Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)
Edited by Maria Chiara Leva, Edoardo Patelli, Luca Podofillini, and Simon Wilson
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5183-4_S27-02-465-cd

Collaborative Kernel-based Nonlinear Degradation Modeling with Transfer Learning for Remaining Useful Life Prediction

Zhen Chen

Department of Industrial Engineering and Management, Shanghai Jiao Tong University, China; Department of Energy, Politecnico di Milano, Italy. E-mail: chenzhendr@sjtu.edu.cn

Lanxiang Liu

School of Electrical Engineering and Automation, Harbin Institute of Technology, China; Department of Energy, Politecnico di Milano, Italy. E-mail: lanxiang.liu@polimi.it

Enrico Zio

Centre for Research on Risk and Crises, Ecole des Mines, ParisTech, PSL University, France; Department of Energy, Politecnico di Milano, Italy. E-mail: enrico.zio@polimi.it

Ershun Pan

Department of Industrial Engineering and Management, Shanghai Jiao Tong University, China. E-mail: pes@sjtu.edu.cn

Abstract:
A novel nonlinear collaborative modeling method for remaining useful life (RUL) prediction is proposed. This method uses a kernel-based Wiener process (KWP) model, which formulates a nonlinear drift function with the weighted combination of kernel functions. Compared with the existing Wiener process models, this kind of modeling allows characterizing the nonlinearity of degradation more accurately and flexibly. To address the problem of error accumulation and lack of data in long-term prediction, a transfer learning scheme, based on the KWP models, is developed by leveraging multiple historical degradation trends from different units to collaboratively describe the degradation process of interest, even with limited data. The positive transfer learning is realized by introducing cross-correlations into the drift functions for obtaining more robust and accurate results than those obtained by constructing multiple individual models, one for each unit. The unknown model parameters are estimated by a spare estimation algorithm. Then, based on the KWP model, a close-form expression of the RUL distribution is derived for uncertainty quantification. An online framework is also proposed to iteratively predict the RUL. Finally, the proposed method is verified on lithium-ion battery datasets and compared to existing methods. The outcomes demonstrate the effectiveness and superiority of the proposed method for RUL prediction.

Keywords: Degradation modeling, nonlinearity, transfer learning, remaining useful life (RUL).

1. Introduction

Prognostics and health management (PHM) contributes to satisfy the reliability, safety and economical requirements of industrial components and systems. The prediction of the remaining useful life (RUL) of a component or system, i.e., the residual time for losing its intended function, is a most challenging task of PHM (Cheng *et al*. 2022). Accurate RUL predictions can help making effective and economical maintenance decisions.

Due to differing operating environments, unit-to-unit variability and imperfect inspection, most degradation processes show stochastic dynamics characteristics (Chen *et al*. 2020). Hence, stochastic process models are widely used in degradation modeling and RUL prediction (Ye *et al.* 2015). Wiener process, Gamma process, inverse Gaussian process, Lévy process and Cauchy process are the common stochastic processes in degradation analysis (Ye *et al.* 2015, Tseng *et al.* 2009, Wang and Xu 2010, Liu *et al.* 2017, Duan *et al*. 2021, Hong *et al.* 2022). Since the Wiener process is more suitable used to describe non-monotonous degradation paths than the other processes, it has attracted more

attention and been used for various systems (Si 2015, Chen *et al.* 2021).

In addition to stochasticity, nonlinearity is another key property of the degradation processes of components and systems. To model the nonlinearity, The state transformations (Gebraeel *et al.* 2005) or time-scale transformations (Chen et al. 2021) can be applied to the drift function of Wiener process to linearize the nonlinear degradation paths. Among the types of nonlinear Wiener process models, the power-law and exponential-law models have gained popularity (Son *et al*. 2013, Deng *et al*. 2016). However, these two models struggle to capture complicated nonlinear behaviours. Zhang *et al.* (2017) used a linear combination of simple functions to construct the drift rate as a general representation of non-linearity, which depended heavily on prior knowledge of the processes. Zhai and Ye (2017) developed a new adaptive Wiener process model that utilized a Brownian motion to describe the time-varying drift rate for RUL prediction. Li et al. (2017) proposed a generalized Wiener process degradation model with measurement errors to represent the nonlinear degradation paths. Wang et al. (2019) developed an improved Wiener process model with adaptive drift and diffusion for online RUL prediction. Zhang *et al.* (2021) combined a nonlinear Wiener process model with a relevance vector machine to depict the nonlinearity of degradation without priors. Zhang *et al.* (2022) studied a nonlinear Wiener process model with a random time-varying covariate for degradation modeling and RUL prediction. But these methods ignored some relationships in the degradation data that can significantly affect the modeling accuracy.

