

Article

Q-Learning-Incorporated Robust Relevance Vector Machine for Remaining Useful Life Prediction

Xiuli Wang ^{1,2}, Zhongxin Li ¹, Xiuyi Wang ^{3,*} and Xinyu Hu ⁴

¹ College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China; sherrywang@zjut.edu.cn (X.W.); lzx2524408535@163.com (Z.L.)

² College of Software, Nankai University, Tianjin 300350, China

³ Zhejiang Academy of Special Equipment Science, Hangzhou 310020, China

⁴ State Grid Nantong Power Supply Company, Nantong 226533, China; 13914378978@163.com

* Correspondence: zq_wxy93@163.com

Abstract: Accurate and reliable remaining useful life (RUL) prediction is crucial for improving equipment reliability and safety, realizing predictive maintenance. The relevance vector machine (RVM) method is commonly utilized for RUL prediction, profiting from its sparse property under a Bayesian framework. However, the RVM faces the issue of poor robustness, which is mainly manifested as poor prediction accuracy and difficulty in fitting when the predicted data fluctuate greatly. This is due to weights and random errors following Gaussian distributions, which are highly sensitive to outliers. Also, the traditional model training process heavily relies on an additional feature extraction process, which suffers from the problem of effective data loss as well as the risk of overfitting. Thus, a robust regression framework against outliers is developed by incorporating *t*-distribution into the RVM. And a Q-learning (QL) algorithm is embedded into the constructed robust RVM model to replace the feature extraction process. In addition, this paper firstly predicts the degradation trend of RUL to enhance the accuracy and interpretability of RUL prediction. Finally, a comparative experiment on the performance degradation of capacitors in the traction system is designed, and the root mean square errors for the QL-RRVM, QL-RVM, RRVM, and RVM models are obtained as 0.751, 8.599, 38.316, and 41.892, respectively. The experimental results confirm the superiority of the proposed method.



Citation: Wang, X.; Li, Z.; Wang, X.; Hu, X. Q-Learning-Incorporated Robust Relevance Vector Machine for Remaining Useful Life Prediction.

Processes **2024**, *12*, 2536. <https://doi.org/10.3390/pr12112536>

Academic Editor: Yuhe Wang

Received: 8 October 2024

Revised: 7 November 2024

Accepted: 12 November 2024

Published: 13 November 2024



Copyright: © 2024 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/>).

Keywords: Q-learning (QL); relevance vector machine (RVM); robustness; RUL prediction

1. Introduction

Remaining Useful Life (RUL) prediction plays a crucial role in industrial systems for achieving efficient equipment maintenance and management [1]. Accurately predicting the RUL of equipment or components enables engineers to effectively avoid unexpected equipment failures in advance, and to strategically implement preventive maintenance, minimizing downtime and reducing production costs [2,3].

RUL prediction methods can be broadly classified into two main categories: model-based methods and data-driven methods. The model-based methods utilize physics models [4–6], mechanistic models [7,8] or stochastic process [9] models for RUL prediction. A physical equation was established and converted into a multivariate linear regression model to improve generalization capability of cold load prediction in Reference [10]. The storage life of a certain electronic component was predicted by analyzing its failure process based on a reaction-rate model in Reference [11]. A nonlinear dynamic system was established by the tensor product-based model transformation technique in Reference [12]. However, heavy reliance on prior knowledge and expert experience poses a significant obstacle to model-based approaches. It is extremely difficult to accurately obtain the physical models of failure mechanism with the increasing complexity of engineering equipment [13].

Data-driven approaches are booming in RUL prediction, as they do not require physical knowledge to infer the degradation process. Data-driven RUL prediction methods, such as neural networks [14–16] and the relevance vector machine (RVM) [17], aim to analyze historical and monitoring data to uncover the underlying relationships among data, enabling the prediction of degradation trends and subsequent analysis of RUL [18]. Neural network methods play a dominant role in the data-driven approach [19]. For instance, an improved Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) were combined to enhance the prediction inaccuracy of lithium battery-life prediction in Reference [15]. A deep Bayesian learning method was proposed for fast and uncertainty-aware RUL prediction in Reference [16]. However, the prediction accuracy derived from neural networks is limited, due to the number of available data. In contrast, the RVM method, a small-sample prediction approach, demonstrates excellent performance when faced with insufficient data. The RVM method is under a Bayesian framework, where the weights for each input are controlled by a set of parameters [17,20]. RVM is widely applied in regression tasks, due to its sparsity and excellent prediction performance [21–23]. For example, Ref. [24] utilized an RVM method to describe a degradation process from initial fault to failure. Additionally, a degradation path-based RUL prediction framework is established using a dynamic multivariate RVM model in Reference [25].

While RVMs demonstrate considerable promise in achieving high accuracy and sparsity, their predictive prowess can falter in the presence of substantial data variability and outliers [26]. This limitation stems from the fact that the weights and random noise in RVMs adhere to Gaussian distributions, which are inherently susceptible to the influence of outliers and atypical data points. Furthermore, the application of RVM models for predictive or classification tasks typically necessitates an auxiliary feature-extraction procedure. For example, in Reference [27], features are derived from the battery-charging voltage curve through the application of random forest and Gaussian regression techniques. Similarly, in Reference [28], Principal Component Analysis (PCA) is employed to both reduce data dimensionality and extract pertinent features. However, these supplementary feature extraction techniques are not only contingent upon human expertise but also carry the risk of inadvertently leading to the loss or degradation of crucial information.

Motivated by these discussions, the RVM is extended into a robust one by integrating t -distribution into the RVM framework, for striving against outliers. And a Q -learning algorithm from the realm of reinforcement learning is introduced into the newly formulated RRVM model. In Q -learning, the state is the candidate feature, which transfers to the next state via random action. After all actions are performed, the optimal feature is selected corresponding to each time point. This integration culminates in the development of a QL-RRVM model that is capable of directly processing raw data and delivering precise predictions without the need for intermediate feature extraction steps. Specifically, the contributions of this study are summarized as follows.

1. It is inferred that the weights and random errors of the proposed model obey t -distributions, with the assumption that hyperparameters are gamma distributions, which are more robust for outliers than classical RVM.
2. The QL algorithm is introduced into the RRVM model for feature extraction to improve the model prediction accuracy.
3. In the field of RUL prediction, the model transitions from static to dynamic, employing a forward-rolling prediction within a time-series context. It initially projects degradation trends essential for RUL estimation and subsequently refines these projections to forecast RUL, thereby improving the interpretability of the predictions.

The remainder of this paper is organized as follows. The RRVM model is given in Section 2. RUL prediction is performed in Section 3. QL and RRVM models are combined in Section 4. In Section 5, the effectiveness of the proposed approach is validated by a case study. Finally, a conclusion is summarized in Section 6.

2. Robust RVM Modelling

2.1. Modelling Process

For a nonlinear regression task, it is common to use the kernel function to represent the nonlinear relationship between the input vectors $\{x_i\}_{i=1}^N$ and corresponding targets $\{t_i\}_{i=1}^N$. Typically, predictions are based on a function $y(x)$ defined over the input space, with “learning” being the process of inferring this function (potentially its parameters). A flexible and popular candidate set for $y(x)$ is as follows:

$$y(x) = \sum_{i=1}^N w_i \varphi_i(x) = \Phi W \quad (1)$$

where $\Phi = [\varphi(x_1), \varphi(x_2), \dots, \varphi(x_N)]^T \in \mathbb{R}^{N \times N+1}$ is a design matrix, in which $\varphi(x_i) = [1, k(x_i, x_1), k(x_i, x_2), \dots, k(x_i, x_N)] \in \mathbb{R}^{1 \times (N+1)}$ and $k(x_i, x_k)$ is a Gaussian kernel function between the vector x_i and x_k ($k = 1, \dots, N$) in this study; $W = [w_1, w_2, \dots, w_{N+1}]^T \in \mathbb{R}^{(N+1) \times 1}$ denotes the weight matrix.

