engineering & Construction Corp. Beijing Branch. The remaining authors declare that the research was conducted

in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References


1. Torabi, A.J.; Er, M.J.; Li, X.; Lim, B.S.; Peen, G.O. Application of clustering methods for online tool condition monitoring and fault diagnosis in high-speed milling processes. *IEEE Syst. J.* **2016**, *10*, 721–732. https://doi.org/10.1109/JSYST.2015.2425793.
2. Salonitis, K.; Kolios, A. Reliability assessment of cutting tool life based on surrogate approximation methods. *Int. J. Adv. Manuf. Technol.* **2014**, *71*, 1197–1208. https://doi.org/10.1007/s00170-013-5560-2.
3. Vetrichelvan, G.; Sundaram, S.; Kumaran, S.S.; Velmurugan, P. An investigation of tool wear using acoustic emission and genetic algorithm. *J. Vib. Control* **2015**, *21*, 3061–3066. https://doi.org/10.1177/1077546314520835.
4. Aramesh, M.; Shaban, Y.; Yacout, S.; Attia, M.H.; Kishawy, H.A.; Balazinski, M. Survival life analysis applied to tool life estimation with variable cutting conditions when machining titanium metal matrix composites (Ti-MMCs). *Mach. Sci. Technol.* **2016**, *20*, 132–147. https://doi.org/10.1080/10910344.2015.1133916.
5. Dutta, S.; Pal, S.; Mukhopadhyay, S.; Sen, R. Application of digital image processing in tool condition monitoring: A review. *CIRP J. Manuf. Sci. Technol.* **2013**, *6*, 212–232. https://doi.org/10.1016/j.cirpj.2013.02.005.
6. Yao, C.-W.; Hong, T.W. Evaluating tool wear by measuring the real-time contact resistance. *Int. J. Adv. Manuf. Technol.* **2019**, *100*, 2349–2355. https://doi.org/10.1007/s00170-018-2815-y.
7. Zhou, Y.; Xue, W. Review of tool condition monitoring methods in milling processes. *Int. J. Adv. Manuf. Technol.* **2018**, *96*, 2509–2523. https://doi.org/10.1007/s00170-018-1768-5.
8. Wang, J.; Xie, J.; Zhao, R.; Zhang, L.; Duan, L. Multisensory fusion based virtual tool wear sensing for ubiquitous manufacturing. *Robot. Comput.-Integr. Manuf.* **2017**, *45*, 47–58. https://doi.org/10.1016/j.rcim.2016.05.010.
9. Zhu, R.; Peng, W.; Han, Y.; Huang, C.-G. Intelligent health monitoring of machine tools using a bayesian multibranch neural network. *IEEE Syst. J.* **2022**, *22*, 12183–12196. https://doi.org/10.1109/JSEN.2022.3175722.
10. Quan, Y.; Liu, C.; Yuan, Z.; Zhou, Y. An intelligent multiscale spatiotemporal fusion network model for TCM. *IEEE Sens. J.* **2023**, *23*, 6628–6637. https://doi.org/10.1109/JSEN.2023.3244587.
11. Ambhore, N.; Kamble, D.; Chinchanikar, S.; Wayal, V. Tool condition monitoring system: A review. *Mater. Today Proc.* **2015**, *2*, 3419–3428. https://doi.org/10.1016/j.matpr.2015.07.317.
12. Wang, G.; Zhang, F. A sequence-to-sequence model with attention and monotonicity loss for tool wear monitoring and prediction. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 1–11. https://doi.org/10.1109/TIM.2021.3117082.
13. Oguamanam, D.; Raafat, H.; Taboun, S. A machine vision system for wear monitoring and breakage detection of single-point cutting tools. *Pergamon* **1994**, *26*, 575–598. https://doi.org/10.1016/0360-8352(94)90052-3.
14. Lanzetta, M. A new flexible high-resolution vision sensor for tool condition monitoring. *J. Mater. Process. Technol.* **2001**, *119*, 73–82. https://doi.org/10.1016/S0924-0136(01)00878-0.
15. Kuntoğlu, M.; Sağlam, H. Investigation of signal behaviors for sensor fusion with tool condition monitoring system in turning. *Measurement* **2021**, *173*, 108582. https://doi.org/10.1016/j.measurement.2020.108582.
16. An, Q.; Tao, Z.; Xu, X.; El Mansori, M.; Chen, M. A data-driven model for milling tool remaining useful life prediction with convolutional and stacked LSTM network. *Measurement* **2020**, *154*, 107461. https://doi.org/10.1016/j.measurement.2019.107461.
17. Wang, J.; Yan, J.; Li, C.; Gao, R.X.; Zhao, R. Deep heterogeneous GRU model for predictive analytics in smart manufacturing: Application to tool wear prediction. *Comput. Ind.* **2019**, *111*, 1–14. https://doi.org/10.1016/j.compind.2019.06.001.
18. Lubis, S.; Rosehan; Darmawan, S.; Indra, B. Tool wear analysis of coated carbide tools on cutting force in machining process of AISI 4140 steel. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *852*, 012083. https://doi.org/10.1088/1757-899X/852/1/012083.
19. Duan, J.; Shi, T.; Zhou, H.; Xuan, J.; Wang, S. A novel ResNet-based model structure and its applications in machine health monitoring. *J. Vib. Control* **2021**, *27*, 1036–1050. https://doi.org/10.1177/1077546320936506.
20. Liu, Y.; Hu, X.; Yan, S.; Sun, S. Tool condition monitoring and degradation estimation in rotor slot machining process. *International J. Adv. Manuf. Technol.* **2017**, *91*, 39–48. https://doi.org/10.1007/s00170-016-9706-x.
21. Dimla, D.; Lister, P. On-line metal cutting tool condition monitoring: I: Force and vibration analyses. *Int. J. Mach. Tools Manuf.* **2000**, *40*, 739–768. https://doi.org/10.1016/S0890-6955(99)00084-X.
22. Vincent, P.; LaRochelle, H.; Bengio, Y.; Manzagol, P.-A. 2008 Extracting and composing robust features with denoising autoencoders. In Proceedings of the 25th International Conference on Machine Learning (ICML '08), Helsinki, Finland, 5–9 July 2008; Association for Computing Machinery: New York, NY, USA, 2008; pp. 1096–1103. https://doi.org/10.1145/1390156.1390294.

23. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88.