Specifically, temporal dependency and cross-correlation are the two typical relationships in the monitoring data (Zhang *et al.* 2019). The former indicates that the early degradation of a specific system can affect its advanced stages (Chen *et al.* 2019). The latter describes the statistical relationships between two systems which show some common features (Ding *et al.* 2022). Due to unrevealed failure mechanism and lack of degradation data, it is difficult to obtain accurate results by a single model constructed on the limited degradation data of a target sample. Chehade and Hussein (2020) used a multi-output Gaussian process to collaboratively cross-correlate degradation trends of different units for long-term prediction with limited data. However, they did not derive the distribution of the RUL to account for the uncertainty in the prediction.

This paper proposes a transfer learning (TL) - based collaborative model with kernel-based Wiener process model for the RUL prediction. The main contributions are as follows:

- Considering both stochasticity and non-linearity, an improved Wiener process model, whose drift function is formulated by kernel functions, is used to obtain flexible and accurate degradation modeling.
- As an extension to the independent models, a collaborative modeling framework with TL-based KWP models is developed for RUL prediction by leveraging historical data.
- Both the temporal dependency and the cross-correlations in the degradation data are captured to address the problem of lack of data at early stages and error accumulation for long-term prediction.
- The distribution of the real-time RUL is predicted, for the uncertainty quantification.

The remainder of this paper is organized as follows. Section 2 develops the Kernel-based Wiener process model. In Section 3, we propose a collaborative modeling method for RUL prediction. In Section 4, a numerical example is presented. Conclusions are made in Section 5.

2. Theoretical Background

The degradation paths of industrial components or systems are assumed to follow the Wiener process (WP) model, i.e.,

$$x(t) = x_0 + \eta(t;\boldsymbol{\omega}) + \sigma_B B(t), \qquad (1)$$

where $\eta(t;\boldsymbol{\omega})$ is the drift function, σ_B is the diffusion parameter, and $B(t)$ is the standard Brownian motion. The initial value x_0 is assumed to be 0. Let $\Delta x_j = x_j - x_{j-1}$, where x_j is the degradation observation at time t_j; the drift increment $\Lambda(t_j;\boldsymbol{\omega}) = \eta(t_j;\boldsymbol{\omega}) - \eta(t_{j-1};\boldsymbol{\omega})$, $t_0 = 0$.

To describe the time-varying dynamics of the stochastic degradation process, $\eta(t;\boldsymbol{\omega})$ is usually assumed to be nonlinear time-dependent. The commonly used formulations for $\eta(t;\boldsymbol{\omega})$ are the power-law or the exponential-law form. Though they can describe most of the typical nonlinear degradation trajectories, it might be difficult to

deal with some complex cases in time-varying nonlinearity. To improve the generality of the degradation modeling without prior, an alternative idea of introducing the kernel function into the drift increment is proposed as

$$\Lambda_j(\boldsymbol{\omega}) = \Lambda(t_j;\boldsymbol{\omega}) = \sum_{i=1}^{m} \omega_i \kappa(t_i, t_j) + \omega_0, \quad (2)$$

where ω_i is the weight of the kernel function $\kappa(t_i, t_j)$, ω_0 is the bias, m is the sample size and $\boldsymbol{\omega} = [\omega_0, \omega_1, \mathrm{L}\ , \omega_m]^{\mathrm{T}}$. The scaled Gaussian kernel is usually selected as the basis function in nonlinear prediction. Without loss of generality, the drift rate coefficient can be derived by taking the first partial derivative of $\Lambda(t_j;\boldsymbol{\omega})$ with t_j as

$$\mu(t_j;\boldsymbol{\omega}) = d\Lambda(t_j;\boldsymbol{\omega})/dt_j . \quad (3)$$

The rate of change of the degradation measurement can be captured more effectively by (2) for short-term prediction. It can avoid scaling issues, and be more robust for non-smooth degradation paths. Then, a generalized nonlinear model (Kernel-based Wiener Process, KWP) is formulated based on (2) and (3).

Considering that some elements in $\boldsymbol{\omega}$ could be extremely small, the sparsity optimization strategy is adopted to avoid the overfitting in the maximum likelihood estimation (MLE). The non-zero weights are penalized to reduce the variance of the solutions and computational burden. The l_1-*norm* regularization is selected to realize the sparsity of $\boldsymbol{\omega}$. Assume the degradation process is observed at equal inspection intervals, $\tau = t_j - t_{j-1}$. The log-likelihood of the degradation increment dataset $\Delta\boldsymbol{x} = \{\Delta x_j\}_{j=1}^{m}$ with a sparsity term is written as

$$\ln\left(p(\Delta\boldsymbol{x}|\boldsymbol{\omega},\sigma_B)\right) = -\frac{m}{2}\ln(2\pi\sigma_B^2\tau) - \left\|\Delta\boldsymbol{x} - \boldsymbol{\Omega}\boldsymbol{\omega}\right\|^2 / (2\sigma_B^2\tau) - \lambda\sum_{i=0}^{m}|\omega_i| , \quad (4)$$