In the context of the RVM model, a time-series dataset $\{x_i\}_{i=1} \in \mathbb{R}^{1 \times (N+L)}$ is provided, with $x_i = (x_i, x_{i+1}, \dots, x_{i+L-1}) \in \mathbb{R}^{1 \times L}$, where N denotes the number of target groups, and L represents the stride length of the sliding window. The target value can be represented as the sum of additive noise and the actual value

$$t_i = y(x_i) + \varepsilon \quad (2)$$

where ε is assumed to follow a Gaussian distribution with zero mean and variance σ^2 . The precision β is introduced as $\beta^{-1} = \sigma^2$. Then, the probability density function (PDF) of the predicted value t_i with respect to w_i and β^{-1} can be expressed as

$$p(t_i | x_i, w_i, \beta^{-1}) \sim \mathcal{N}(y(x_i), \beta^{-1}) = (2\pi\beta^{-1})^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\beta^{-1}}(t_i - w_i\varphi(x_i))^2\right\} \quad (3)$$

N groups of data are considered in the following format:

$$p(T | X, W, \beta^{-1}) \sim \prod_{i=1}^N \mathcal{N}(\Phi W, \beta^{-1}) = (2\pi\beta^{-1})^{-\frac{N}{2}} \exp\left\{-\frac{1}{2\beta^{-1}}\|T - \Phi(X)W\|^2\right\} \quad (4)$$

where T, X, W correspond to the matrix forms of t_i, x_i, w_i , respectively.

The parameter W is defined as obeying a Gaussian distribution with zero mean and different variances to guarantee sparsity prediction in RVM. Its PDF is given as follows.

$$p(W | \alpha) = \prod_{i=1}^{N+1} \mathcal{N}(w_i | 0, \alpha_i^{-1}) \quad (5)$$

where $\alpha = (\alpha_1, \dots, \alpha_{N+1})^T \in \mathbb{R}^{N+1}$ and $\alpha_i, i \in [1, N+1]$ denote precision corresponding with W .

The training objective is to estimate the parameters W, β and α by maximizing the conditional PDF $p(W, \alpha, \beta | t_i)$ with the target t_i

$$p(W, \alpha, \beta | t_i) = p(W | t_i, \alpha, \beta) p(\alpha, \beta | t_i) \quad (6)$$

The posterior probability distribution of the parameter W remains Gaussian. Then, there is

$$p(W | t_i, \alpha, \beta) = \mathcal{N}(W | \Lambda, \Sigma) \quad (7)$$

where $\Lambda = \beta \Sigma \Phi^T T$ is the mean; $\Sigma = (A + \beta \Phi^T \Phi)^{-1}$ denotes the variance, in which $A = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_{N+1}) \in \mathbb{R}^{N+1}$.

The poor robustness of the traditional RVM model stems from the assumption that both the weights and random errors follow Gaussian distributions, which are sensitive to outliers and anomalous data points. Thus, the parameters α and β are assumed to follow a gamma distribution in this study. Then, the predicted values follow the t -distribution, which is more robust than the Gaussian distribution. A gamma distribution is introduced for the parameters α and β . And their corresponding PDFs are given below.

$$p(\alpha) = \prod_{i=0}^l \Gamma(\alpha_i | a, b), \quad p(\beta) = \Gamma(\beta | c, d) \quad (8)$$

where a, c are shape parameters; b and d are scale parameters; $\Gamma(x|y, z) = \Gamma(y)^{-1} z^y x^{y-1} e^{-zx}$ in which $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$.

Then, the posterior PDFs of weights $\mathbf{W}$ and random errors ε are

$$p(\mathbf{W}) = \int p(\mathbf{W}|\alpha) p(\alpha) d\alpha = \frac{\Gamma\left(\frac{\nu}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\lambda}{\pi\nu}\right)^{\frac{1}{2}} \left(1 + \frac{\lambda \mathbf{W}^2}{\nu}\right)^{-\left(\frac{\nu}{2} + \frac{1}{2}\right)} = t(\mathbf{W}) \quad (9)$$

$$p(\varepsilon) = \int p(\varepsilon|\beta) p(\beta) d\beta = \frac{\Gamma\left(\frac{\nu'}{2} + \frac{1}{2}\right)}{\Gamma\left(\frac{\nu'}{2}\right)} \left(\frac{\lambda'}{\pi\nu'}\right)^{\frac{1}{2}} \left(1 + \frac{\lambda' \varepsilon^2}{\nu'}\right)^{-\left(\frac{\nu'}{2} + \frac{1}{2}\right)} = t(\varepsilon) \quad (10)$$

where the parameters $\lambda = a/b$ ($\lambda' = c/d$) and $\nu = 2a$ ($\nu' = 2c$) present the precision and degrees of freedom of the weights $\mathbf{W}$ (error ε) obeying the t -distribution.

The Equations (9) and (10) suggest that both the weights $\mathbf{W}$ and the random errors ε adhere to t -distribution. It is noteworthy that the t -distribution is distinguished by the presence of extended tails, a characteristic that sets it apart from the Gaussian distribution. Thus, the established regression model (2) is robust by assuming that the regression model coefficients $\mathbf{W}$ and the random errors ε follow a long-tailed t -distribution.

The subsequent section deduces the forecasts of the model, which is predicated on the completion of parameter estimation (the parameter estimation process is detailed in Section 3), assuming optimal estimation of parameters α and β has been achieved. Initially, the model is trained assuming that under the condition at time i , the most fitting α and β for the current data, obtained by training the model with data from time $i - 1$, is denoted as α_{i-1}^{best} and β_{i-1}^{best} . Then, we calculate the predictive distribution of data X_{i-1} by combining Equations (10) and (11) as follows.

$$p(t_i | t_{i-1}) = \int p(t_i | \mathbf{W}, \alpha, \beta) p(\mathbf{W}, \alpha, \beta | t_{i-1}) d\mathbf{W} d\alpha d\beta \quad (11)$$

$$p(t_i | t_{i-1}, \alpha_{i-1}^{best}, \beta_{i-1}^{best}) = \int p(t_i | \mathbf{W}, \beta_{i-1}^{best}) p(\mathbf{W} | t_{i-1}, \alpha_{i-1}^{best}, \beta_{i-1}^{best}) d\mathbf{W} \quad (12)$$

Upon incorporating the PDF of the respective distribution into the aforementioned equation and performing simplifications, it is determined that the resulting form aligns with the properties of the Gaussian distribution, i.e., $p(t_i | t_{i-1}, \alpha_{i-1}^{best}, \beta_{i-1}^{best}) = \mathcal{N}(t_i | y_i, \beta_i)$, where $y_i = \mathbf{W}^T \Phi(\mathbf{x}_{i-1})$ and $\beta_i = \beta_{i-1}^{best} + \Phi(\mathbf{x}_{i-1})^T \Sigma \Phi(\mathbf{x}_{i-1})$. In this study, the mean value y_i is considered as the predicted value for the data $\mathbf{x}_{i-1}$. Similarly, a prediction for data at moment $i + 1$ has $p(t_{i+1} | t_i, \alpha_i^{best}, \beta_i^{best}) = \mathcal{N}(t_{i+1} | y_{i+1}, \beta_{i+1})$, where $y_{i+1} = \mathbf{W}^T \Phi(\mathbf{x}_i)$ and $\beta_{i+1} = \beta_i^{best} + \Phi(\mathbf{x}_i)^T \Sigma \Phi(\mathbf{x}_i)$.

Employing this method, the time-series data is sequentially forecasted, and the actual data at time $i + L + 1$ are utilized as the training target for X_i . This approach aims to progressively reduce the error, ultimately achieving accurate rolling forecasts.

2.2. Parameter Estimation

In this section, parameters α and β in the regression model are estimated. The optimal parameters are sought by maximizing the posterior probability, $\max(p(\alpha, \beta|t_i))$.

The PDF of the predicted values with respect to the parameters α, β are calculated according to the convolution formula, as follows [29].

$$p(t_i|\alpha, \beta) = \int_{-\infty}^{+\infty} p(t_i|W, \beta)p(W|\alpha)dW = \left(\frac{\beta}{2\pi}\right)^{\frac{1}{2}} \left(\frac{1}{\beta} + \Phi A^{-1} \Phi^T\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\left(\frac{1}{\beta} + \Phi A^{-1} \Phi^T\right)t_i^2\right) \quad (13)$$

Taking the logarithm of the above relevant equation, the following log-likelihood function is given,

$$LF = \ln p(t_i|\ln\alpha, \ln\beta) + \ln p(\ln\alpha) + \ln p(\ln\beta) \quad (14)$$

As $p(\ln\alpha) = \alpha p(\alpha)$, and $p(t_i|\ln\alpha, \ln\beta) = p(t_i|\alpha, \beta)$, the likelihood function (14), ignoring the the terms unrelated to α, β , is converted as follows.

$$LF = -\frac{1}{2} \left[\ln \left| \frac{1}{\beta} I + \Phi^T A^{-1} \Phi \right| + t_i \left(\frac{1}{\beta} I + \Phi^T A^{-1} \Phi \right)^{-1} t_i \right] + \sum_{i=1}^{N+1} (a - \ln\alpha_i - b\alpha_i) + c \ln\beta - d\beta \quad (15)$$

Then, take the partial derivatives of α, β respectively and let them be 0.

$$\frac{\partial LF}{\partial \ln\alpha} = \frac{1}{2} [1 - \alpha(\Lambda^2 + \Sigma)] + a - b\alpha = 0 \quad (16)$$

$$\frac{\partial LF}{\partial \ln\beta} = \frac{1}{2} \left[\frac{N}{\beta} - (t_i - \Phi\Lambda)^2 - \text{tr}(\Sigma\Phi^T\Phi) \right] + c - d\beta = 0 \quad (17)$$

An implicit expression for the parameters α and β can be obtained.

$$\alpha_i = \frac{\gamma_i + 2a}{\Lambda_i^2 + 2b}, \quad \beta = \frac{N - \Sigma\Lambda + 2c}{(t_i - \Phi\Lambda)^2 + 2d}, \quad i = 1, \dots, N+1. \quad (18)$$

where $\gamma_i = 1 - \alpha_i \Sigma_{ii}$, $i = 1, \dots, N+1$. and Σ_{ii} is the corresponding element in the i -th row and i -th column of matrix Σ .

Specifically, the step-by-step parameter estimation process according to Equation (18) is as follows: firstly, substitute the optimal parameters obtained in Section 2.2 into the model and provide initial values α_0 and β_0 , then calculate the values of A , Σ and Λ based on α_0 , β_0 using $A = \text{diag}(\alpha_1, \alpha_2, \dots, \alpha_M)$, $\Sigma = (A + \beta\Phi^T\Phi)^{-1}$ and $\Lambda = \beta\Sigma\Phi^T t_n$. Next, substitute the values of A , Σ , Λ , α_0 and β_0 into Equation (18) to calculate the values for the next time step, denoted as α_1 and β_1 . Thirdly, use the newly obtained α_1 and β_1 to recalculate the values of A , Σ and Λ . After obtaining the three new values for A , Σ and Λ , substitute them back into Equation (18) to obtain the values of α_2 and β_2 for the third time step. By iteratively repeating the above steps until traversing the data, a set of values for α and β can be obtained, achieving the objective of parameter estimation for two parameters in this section.

3. RUL Prediction

Assuming that historical and current observations are $\{x_1, x_2, \dots, x_k\}$, the lifetime L of moment $time_l$ can be defined according to First Hitting Time (FHT) as

$$L = \inf\{time_l : t(time_k + time_l) \equiv t_{k+l} \in B | X_{1:k}\} \quad (19)$$

where $time_l$ denotes the length of remaining time from the moment $time_k$ to the moment of failure, and l refers to the prediction step. $X_{1:k}$ denotes the measurement data from moment $time_1$ to moment $time_k$. $t(time_k + time_l)$ denotes the predicted mean value at moment $time_k + time_l$. B is a known bounded set with elements of fault thresholds, the

exact values of which are determined based on expert experience. The physical meaning of Equation (19) is that the moment when the predicted degradation trajectory exceeds the given threshold for the first time is defined as the failure moment, and the time length from the current moment to the failure moment is the remaining lifetime at the current moment.

Assuming that failure occurs when the degradation trajectory reaches the threshold $H \in \mathbf{B}$, the distribution of lifetimes at each moment before the variable reaches H is predicted next. The Cumulative Distribution Function (CDF) of the lifetime L , given the observed data $\{x_1, x_2, \dots, x_k\}$, is calculated by the following equation.

$$p\{L \leq time_l | X_{1:k}\} = 1 - p\left\{z \leq \frac{H - t_{k+l}}{\beta_{k+l}}\right\} = p\left\{z \geq \frac{H - t_{k+l}}{\beta_{k+l}}\right\} = \int_{-\infty}^{g_{k+l}} z dz \triangleq \rho(g_{k+l}) \quad (20)$$

where β_{k+l} denotes the predicted variance at moment $time_k + time_l$; z is a random variable obeying a standard normal distribution; and $g_{k+l} = (t_{k+l} - H)/\beta_{k+l}$ and $\rho(\cdot)$ is the CDF of the random variable z .

From a practical application point of view, for lifetimes there is always $L \geq 0$. Therefore, the truncation condition CDF is introduced based on $L \geq 0$. The truncation CDF for the lifetime L is given by the following equation.

$$p\{L \leq time_l | X_{1:k}, L \geq 0\} = \frac{p\{0 \leq L \leq time_l | X_{1:k}\}}{p\{L \geq 0 | X_{1:k}\}} = \frac{\rho(g_{k+l}) - \rho(g_k)}{1 - \rho(g_k)} \quad (21)$$

Differentiating on $time_l$ yields the truncation condition PDF for the remaining lifetime L as

$$p_{L|X_{1:k}, L \geq 0} = \frac{\mathcal{N}(g_{k+l}) \Delta g_{k+l}}{1 - \rho(g_k)} \quad (22)$$

where Δg_{k+l} denotes the derivative of g_{k+l} with respect to $time_{k+l}$. $\mathcal{N}(\cdot)$ denotes the PDF of a random variable obeying a standard normal distribution. Further, the desired value of RUL can be obtained as follows.

$$E_{L|X_{1:k}, L \geq 0} = \int_{-\infty}^{+\infty} L p_{L|X_{1:k}, L \geq 0} dL \quad (23)$$

4. RRVM with Embedded Q-Learning

4.1. Q-Learning Algorithm

Within the RL paradigm, the QL technique aligns with the structure of Markov Decision Processes (MDPs) [30]. It accomplishes this by mastering a Q-function designed to gauge the prospective returns of selecting a particular action within a defined state. Employing temporal difference learning to refine these estimates, the method unearths the optimal policy devoid of any pre-existing knowledge of the environment's underlying dynamics. In the throes of decision-making, the agent discerns its environs and carries out actions from a spectrum of potential choices.

At time step t , with the current environmental state S_t , the Agent selects and performs action AC_t , causing the environment to transition from S_t to S_{t+1} , and simultaneously provides the reward $RE(S_t, AC_t)$ to the Agent. The Agent continuously repeats this process until the training and learning process is concluded. In the QL algorithm [31], the action-value function $Q(S_t, AC_t)$ represents the maximum cumulative reward that the Agent will receive after choosing and taking action AC_t in state S_t . This value is determined by the immediate reward received by the Agent after choosing and executing the action, plus the value obtained by subsequently following the optimal policy. The value of $Q(S_t, AC_t)$ can be expressed by the formula

$$Q(S_t, AC_t) = RE(S_t, AC_t) + \gamma \max_{ac} Q(S_{t+1}, AC_{t+1}) \quad (24)$$

where AC represents any action from the action set, and the constant parameter $\gamma (0 \leq \gamma \leq 1)$ is known as the discount factor. During the Agent's training and learning process, it

always selects the action corresponding to the state with the maximum Q -value, and then iteratively trains, based on this policy. Through multiple training sessions, the Q -table, which stores the Q -values, is continuously updated. To ensure that QL converges at the appropriate time, an appropriate learning rate, often denoted as χ , is incorporated into the formula. With the learning rate, the update rule for $Q(S_t, AC_t)$ can be expressed as

$$Q(S_t, AC_t) = (1 - \chi)Q(S_t, AC_t) + \chi(RE(S_t, AC_t) + \gamma \max Q(S_{t+1}, AC_{t+1})) \quad (25)$$

4.2. QL-Based Feature Extraction Process

In response to the issues of simplicity in the feature extraction process, susceptibility to loss of effective information, and difficulty in extracting features with high variability in traditional machine learning methods, a feature extraction process is designed based on the QL algorithm under intelligent decision-making. This algorithm is integrated into the front of the RRVM model, providing the model with accurate feature information, which serves as an essential foundation for subsequent precise prediction processes.