where $\Delta\boldsymbol{x} = [\Delta x_1, \mathrm{L}\ , \Delta x_m]^{\mathrm{T}}$, $\boldsymbol{\Omega}$ is an $m\times(m+1)$ matrix with $\boldsymbol{\Omega} = [\phi(t_1), \mathrm{L}\ , \phi(t_m)]^{\mathrm{T}}$, and $\varphi(t_j) = [1, \kappa(t_1,t_j), \mathrm{L}\ , \kappa(t_m,t_j)]$. $\lambda > 0$ is the penalty coefficient with a large value. Then, the measurements corresponding to the nonzero weights are responsible for representing the temporal dependency in the degradation paths.

Since it is difficult to estimate the values of $\boldsymbol{\omega}$ and σ_B by directly maximizing (4) in closed-form, iterative estimation or evolutionary algorithms can be used to obtain the optimal MLEs, which are denoted by $\boldsymbol{\omega}^*$ and σ_B^*. Then, the prediction of x_{m+1} at the next time point t_{m+1} is Gaussian distributed conditional on the estimates, i.e.,

$$\hat{x}_{m+1}|\Delta\boldsymbol{x},\boldsymbol{\omega},\sigma_B : \ \mathrm{N}\left(x_m + \boldsymbol{\omega}^{*\mathrm{T}}\boldsymbol{\phi}(t_{m+1}), \sigma_B^{2*}\tau\right). \quad (5)$$

Generally, the system lifetime T is defined as the first passage time (FPT) of $x(t)$ hitting a critical threshold ξ, i.e.,

$$T = \inf\{t : x(t) \ge \xi | x_0 < \xi\}, \quad (6)$$

According to (6), the RUL at time t_j can be defined as the residual time for the degradation path crossing ξ from x_j, i.e.,

$$R_j = \inf\{r : x(t_j + r) \ge \xi | x_j < \xi\}. \quad (7)$$

Given x_j, the PDF for the RUL at time t_j is

$$f_{R_j}(r|\boldsymbol{\omega},\sigma_B) \cong \frac{1}{\sqrt{2\pi r}}\exp\{-S_j^2(r)/(2r)\} \times\left[S_j(r)/r + \Lambda_j(r;\boldsymbol{\omega})/(\sigma_B r)\right] , \quad (8)$$

where $S_j(r) = (\xi - x_j - \Lambda_j(r;\boldsymbol{\omega}))/\sigma_B$ and $\Lambda_j(r;\boldsymbol{\omega}) = \eta(t_j + r;\boldsymbol{\omega}) - \eta(t_j;\boldsymbol{\omega})$.

3. The Proposed Collaborative Modeling with Transfer Learning for RUL Prediction

Note that Zhang *et al.* (2021) used a Gaussian basis function as part of RVM to represent the degradation drift increment for RUL estimation, which is the same as the proposed KWP model. However, the main limitation of KWP is that the prediction errors can accumulate in the long-term task, where the degradation mechanism is not explicit and the future data is unavailable. Moreover, due to unit-to-unit variability, even if the historical data from other identical units are used to estimate the parameters of KWP, the predictions by the determined KWP could deviate from the actual values. Considering that identical units should share some common features and are often cross-correlated, a novel TL-based collaboratively degradation model is developed to leverage cross-correlations between the degradation measurements of different units, for predicting the RUL more accurately with limited data.

3.1. TL-based Collaboratively Modeling

Unlike developing an independent KWP model for each unit, the TL-based collaborative KWP model

combines the degradation data from all units to capture both cross-correlations between multiple units and temporal dependency among different observations. The TL will transfer the shared knowledge between the target unit with limited data and the unit with complete historical data as much as possible, if the two units are correlated. Otherwise, the negative TL need to be avoided if the two units are nearly uncorrelated.

Only the selection of the kernel functions in the drift term will be used to account for the cross-correlations between different units. Especially at the early degradation stage of the target unit with insufficient data, the collaborative model allows extrapolating the future degradation path by making full use of the complete historical data from other units. The framework of the TL-based collaborative model is shown in Fig. 1.