The training dataset is initialized as an empty set, and a QL action list with n actions is established to determine the processing of original data. For example, action $AC_{3,i} = [0.6 \ 0.4]$ signifies that at the i -th sampling point, the third action is executed to combine the original data's median and mean in a 6:4 ratio, thereby obtaining the feature data for that particular sampling point. Using the predictive error of the RRVM model in its current state as the reward function, the specific calculation process involves first using the data formed by randomly selecting actions at corresponding points to train and validate the RRVM. Then, the training error and prediction error are combined in the same proportion as the corresponding action, to obtain the final reward function. The specific framework of the QL-RRVM model is illustrated in Figure 1.

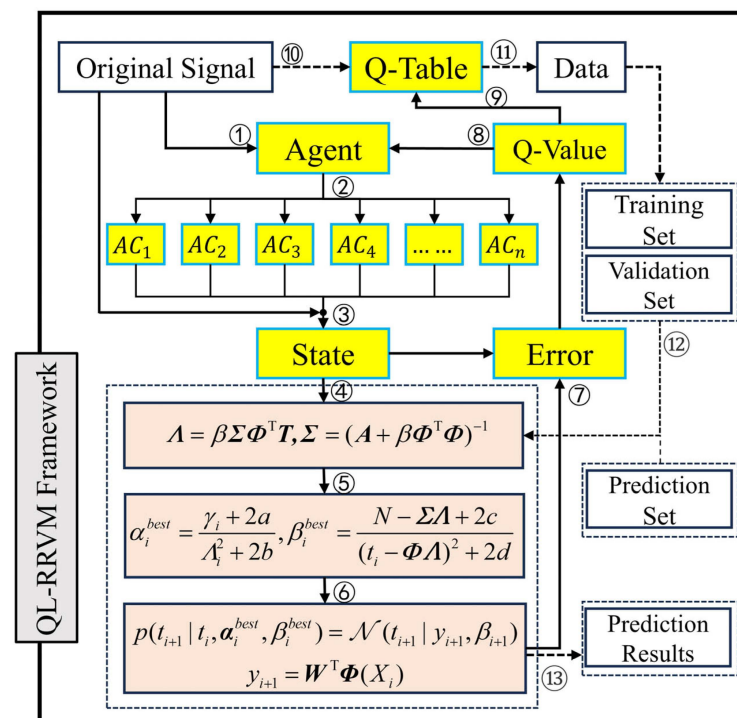


Figure 1. The overall framework of QL-RRVM.

The specific processing steps represented by steps (1) to (13) are as follows:

(1)(2): The agent receives the original signal and selects an action from the action set according to a random selection strategy.

③: The agent processes the original signal at the current moment according to the randomly selected action, using this as the current state S_i .

④⑤⑥: Using the RRVM model for training, first calculate matrices A , Λ and Σ based on the current state S_i and the Gaussian radial basis kernel function. Then, determine the optimal parameters α_i^{best} and β_i^{best} for the current moment i . Finally, calculate the predicted value $y_{i+1} = W^T \Phi(x_i)$ based on $p(t_{i+1}|t_i, \alpha_i^{best}, \beta_i^{best}) \sim \mathcal{N}(t_{i+1}|y_{i+1}, \beta_{i+1})$.

⑦: Combine the true value $\tilde{y}_{i+1}$ corresponding to the current moment's state with the predicted value y_{i+1} at the same moment to calculate the absolute value of the prediction error. In the specific calculation, both the training set and test set errors are considered: $Er_i = AC_{j,i}(|\tilde{y}_{i+1}^{train} - y_{i+1}^{train}| + |\tilde{y}_{i+1}^{test} - y_{i+1}^{test}|)$, where $AC_{j,i}$ is the j -th action randomly selected at the i -th moment, with $1 \leq j \leq 6$; y_{i+1}^{train} is the predicted value for the i -th point in the training set, and $\tilde{y}_{i+1}^{train}$ is its corresponding true value. The same applies to the test set.

⑧: The current Q-Value is saved and transmitted to the agent, which then iteratively performs steps ① to ⑦ until all actions have been traversed.

⑨⑩⑪: The Q-Values obtained from executing all actions are summarized in a Q-Table, and the original data is processed accordingly, based on the data in this table, to obtain the final training dataset, referred to as Data.

⑫⑬: Using the obtained dataset, the model is trained and validated. Finally, unknown labeled signals are input into the model to achieve prediction.

5. Case Study

5.1. Platform Description and Data Preprocessing

In this section, an experimental setup jointly developed by the China CRRZ Zhuzhou Institute and Central South University [32,33] is employed to substantiate the QL-RRVM method. Intermediate DC-link support capacitor-performance degradation data are extracted to predict the RUL of the capacitor. Figure 2 illustrates the primary components of the testbed, encompassing four main sections: (a) data acquisition and real-time monitoring; (b) signal processing; (c) traction control system; and (d) dSPACE. It is important to note that the immediate DC-link is equipped with a pair of DC-link capacitors, namely the upper and lower DC-link capacitors. These capacitors are instrumental in sustaining voltage stability for the high-speed train's traction systems. Given that two voltage sensors are installed within the DC-link, the upper and lower terminal voltages are read, to accurately mirror the capacitors' health status.

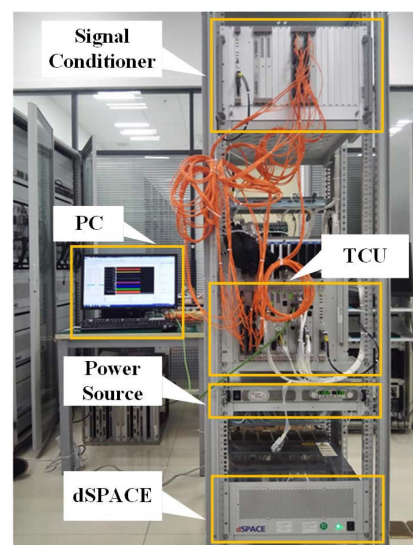


Figure 2. The experimental platform of the CRH2-type high-speed train.

Initially, the experimental platform generates eight distinct sets of original voltage signals under uniform operational conditions, differentiated by their sampling batches. These are sequentially labeled as original dataset 1 through original dataset 8. Each dataset encompasses both the upper and lower terminal voltages. Figure 3 provides a visual representation, encapsulating the entire lifespan of the system, highlighting the stages of extended normal operation, progressive degradation, and eventual shutdown. As this study focuses on the data pertaining to the degradation process, the original data are segmented, to clearly exhibit degradation patterns, which are illustrated in Figure 4.

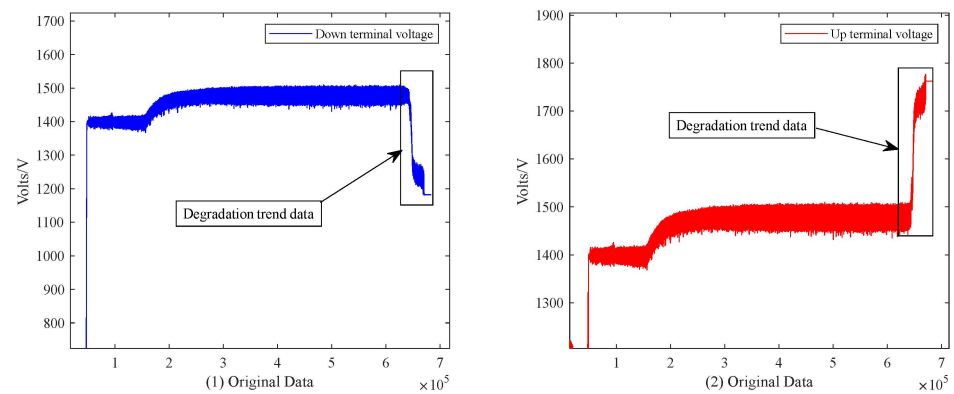


Figure 3. Full-life original voltage signals.

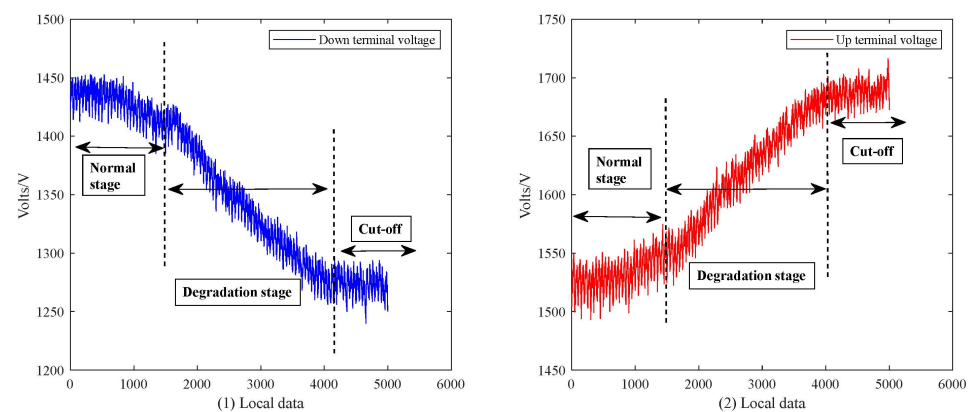


Figure 4. Original voltage signals with three stages.

5.2. QL Algorithm

In this section, the process of feature data extraction is implemented using QL, corresponding to the steps from ① to ⑪ in Figure 1. During this process, the parameters of the QL algorithm are set as follows: the learning rate α is set to 0.1, the discount factor $\gamma = 0.9$, and the exploration rate $\epsilon = 0.1$, with the number of training rounds defined as 100. The actions to be taken are defined in a matrix named 'actions', which consists of six sets of vectors, each representing a data processing action. For instance, the vector $[0.6 \ 0.4]$ signifies that the current data point is processed as 0.6 times the median plus 0.4 times the mean. The Q-table is initialized as a zero-vector matrix, and the exploration rate is designed to be greater than 0.01 and gradually decreasing. After the process, multiple sets of original data (as shown in Figure 5 (1)) are processed, resulting in the new data as depicted in Figure 5 (2).

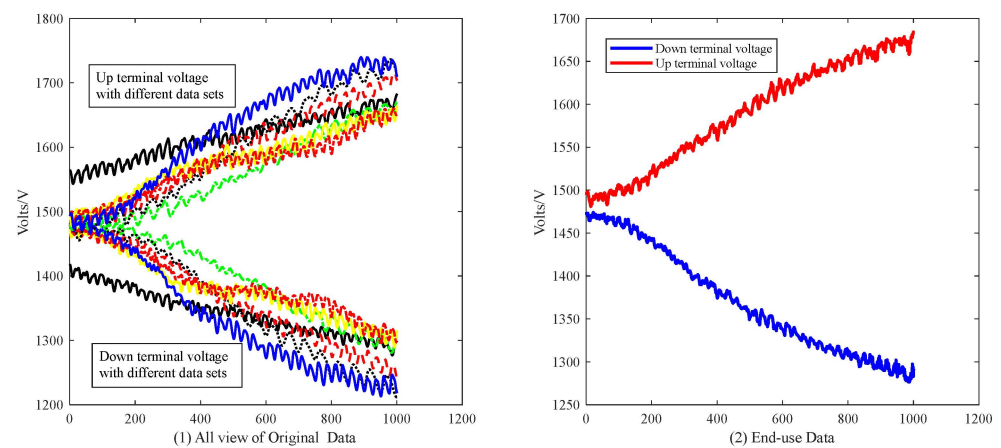


Figure 5. Comparison of original data with end-use data.

5.3. Voltage Degradation-Trend Prediction

In this subsection, a prediction of the voltage degradation trend is implemented. Initially, the RRVM model is trained using the final operational data obtained in Section 2. The hyperparameters are set with the following initial values: scalar hyperparameter to 0.2, inverse noise variance to 0.5, kernel function bandwidth to 3, and the maximum number of iterations to 50. The prediction lag L is set to 10, and the prediction step is set to 1, meaning that for each training instance, every 10 data points are used to predict the corresponding subsequent data points. The final training results of the QL-RRVM model are depicted in Figure 6, where the upper and lower curves in Figure 6a correspond to the degradation trends of the upper and lower voltages, respectively. Figure 6b,c represent the corresponding training errors. From this figure, it can be observed that the QL-RRVM model exhibits a satisfactory prediction performance for this dataset.

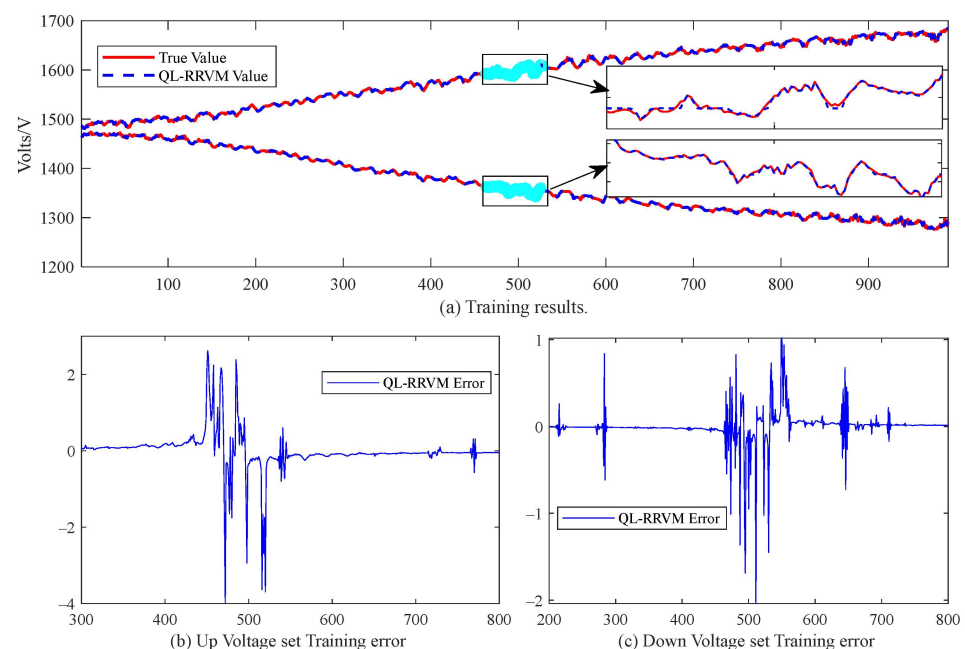


Figure 6. Training effects corresponding to the two voltage datasets.

After the training is completed, the original data from Figure 5a are used as the prediction set to verify the predictive capability of the model. Additionally, a comparative experiment is designed based on the QL-RRVM and QL-RVM models on the original dataset 6, to show the effectiveness of the robustness on outliers. Figure 7a–c correspond

to the overall prediction comparison, the lower voltage-prediction error comparison, and the upper voltage-prediction error comparison, respectively. By comparing the prediction effects of the QL-RRVM model with the QL-RVM model, it can be observed that the predictive accuracy of the improved model presented in this paper is higher, and the prediction trend is smoother, indicating better robustness.