The kernel-based drift increment of sample l cross-correlated with another unit is formulated as

$$\Lambda^*\left(t_j^l;\boldsymbol{\omega}\right)=\sum_{i=1}^{m_k}\omega_i\kappa^*\left(t_i^k,t_j^l\right)+\omega_0, \tag{9}$$

where k and l denote the reference numbers for the samples, and the kernel function is given by

$$\kappa^*\left(t_i^k,t_j^l\right)=\beta_F^2\exp\left\{-\frac{1}{2}\left[\left(t_i-t_j\right)/\beta_L\right]^2\right\}\times\exp\left\{-\frac{1}{2}\left[\left(1-\delta\left(k,l\right)\right)/\beta_{k,l}\right]^2\right\}, \tag{10}$$

where $\delta(k, l)$ is the Dirichlet function whose value equals to 1 if $k = l$ and 0 otherwise. For highly similar units k and l, the hyperparameter $\beta_{k,l}$ should be large when the positive TL occurs. On the contrary, for nearly uncorrelated units, $\beta_{k,l}$ is expected to be small and $\kappa^*\left(t_i^k,t_j^l\right)\to 0$.

Let $\boldsymbol{\beta}=\left\{\beta_F,\beta_L,\beta_{k,l}\right\}$. Given K samples with the degradation increment dataset $\left\{t_j^k,\Delta x_j^k\right\}_{j,k}$, the log-likelihood with l_1-*norm* regularization of $\boldsymbol{\Delta X}$ is expressed by

$$\ln\left(p\left(\Delta\mathbf{X}|\boldsymbol{\omega},\boldsymbol{\beta},\sigma_B\right)\right)=-\lambda\sum_{i=0}^{m}|\omega_i|-\sum_{l=1}^{K}\sum_{k=1}^{K}\left[\frac{m}{2}\ln\left(2\pi\sigma_B^2\tau\right)+\left\|\Delta\boldsymbol{x}^l-\boldsymbol{\Omega}^{l,k}\boldsymbol{\omega}\right\|^2/\left(2\sigma_B^2\tau\right)\right], \tag{11}$$

where $\Delta\mathbf{X}=\left\{\Delta\boldsymbol{x}^l\right\}_{l=1}^{K}$, $\boldsymbol{\Omega}^{l,k}=\left[\boldsymbol{\phi}\left(t_1^{l,k}\right),\mathrm{L},\boldsymbol{\phi}\left(t_m^{l,k}\right)\right]^{\mathrm{T}}$, and $\varphi\left(t_j^{l,k}\right)=\left[1,\kappa^*\left(t_1^k,t_j^l\right),\mathrm{L},\kappa^*\left(t_m^k,t_j^l\right)\right]$. The MLEs of $\boldsymbol{\omega}$, $\boldsymbol{\beta}$ and σ_B can be obtained by maximizing (11), and are denoted by $\boldsymbol{\omega}^*$, $\boldsymbol{\beta}^*$ and σ_B^* for specifically. After the parameters are determined, the trained TL-based collaborative KWP model is used for real-time RUL prediction.

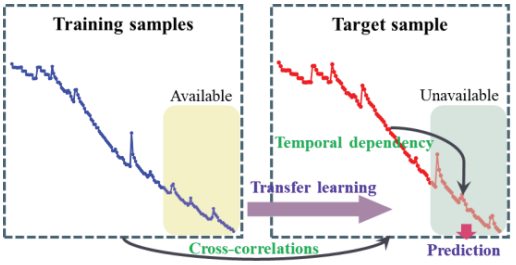


Fig. 1. Schematic of Transfer learning-based collaborative modeling for multiple samples.

To illustrate the necessity of transfer learning compared with Bayesian updating for individual degradation parameters, two shortcomings of the latter are as follows: firstly, if only the degradation data of the target unit is used, many useful information from other identical historical units is neglected; secondly, if both the degradation from the target unit and other identical units are used, the negative or useless information could be introduced into the parameter estimation as the other units' data are adopted for estimating the initial values in the Bayesian updating. Note that, the transfer learning can address these two problems well. The model parameters are estimated based on the degradation data up to date of the target unit and the whole degradation process data of other units. Although this is not an iterative procedure, the parameters' estimates are still updated step by step.

3.2. RUL Prediction

Based on the proposed degradation model, the RUL for the target unit at a given inspection timepoint is predicted in this subsection.

Suppose that the available degradation measurements at time t_1,L,t_m of the target unit are y_1,L,y_m. There are K training samples with complete degradation data observed at time t_1^k,L,t_m^k, i.e., $\mathbf{X}_{1:n}^{1:K}=\left\{x_1^k,\mathrm{L},x_m^k,\mathrm{L},x_n^k\right\}_{k=1}^{K}$, $m<n$. Then, the synthetic dataset $\left\{\mathbf{X}_{1:m}^{1:K},y_{1:m}\right\}$ is used to calculate the MLEs of the unknown parameters of the proposed model by using (11). The MLEs

are denoted by $\boldsymbol{\omega}^*$, $\boldsymbol{\beta}^*$ and σ_B^*, where the hyper-parameter $\boldsymbol{\beta}=\{\beta_F,\beta_L,\beta_1,\mathrm{L}\ ,\beta_K\}$. β_k, k = 1, …, K, represents the cross-correlation between the target unit and the k-th training sample.