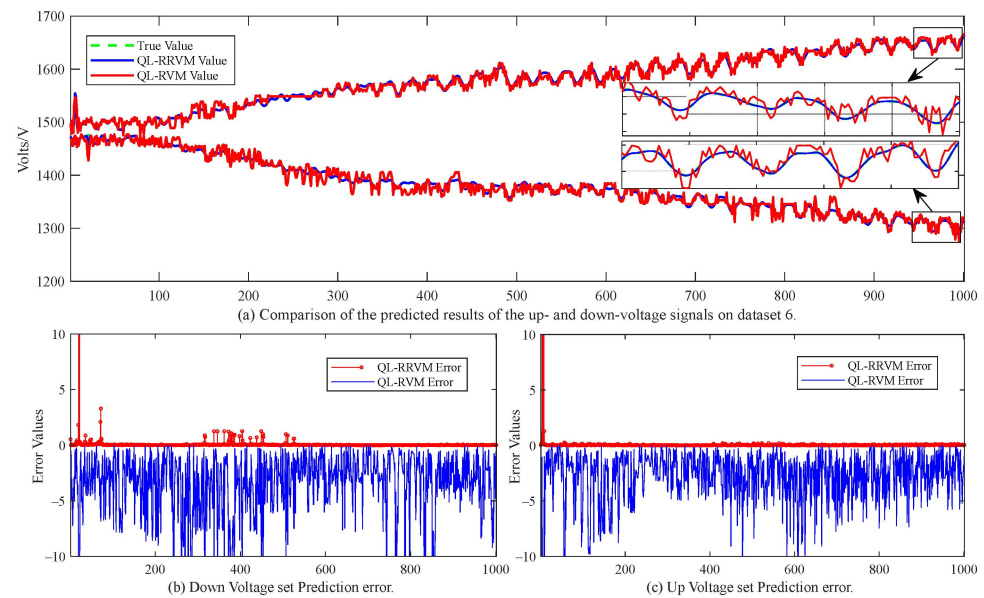


Figure 7. Comparison of the predicted results of the up- and down-voltage signals on dataset 6.

To validate the superiority of incorporating QL in the proposed QL-RRVM model, the QL algorithm component is removed, and the following comparative experiments are conducted. Figure 8 illustrates the degradation trend prediction experiment using the RRVM model and the RVM model. It can be observed that significant prediction errors occur at both the beginning and the end of the dataset. This is attributed to the unprocessed dataset lacking representativeness, which makes it difficult to make accurate predictions for other data.

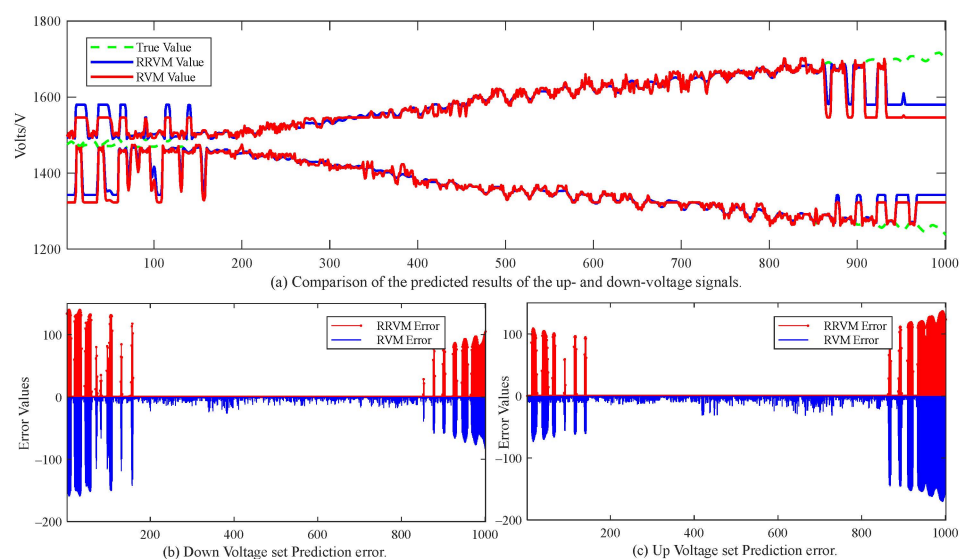


Figure 8. Model predictions after removing the QL algorithm.

In this study, four evaluation criteria—Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE) and Pearson Correlation

Coefficient—are employed to further assess the prediction performance. The criteria are given below.

$$MAE = \frac{1}{M} \sum_{i=1}^M |\tilde{y}_i - y_i| \quad (26)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\tilde{y}_i - y_i)^2} \quad (27)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{\tilde{y}_i - y_i}{y_i} \right| \times 100 \quad (28)$$

$$R = \frac{\sum_{i=1}^N (y_i - \bar{y}_*)(\tilde{y}_i - \bar{y})}{\sqrt{\sum_{i=1}^N (y_i - \bar{y}_*)^2 (\tilde{y}_i - \bar{y})^2}} \quad (29)$$

where N , $\tilde{y}_i$, y_i , $\bar{y}_*$ and $\bar{y}$ represent the number of samples, the true values, the predicted values, the true value means and the predicted value means, respectively.

Utilizing four evaluation metrics, trend predictions are conducted among multiple models, which are shown in Table 1. By comparing the data in the table, it is evident that the QL-RRVM model provides the best prediction for the degradation trend, followed by the RRVM model. This further demonstrates the effectiveness of the introduction of the QL algorithm and the robustness improvements made in this paper.

Table 1. Evaluation of the results of four models for predicting degradation trends on different datasets.

DATA	MODEL	MAE (V)	RMSE (V)	MAPE (%)	R
DATA SET 1 (Up and Down)	QL-RRVM	0.0109	0.0393	0.0008	0.99
		0.0576	0.751	0.0039	0.99
	QL-RVM	5.161	6.518	0.319	0.983
		6.749	8.599	0.501	0.969
	RRVM	15.790	41.714	0.968	0.818
		14.270	38.316	1.039	0.846
	RVM	24.317	50.758	1.483	0.729
		19.432	41.892	1.399	0.815
DATA SET 6 (Up and Down)	QL-RRVM	0.203	3.441	0.0137	0.997
		0.057	0.751	0.003	0.998
	QL-RVM	5.652	7.861	0.359	0.986
		6.967	9.041	0.504	0.982
	RRVM	1.417	10.835	0.095	0.974
		4.829	24.082	0.328	0.872
	RVM	8.051	12.581	0.513	0.964
		11.064	29.578	0.775	0.806

5.4. RUL Prediction

Finally, the predicted degradation trends obtained in Section 5.4 are utilized to forecast the RUL of the components. In this paper, the RUL of a component is linked to its corresponding degradation trend, with the final degradation moment defined as the failure state, i.e., $RUL = 0$. Figure 9 presents the training results of the RUL prediction for the components using the QL-RRVM model based on the degradation trends derived from Figure 7. It is observed that the model presented in this paper accurately predicts the trend of RUL changes.

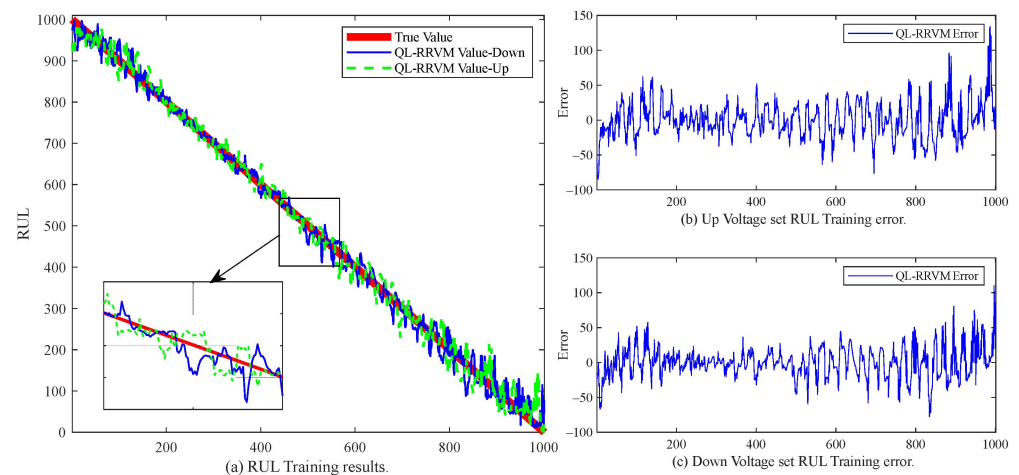


Figure 9. RUL training results.