When only considering the historical data $y_{1:m}$ of the target unit itself, the predicted mean of the degradation measurement involving the temporal dependency can be given by

$$\mathrm{E}\left(\hat{y}_{m+1,Self}\left|\mathbf{X}_{1:m+1}^{k},y_{1:m}\right.\right)=y_m+\boldsymbol{\omega}^{*\mathrm{T}}\boldsymbol{\phi}\left(t_{m+1}\right). \quad (12)$$

When only considering the transfer learning between the target unit and the k-th sample, the mean of the degradation measurement at t_{m+1} for the target unit can be predicted by

$$\mathrm{E}\left(\hat{y}_{m+1,TL}^{k}\left|\mathbf{X}_{1:m+1}^{k},y_{1:m}\right.\right)=y_m+\boldsymbol{\omega}^{*\mathrm{T}}\boldsymbol{\phi}^{k}\left(t_{m+1}\right), \quad (13)$$

where $\boldsymbol{\phi}^{k}\left(t_{m+1}\right)=\left[1,\kappa^*\left(t_1^k,t_{m+1}\right),\mathrm{L}\ ,\kappa^*\left(t_{m_k}^k,t_{m+1}\right)\right]$.

Then, considering all the training samples and the target unit itself, the prediction of y_{m+1} can be approximated by leveraging the cross-correlations and auto-correlation as follows

$$\hat{y}_{m+1}:\ \mathrm{N}\left(\frac{\hat{y}_{m+1,Self}+\sum_{k=1}^{K}\hat{y}_{m+1,TL}^{k}}{K+1},\sigma_B^{2*}\tau\right). \quad (14)$$

Note that only the mean is assumed to be correlated and the variance completely depends on the temporal dependency.

Assume that $r\cong g\tau$ and g is an integer. Given $y_1,\mathrm{L}\ ,y_m$, the PDF of the RUL for the proposed TL-based collaborative KWP model at time t_m can be given by

$$f_{R_m}\left(r\left|\boldsymbol{\omega}^*,\sigma_B^*\right.\right)\cong\frac{1}{\sqrt{2\pi r}}\exp\left[-S_m^2\left(r\right)/\left(2r\right)\right]\times\left[S_m\left(r\right)/r+\Lambda\left(t_m+r;\boldsymbol{\omega}^*\right)/\left(\sigma_B^*\tau\right)\right] \quad (15)$$

where $S_m\left(r\right)=\left(\xi-y_m-\bar{\Lambda}_m\left(r\right)\right)/\sigma_B^*$,

$$\bar{\Lambda}_m\left(r\right)=\left\{\sum_{k=1}^{K}\left[\sum_{i=1}^{m}\omega_i^*\sum_{v=1}^{g}\kappa\left(t_i^k,t_m+v\tau\right)+g\omega_0\right]+\sum_{i=1}^{m}\omega_i^*\sum_{v=1}^{g}\kappa\left(t_i,t_m+v\tau\right)+g\omega_0\right\}\Big/\left(K+1\right),$$

and $\Lambda\left(t_m+r;\boldsymbol{\omega}^*\right)=\sum_{i=1}^{m}\omega_i^*\kappa\left(t_i,t_m+r\right)+\omega_0$.

The real-time RUL prediction algorithm is summarized as Algorithm 1. Specifically, the flowchart of RUL prediction is shown in Fig. 2.

Algorithm 1: Real-time RUL prediction.

Input: Historical degradation data of K training samples $\mathbf{X}_{1:n}^{1:K}$; the degradation data of the target unit up to time t_m, $y_1,\mathrm{L}\ ,y_m$

Output: The PDF of the degradation value y_{m+1} at t_{m+1}; the PDF of the RUL at t_m, $f_{R_m}\left(r\right)$.

1: Obtain the degradation increment dataset $\Delta\mathbf{X}$.
2: Calculate the MLEs of $\boldsymbol{\omega}$, $\boldsymbol{\beta}$ and σ_B by (11).
3: Predict the degradation measurement $\hat{y}_{m+1}$ of the target unit at time t_{m+1} by (14).
4: Calculate the PDF of the RUL $f_{R_m}\left(r\left|\boldsymbol{\omega}^*,\sigma_B^*\right.\right)$ at time t_m by (15).
5: Set $m = m + 1$ and collect the new degradation measurement of the target unit at time t_{m+1}.
6: Go back to step 1.

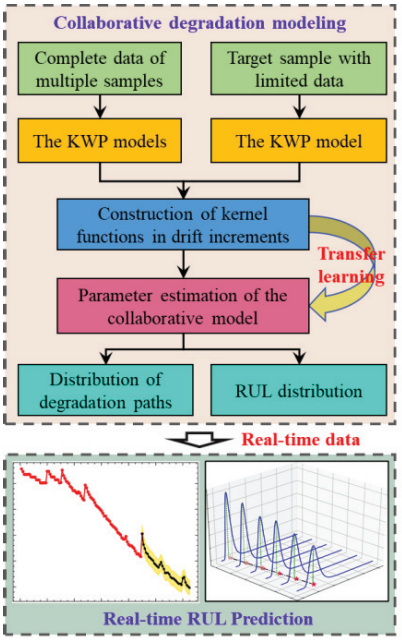


Fig. 2. Flowchart for the RUL prediction by the TL-based collaborative KWP model.

4. Illustrative Example

The lithium-ion battery datasets from the NASA Ames Prognostics Center of Excellence (Saha and Goebel 2007) are used to validate the proposed TL-based collaborative KWP model. The state parameters of batteries, such as voltage, current, temperature and capacity, were observed for different running cycles during charge-discharge experiments performed on a test bed at room

temperature. Repeated charge and discharge cycles resulted in performance degradation of the batteries. The battery is defined failed when its capacity declines to below 1.4 Ah. The datasets of three battery cells (B05, B06, B07) are analysed.

Fig. 3 shows the results of the correlation analysis of the capacity data. Not only the temporal dependency exists in the capacity degradation, but also the cross-correlation between two batteries is significant. Hence, it is necessary to consider both in the degradation modeling.

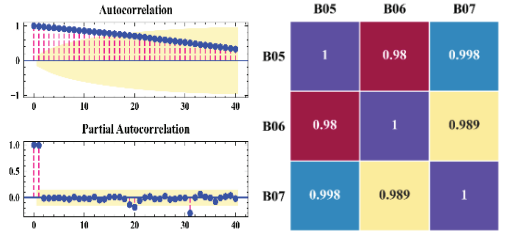


Fig. 3. Illustration of the correlations in degradation data. (a) Autocorrelation in the capacity data of B05. (b) Cross-correlations between the capacity data of two different batteries.

Fig. 4 shows the multi-step prediction results obtained by the proposed model. In Fig. 4(a), the capacity observations for the first 168 cycles for B06 and B07 are available, and only limited observations for B05 are given. In Fig. 4(b), the capacity data for the first 91 cycles of B06 are given and those for the first 168 cycles of B05 and B07 are known. It can be clearly seen from Fig. 4(a) that the TL-based collaborative KWP model is able to transfer knowledge from B06 and B07 to B05. This verifies that there is high cross-correlation between B05 and the other two batteries. Similarly, we can see from Fig. 4(b) the excellent predictability of the proposed model for the future degradation trend of B06. In addition, the 95% confidence intervals of prediction are presented in shaded yellow. Almost all the true capacity data fall within the intervals. It can be asserted that the proposed model can yield a convincing uncertainty quantification in long-term prognostics tasks, which is important in practical applications.

Next, the proposed model is compared to other models: (i) Artificial neural network (ANN) with one hidden layer (10 neurons), (ii) independent KWP (IKWP) for each sample, (iii) simple KWP (SKWP) by considering all the samples, which ignores the cross-correlations; (iv) a linear WP

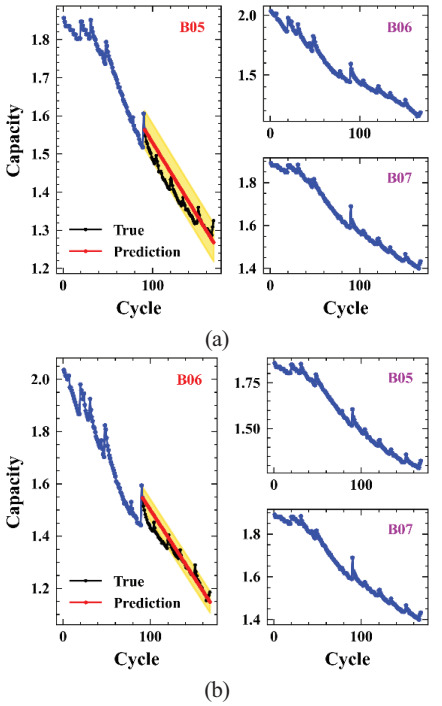


Fig. 4. Prediction results for the battery capacity by the proposed model. (a) Predictions for the capacity data of B05. (b) Predictions for the capacity data of B06.

(LWP) with the drift function $\eta(t;\omega)=\omega t$ for all samples; (v) an improved WP (IWP) model with the power-law drift function for all samples. Two evaluation metrics, Mean-absolute Error (MAE), and Mean Squared Error (MSE) are calculated for the predictions of these four models based on the scenarios as shown in Fig. 4. The corresponding results are presented in Table 1, where the TL-based collaborative KWP model is abbreviated as CKWP.