Similarly, multiple comparative experiments are designed to verify the superiority of the model proposed in this paper. Figure 10 compares the RUL prediction performance of the RRVM and RVM models when the QL algorithm is incorporated. Figure 10a,b represent the prediction results of the two models under the lower- and upper-voltage data, respectively, while Figure 10c–f correspond to the respective prediction errors. From Figure 10, it can be seen that the RRVM model proposed in this paper provides more accurate RUL predictions compared to the RVM model, with less fluctuation.

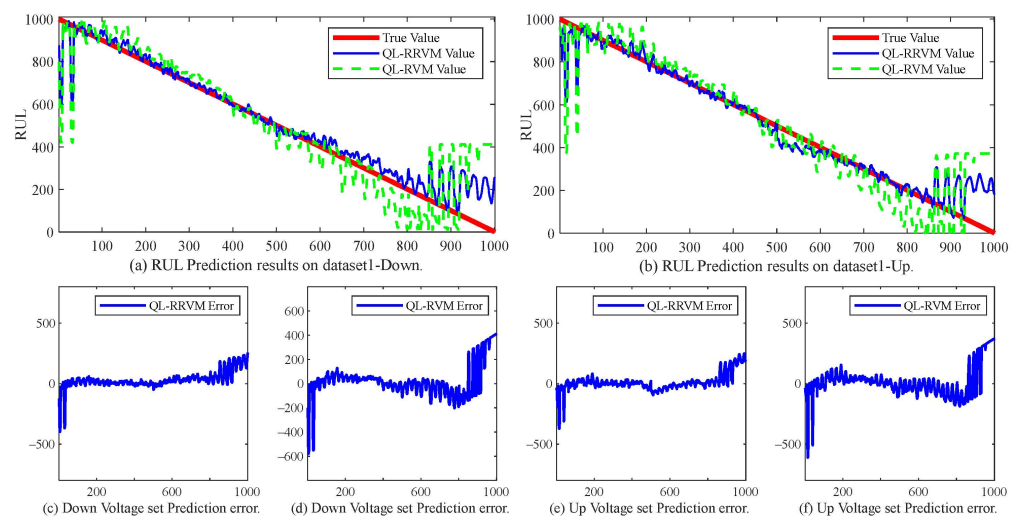


Figure 10. RUL prediction results with QL-RRVM and QL-RVM model.

Figure 11 presents the RUL predictions for the same dataset by four different models. Figure 11a–d correspond to the prediction results of the four models, respectively, with the bar charts in the figures representing the absolute values of the prediction errors. From Figure 11, one can visually discern the differences in the prediction results among the four models, where the QL-RRVM model yields the best results, and the RVM model the worst. Similarly, comparative experiments are conducted on multiple datasets, and the results are accurately assessed using three evaluation metrics, as shown in Table 2. The comparison of corresponding metrics between the QL-RRVM and QL-RVM models demonstrates the effectiveness of the robustness improvements made in this paper. The comparison of corresponding metrics between the QL-RRVM and RRVM models proves the efficacy of incorporating the QL algorithm in this study.

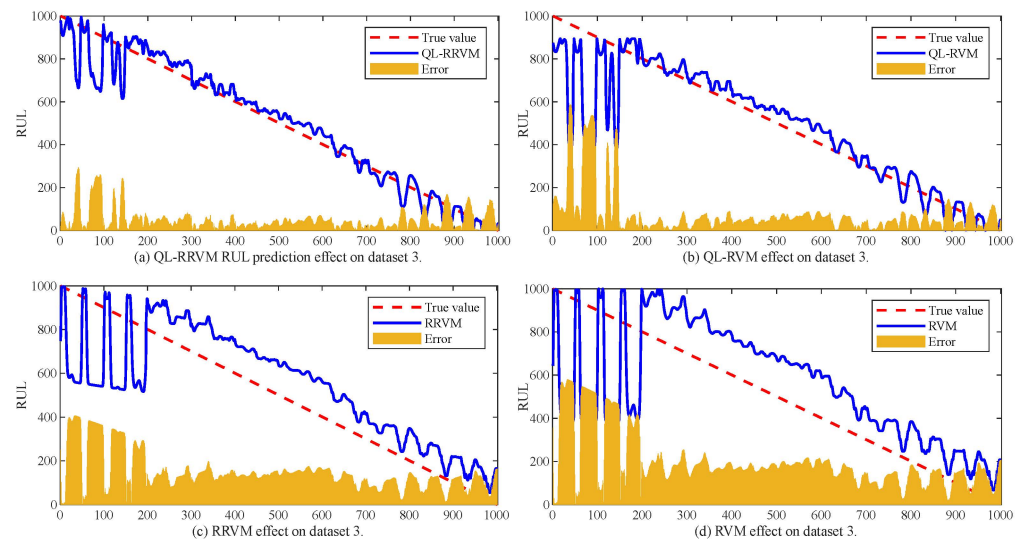


Figure 11. RUL prediction results with all models.

Table 2. Evaluation of RUL prediction results for four models on different datasets.

DATA	MODEL	MAE (RUL)	RMSE (RUL)	R
DATA SET 1	QL-RRVM	63.023	77.391	0.996
	QL-RVM	86.729	108.418	0.983
	RRVM	131.399	157.154	0.984
	RVM	110.987	147.024	0.907
DATA SET 3	QL-RRVM	53.386	75.552	0.967
	QL-RVM	78.646	133.706	0.890
	RRVM	149.002	172.939	0.819
	RVM	197.572	235.4965	0.654
DATA SET 6	QL-RRVM	44.581	55.7997	0.983
	QL-RVM	96.39	119.922	0.969
	RRVM	86.472	112.719	0.958
	RVM	110.987	147.024	0.907

6. Conclusions

This paper proposes a robust RVM model based on QL for RUL prediction. The robustness issue of the traditional RVM model is addressed by introducing a t -distribution; the QL algorithm is incorporated into the model, which processes the original signals by selecting different actions to obtain the features data. The article achieves RUL prediction based on the rolling prediction of degradation trends, enhancing the prediction interpretability.

The effectiveness and superiority of the proposed method are verified under identical data and operational conditions using four models: QL-RRVM, QL-RVM, RRVM, and RVM. As the corresponding RMSE values obtained are 75.552, 133.706, 172.939, and 235.4965, respectively, it indicates that the QL-RRVM model exhibits an approximate 67% improvement in prediction performance. Furthermore, it is evident that the present study significantly enhances feature extraction and robustness.

The proposed QL-RRVM model involves multiple nested loops and complex parameter optimization, which leads to substantial time consumption during training and prediction. In future research, the QL algorithm is planned to be optimized to address this issue. Additionally, within the RRVM module, the large number of kernel functions corresponding to each variable affects computational efficiency. Therefore, future focuses

include implementing pruning on the kernel functions to enhance the model's computational efficiency.

Author Contributions: Conceptualization, methodology and funding acquisition, X.W. (Xiuli Wang); data curation, writing, software and validation, Z.L.; formal analysis, supervision, writing—review and editing, X.W. (Xiuyi Wang); visualization and investigation, X.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China under Grant 62303414, China Postdoctoral Science Foundation under Grant 2023M741821, Natural Science Foundation of Zhejiang Province under Grant LQ23F030016, Zhejiang Province Postdoctoral Selected Foundation under Grant ZJ2023143.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: Author Xinyu Hu was employed by the State Grid Nantong Power Supply Company. 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