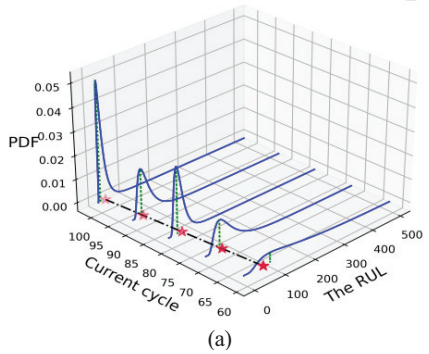
It can be seen from Table 1 that the proposed model which the lowest MAE and MSE values. This is because both the temporal dependency and cross-correlations are considered. The IKWP model uses only observations from the sample of interest and leads to poor predictions with limited data. The SKWP model assumes that the degradation data from all the samples are identical and ignores the cross-correlations. By comparing the MAE and MSE values of IKWP and SKWP, we can conclude that considering data available from different samples significantly improve the prediction accuracy. From the comparative results

of LWP, IWP and CKWP, we can see that the Wiener process with a power-law term cannot capture the degradation nonlinearity well, while the CKWP with kernel-based function represent the nonlinear degradation paths better.

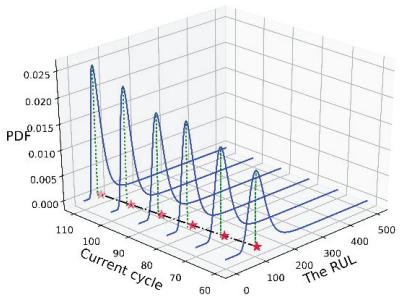
Table 1. The comparative results of different models based on the battery datasets.

Model	MAE	MSE
ANN	1.056×10^{-1}	1.399×10^{-2}
IKWP	5.188×10^{-2}	3.918×10^{-3}
SKWP	4.142×10^{-2}	3.049×10^{-3}
LWP	1.204×10^{-1}	2.295×10^{-2}
IWP	1.204×10^{-1}	2.327×10^{-2}
CKWP	$\mathbf{2.735\times10^{-2}}$	$\mathbf{1.127\times10^{-3}}$

The results are based on the scenarios in Fig. 4.



(a)



(b)

Fig. 5. Predicted PDFs of the RUL by the proposed model for (a) B05, (b) B06.

The effectiveness of the real-time RUL prediction algorithm is also validated here. Four different starting timepoints of available capacity data for the RUL prediction are selected, i.e., 61 cycles, 71 cycles, 81 cycles, 91 cycles and 101 cycles. The predicted PDFs of the RUL distribution for B05 and B06 are shown in Fig. 5. The true RUL values are marked as asterisks. It can be observed that the PDF curves cover the true RULs better and better over time, whose peaks get narrower and higher. That is, the prediction results are robust and accurate with different quantity of available data, and the uncertainty declines with more available data as expected. It further indicates that the proposed model can perform stably, and provide effective and accurate predicted RULs in the task of long-term prediction, as the inspection keeps rolling over time.

5. Conclusions

This paper proposed a collaborative kernel-based Wiener process model for RUL prediction. The cross-correlations between different samples and the temporal dependency in the specific sample of interest were captured by introducing a transfer learning term into the kernel functions of the drift increments. To address the problem of error accumulation and lack of data in long-term prediction, multiple historical degradation data were used. Then, the distributions of the RUL of the target unit with limited data can be predicted accurately. An application to lithium-ion battery datasets was conducted to illustrate the effectiveness of the proposed method.

Some future work is worth pursuing. The Bayesian updating for real-time iteratively estimating the model parameter can be introduced in the transfer learning to improve the computational efficiency.

Acknowledgement

This work was supported in part by National Natural Science Foundation of China under Grant 72001138 and Grant 72071127, China Postdoctoral Science Foundation under Grant 2019M661532, and in part by the National Key Research and Development Program of China under Grant 2020YFB1711100.

References

Y. Cheng, K. Hu, J. Wu, H. Zhu and X. Shao (2022). Auto-encoder quasi-recurrent neural networks for remaining useful life prediction of engineering systems. *IEEE/ASME Transactions on Mechatronics*, doi: 10.1109/TMECH.2021.3079729.

Z. Chen, T. Xia, Y. Li, E. Pan (2020). Tweedie exponential dispersion processes for degradation modeling, prognostic, and accelerated degradation test planning. *IEEE Transactions on Reliability*, 69(3), 887-902.

Z. S. Ye and M. Xie (2015). Stochastic modelling and analysis of degradation for highly reliable products. *Applied Stochastic Models in Business and Industry*, 31(1), 16–32, 2015.

Z. S. Ye, N. Chen, and Y. Shen (2015). A new class of Wiener process models for degradation analysis. *Reliability Engineering & System Safety*, 139, 58–67.