- Li, Y.; Xu, S.; Chen, H.; Jia, L.; Ma, K. A general degradation process of useful life analysis under unreliable signals for accelerated degradation testing. *IEEE Trans. Ind. Inform.* **2023**, *19*, 7742–7750. [\[CrossRef\]](#)
- Achouch, M.; Dimitrova, M.; Ziane, K.; Sattarpanah Karganroudi, S.; Dhouib, R.; Ibrahim, H.; Adda, M. On predictive maintenance in industry 4.0: Overview, models, and challenges. *Appl. Sci.* **2022**, *12*, 8081. [\[CrossRef\]](#)
- Ahmad, T.; Zhu, H.; Zhang, D.; Tariq, R.; Bassam, A.; Ullah, F.; AlGhamdi, A.S.; Alshamrani, S.S. Energetics systems and artificial intelligence: Applications of industry 4.0. *Energy Rep.* **2022**, *8*, 334–361. [\[CrossRef\]](#)
- Feng, K.; Ji, J.C.; Ni, Q.; Beer, M. A review of vibration-based gear wear monitoring and prediction techniques. *Mech. Syst. Signal Process.* **2023**, *182*, 109605. [\[CrossRef\]](#)
- Thelen, A.; Li, M.; Hu, C.; Bekyarova, E.; Kalinin, S.; Sanghadasa, M. Augmented model-based framework for battery remaining useful life prediction. *Appl. Energy* **2022**, *324*, 119624. [\[CrossRef\]](#)
- Li, Y.; Gao, H.; Chen, H.; Liu, C.; Yang, Z.; Zio, E. Accelerated degradation testing for lifetime analysis considering random effects and the influence of stress and measurement errors. *Reliab. Eng. Syst. Saf.* **2024**, *247*, 110101. [\[CrossRef\]](#)
- Jiang, X.; Xu, J.; He, Q.; Wang, C.; Jiang, L.; Xu, K.; Xiang, J. A study of the relationships between coal heterogeneous chemical structure and pyrolysis behaviors: Mechanism and predicting model. *Energy* **2023**, *282*, 128715. [\[CrossRef\]](#)
- Jorner, K.; Brinck, T.; Norrby, P.O.; Buttar, D. Machine learning meets mechanistic modelling for accurate prediction of experimental activation energies. *Chem. Sci.* **2021**, *12*, 1163–1175. [\[CrossRef\]](#)
- Li, N.; Gebraeel, N.; Lei, Y.; Fang, X.; Cai, X.; Yan, T. Remaining useful life prediction based on a multi-sensor data fusion model. *Reliab. Eng. Syst. Saf.* **2021**, *208*, 107249. [\[CrossRef\]](#)
- Chen, S.; Zhou, X.; Zhou, G.; Fan, C.; Ding, P.; Chen, Q. An online physical-based multiple linear regression model for building's hourly cooling load prediction. *Energy Build.* **2022**, *254*, 111574. [\[CrossRef\]](#)
- Ramirez, J.G.; Gore, W.L.; Johnston, G. New methods for modeling reliability using degradation data. *Stat. Data Anal. Data Min.* **2001**, *9*, 226–263.
- Hedrea, E.L.; Precup, R.E.; Bojan-Dragos, C.A. Results on tensor product-based model transformation of magnetic levitation systems. *Acta Polytech. Hung.* **2019**, *16*, 93–111. [\[CrossRef\]](#)
- Lei, Y.; Li, N.; Guo, L.; Li, N.; Yan, T.; Lin, J. Machinery health prognostics: A systematic review from data acquisition to RUL prediction. *Mech. Syst. Signal Process.* **2018**, *104*, 799–834. [\[CrossRef\]](#)
- Chen, H.; Luo, H.; Huang, B.; Jiang, B.; Kaynak, O. Transfer learning-motivated intelligent fault diagnosis designs: A survey, insights, and perspectives. *IEEE Trans. Neural Networks Learn. Syst.* **2023**, *35*, 2969–2983. [\[CrossRef\]](#)
- Ren, L.; Dong, J.; Wang, X.; Meng, Z.; Zhao, L.; Deen, M.J. A data-driven auto-CNN-LSTM prediction model for Lithium-Ion Battery Remaining useful Life. *IEEE Trans. Ind. Inform.* **2021**, *17*, 3478–3487. [\[CrossRef\]](#)
- Lin, Y.H.; Li, G.H. A Bayesian deep learning framework for RUL prediction incorporating uncertainty quantification and calibration. *IEEE Trans. Ind. Inform.* **2022**, *18*, 7274–7284. [\[CrossRef\]](#)
- Tipping, M.E. Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* **2001**, *1*, 211–244.
- Feng, Y.; Wang, Y.; Wang, J.W.; Li, H.X. Backstepping-Based distributed abnormality localization for linear parabolic distributed parameter systems. *Automatica* **2021**, *135*, 109930. [\[CrossRef\]](#)
- Li, Y.; Kaynak, O.; Jia, L.; Liu, C.; Wang, Y.; Zio, E. A Generalized Testing Model for Interval Lifetime Analysis Based on Mixed Wiener Accelerated Degradation Process. *IEEE Internet Things J.* **2024**; preprint. [\[CrossRef\]](#)
- Tipping, M.E.; Faul, A.C. Fast marginal likelihood maximisation for sparse Bayesian models. In Proceedings of the International Workshop on Artificial Intelligence and Statistics, PMLR, Key West, FL, USA, 3–6 January 2003; pp. 276–283.

21. Song, Y.; Liu, D.; Hou, Y.; Yu, J.; Peng, Y. Satellite lithium-ion battery remaining useful life estimation with an iterative updated RVM fused with the KF algorithm. *Chin. J. Aeronaut.* **2018**, *31*, 31–40. [CrossRef]
22. Sanz-Gorrachategui, I.; Pastor-Flores, P.; Pajovic, M.; Wang, Y.; Orlik, P.V.; Bernal-Ruiz, C.; Artal-Sevil, J.S. Remaining useful life estimation for LFP cells in second-life applications. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 2505810. [CrossRef]
23. Zheng, X.; Fang, H. An integrated unscented kalman filter and relevance vector regression approach for lithium-ion battery remaining useful life and short-term capacity prediction. *Reliab. Eng. Syst. Saf.* **2015**, *144*, 74–82. [CrossRef]
24. Wang, X.; Jiang, B.; Lu, N. Adaptive relevant vector machine based RUL prediction under uncertain conditions. *ISA Trans.* **2019**, *87*, 217–224. [CrossRef] [PubMed]
25. Wang, X.; Jiang, B.; Wu, S.; Lu, N.; Ding, S.X. Multivariate relevance vector regression based degradation modeling and remaining useful life prediction. *IEEE Trans. Ind. Electron.* **2021**, *69*, 9514–9523. [CrossRef]
26. Chen, H.; Liu, Z.; Alippi, C.; Huang, B.; Liu, D. Explainable intelligent fault diagnosis for nonlinear dynamic systems: From unsupervised to supervised learning. *IEEE Trans. Neural Networks Learn. Syst.* **2022**, *35*, 6166–6179. [CrossRef]
27. Chen, Z.; Shi, N.; Ji, Y.; Niu, M.; Wang, Y. Lithium-ion batteries remaining useful life prediction based on BLS-RVM. *Energy* **2021**, *234*, 121269. [CrossRef]
28. Liu, B.; Zhang, Y. Calibration of miniature air quality detector monitoring data with PCA-RVM-NAR combination model. *Sci. Rep.* **2022**, *12*, 9333. [CrossRef]
29. Li, D.H. The approach to get the decomposition of the Kronecker product of matrix. *J. Minjiang Univ.* **2007**. Available online: <https://api.semanticscholar.org/CorpusID:124306738> (accessed on 7 October 2024).
30. Haobo, J.; Guangyu, L.; Jin, X.; Jian, Y. Action candidate driven clipped double Q-Learning for discrete and continuous action tasks. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 5269–5279.
31. Ding, W.; Haiming, H.; Mingming, Z. Model-free optimal tracking design with evolving control strategies via Q-learning. *IEEE Trans. Circuits Syst. II Express Briefs* **2024**, *71*, 3373–3377.
32. Li, Y.; Fei, M.; Jia, L.; Lu, N.; Kaynak, O.; Zio, E. Novel Outlier-Robust Accelerated Degradation Testing Model and Lifetime Analysis Method Considering Time-Stress-Dependent Factors. *IEEE Trans. Ind. Inform.* **2024**, *20*, 9907–9917. [CrossRef]
33. Yang, C.; Gui, W.; Chen, Z.; Zhang, J.; Peng, T.; Yang, C.; Ding, S.X. Voltage difference residual-based open-circuit fault diagnosis approach for three-level converters in electric traction systems. *IEEE Trans. Power Electron.* **2019**, *35*, 3012–3028. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.