S. T. Tseng, N. Balakrishnan, and C. C. Tsai (2009) Optimal step-stress accelerated degradation test plan for gamma degradation processes. *IEEE Transactions on Reliability,* 58(4), 611-618.

X. Wang and D. Xu (2010) An inverse Gaussian process model for degradation data. Technometrics, 52(2), 188–197.

L. Liu, X. Y. Li, E. Zio, R. Kang, and T. M. Jiang (2017). Model uncertainty in accelerated degradation testing analysis. *IEEE Transactions on Reliability*, 66(3), 603–615.

S. Duan, W. Song, E. Zio, C. Cattani, and M. Li (2021). Product technical life prediction based on multi-modes and fractional Lévy stable motion. *Mechanical Systems and Signal Processing*, 161, 107974.

G. Hong, W. Song, Y. Gao, E. Zio, and A. Kudreyko, (2022). An iterative model of the generalized Cauchy process for predicting the remaining useful life of lithium-ion batteries. *Measurement*, 187, 110269.

X. S. Si (Si 2015). An adaptive prognostic approach via nonlinear degradation modeling: Application to battery data. *IEEE Transactions on Industrial Electronics*, 62(8), 5082–5096.

Z. Chen, T. Xia, D. Zhou, E. Pan. (2021). A Health index construction framework for prognostics based on feature fusion and constrained optimization. *IEEE Transactions on Instrumentation and Measurement*, 70, 1-15.

N. Z. Gebraeel, M. A. Lawley, R. Li, and J. K. Ryan (2005). Residual-life distributions from component degradation signals: A Bayesian approach. *IIE Transactions,* 37(6), 543–557.

Z. Chen, T. Xia, Y. Li, E. Pan (2021). A hybrid prognostic method based on gated recurrent unit network and an adaptive Wiener process model considering measurement errors. *Mechanical Systems and Signal Processing*, 158, 107785.

K. Le Son, M. Fouladirad, A. Barros, E. Levrat, and B. Iung (2013). Remaining useful life estimation based on stochastic deterioration models: A comparative study. *Reliability Engineering & System Safety*, 112, 165–175.

Y. Deng, A. Barros, and A. Grall (2016). Degradation modeling based on a time-dependent ornstein-uhlenbeck process and residual useful lifetime estimation. *IEEE Transactions on Reliability*, 65(1), 126–140.

Z. Zhang, C. Hu, X. Si, J. Zhang, and J. Zheng (2017). Stochastic degradation process modeling and remaining useful life estimation with flexible random-effects. *Journal of the Franklin Institute*, 354(6), 2477–2499.

Q. Zhai, Z. S. Ye (2017). RUL prediction of deteriorating products using an adaptive Wiener process model. *IEEE Transactions on Industrial Informatics*, 13(6): 2911-2921.

J. Li, Z. Wang, Y. Zhang, H. Fu, C. Liu, and S. Krishnaswamy (2017). Degradation data analysis based on a generalized Wiener process subject to measurement error. *Mechanical Systems and Signal Processing*, 94, 57-72.

H. Wang, X. Ma, and Y. Zhao (2019). An improved Wiener process model with adaptive drift and diffusion for online remaining useful life prediction. *Mechanical systems and signal processing*, 127, 370-387.

Y. Zhang, Y. Yang, H. Li, X. Xiu and W. Liu (2021). A data-driven modeling method for stochastic nonlinear degradation process with application to RUL estimation. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, doi: 10.1109/TSMC.2021.3073052.

S. Zhang, Q. Zhai, X. Shi and X. Liu (2022). A Wiener process model with dynamic covariate for degradation modeling and remaining useful life prediction. *IEEE Transactions on Reliability*, doi: 10.1109/TR.2022.3159273.

Y. Zhang, Z. Y. Dong, W., Kong, K. Meng (2019). A composite anomaly detection system for data-driven power plant condition monitoring. *IEEE Transactions on Industrial Informatics*, 16(7), 4390-4402.

Z. Chen, Y. Li, T. Xia, and E. Pan (2019). Hidden Markov model with auto-correlated observations for remaining useful life prediction and optimal maintenance policy. *Reliability Engineering & System Safety*, 184, 123-136.

Y. Ding, P. Ding, X. Zhao, Y. Cao and M. Jia (2022). Transfer learning for remaining useful life prediction across operating conditions based on multisource domain adaptation. *IEEE/ASME Transactions on Mechatronics*, doi: 10.1109/TMECH.2022.3147534.

A. A. Chehade, A. A. Hussein (2020). A collaborative Gaussian process regression model for transfer learning of capacity trends between li-ion battery cells. *IEEE Transactions on Vehicular Technology*, 69(9), 9542-9552.

B. Saha and K. Goebel (2007). Battery Data Set. Moffett Field, CA, USA: NASA Ames Research Center. [Online]. Available: http://ti.arc.nasa.gov/project/prognostic-data-repository