A Novel Pattern-driven Stochastic Process for End-of-Life Forecasting

Zihan Zhang*, Chi-Guhn Lee*

** Center for Maintenance Optimization and Reliability Engineering (C-MORE), Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, M5S 2E4, Canada {zhzhang, cglee}@mie.utoronto.ca*

Abstract. In recent decades, there has been significant growth in the development of rechargeable battery-powered devices, such as electric vehicles, leading to a huge demand for batteries with high reliability and quality. End of life (EoL) is a critical indicator of battery health and can be estimated by either adaptive stochastic processes or advanced machine learning techniques. However, these approaches follow the degradation path that can be modelled as simple mathematical form such as linear or exponential function and lack interpretability due to its black-box nature. To address these shortfalls, an GRU-driven degradation process is proposed to learn complex battery degradation patterns, in which degradation progression is controlled by a recursive Gaussian distribution with its mean learnt from an GRU-driven degradation pattern. Due to the non-Markovian state transitions, a joint-learning sampling-based expectation maximization algorithm is developed to estimate model parameters based on historical observations. To validate the superiority of the proposed methods, a case study of battery data was implemented. The results show a better performance with respect to EoL accuracy than that achieved with traditional methods.

1. Introduction

As rechargeable batteries have been widely used in electric vehicles (EVs) and hybrid EVs, the reliability of battery has become increasingly critical to the safe and reliable operation of a vehicle. This has led to increasing efforts of enhancing battery health management [1]. The prediction of state-of-health (SoH), a critical indicator of battery state, has been recognized as the central metric in battery health management and prognostic methods for SoH can be classified as model-based and data-driven forecasting methods.

Model-based forecasting methods include electrochemical model, equivalent circuit model, empirical model and stochastic model [2-3]. The former two models attempt to utilize the physical or chemical dynamics of the battery properties by exploring complex aging mechanisms [4]. However, the high cost of measuring the internal degradation parameters makes their practical application prohibitively expensive [5]. In contrast, empirical models may be useful, because the mathematical forms of the degradation trajectory can be mined from historical battery degradation processes, such as linear model [6] and exponential model [7]. As a special empirical model, stochastic models can capture the degradation uncertainty based on the nonlinear Wiener process with specific drift terms, such as linear [8] and exponential [9]. Furthermore, advanced filtering techniques, such as Kalman filter [10] and extended Kalman filter (EKF) [11], can help empirical models reduce impacts of measurement noise and external disturbance to further increase forecasting accuracy. However, without prior knowledge of the underlying degradation mechanisms and limited by specific expressions of degradation path (i.e., mathematical form in empirical models and drift terms in stochastic models), these models are difficult to generalize for batteries of diverse types.

Data-driven forecasting methods rely only on historical measurements to establish a degradation model, without explicit expressions to represent the prior knowledge of the inherent degradation behavior. Optimization methods integrating kernel techniques are popular solution tools to improve the prediction efficiency of data-driven methods, such as support vector machine (SVM) [12], relevant vector machine (RVM) [13] and Gaussian process regression (GPR) [14]. But it is difficult to generalize them due to their sensitivity to the kernel parameters [15]. To overcome this drawback, neural network-based methods have been implemented, such as Elman neural network [16]. But their prediction power is limited to short-data

windows. Recurrent neural networks (RNNs) can model long-range dependencies [17], but it is hard to implement back propagation in practice because the squashing nonlinear activation function causes the prediction power of the model exponentially decaying over time. Besides, gated recurrent unit (GRU) achieves more efficient operation for real-time prediction with less training parameters than other RNNs, such as LSTM (Long Short-term Memory) [18]. Although GRU and LSTM exhibit better performance in learning long-term dependence than the traditional RNN, they cannot describe stochastic degradation characteristics.

To close the gap, a GRU-driven degradation process is proposed for tracking battery degradation along with its increasing age. Specifically, the hidden state of the GRU neural network is introduced into the degradation process as a pattern term for capturing long dependencies among degradation rates. Then, a mapping function maps the pattern term to a mean of the degradation rate space. Thus, degradation rate is updated based on the entire degradation history instead of only the last time step, leading to a non-Markovian transitions. To cope with the non-Markovian transition, we developed a joint-learning sampling-based algorithm to estimate model parameters under the framework of Expectation Maximization (EM) method. Specifically, Sequential Monte Carlo (SMC) embedded in E-step provides posterior samples of degradation rates to M-step. Finally, a forecasting algorithm is proposed to estimate battery EoL based on the generative process of predicted degradation trajectory.

The remainder of this paper is organized as follows. Section 2 presents the model formulation, where a novel GRU-driven degradation process is developed. Section 3 investigates the parameter estimation under the framework of Expectation Maximization method. Section 4 explains how EoL can be estimated. Section 5 validates the proposed model using a case study of battery datasets. Finally, Section VI concludes the paper.

2. GRU-driven Degradation Process

Let $\{S_t; t \geq 0\}$ be a discrete-cycle battery degradation process, where S_t represents the degradation observation, which can be calculated by $S_t = 1 - C_t/C_0$, where C_t is the capacity at charge and discharge cycle t and C_0 is the initial capacity.

The evolution of the degradation process is controlled by degradation rate κ_t which represents the degradation speed. Additionally, the update of degradation rate is assumed to follow a recursive Gaussian distribution which is Markovian, i.e., $p(\kappa_t|\kappa_{t-1}) = \mathcal{N}(\kappa_t; \kappa_{t-1}, \xi^2)$. Note that ξ is the time-invariant standard deviation and the initial degradation state is κ_0 , which are to be estimated.

We introduce a GRU-driven pattern term $\mathbf{Q}_t$ to model long dependency in degradation rate κ_t , that is

$$\mathbf{Q}_t = \text{GRU}(\mathbf{Q}_{t-1}, \kappa_{t-1}; \boldsymbol{\theta}_G), \quad (1)$$

where $\mathbf{Q}_t$ characterizes the time-varying degradation pattern, and $\boldsymbol{\theta}_G$ is the vector of parameters of GRU neural network. Because (1) incorporates the entire evolution information, the Markovian assumption of traditional update of degradation rate will be violated. To enable the degradation rate κ_t controlled by the degradation pattern $\mathbf{Q}_t$, we use a mapping function $h_\mu(\cdot)$ to project $\mathbf{Q}_t$ from pattern space to the mean of the Gaussian space, i.e.,

$$p(\kappa_t|\kappa_{1:t-1}) = \mathcal{N}(\kappa_t; h_\mu(\mathbf{Q}_t; \boldsymbol{\theta}_h), \xi^2), \quad (2)$$

where $\boldsymbol{\theta}_h$ is the parameter vector of mapping function. For notational convenience, $\boldsymbol{\theta}_G$ and $\boldsymbol{\theta}_h$ are omitted in expressions of (1) and (2) in the rest of the paper.

By combining (1) and the mapping function $h_\mu(\cdot)$ in a single GRU-driven (Gd) neural network, we develop our GRU-driven degradation process model as follows:

$$\kappa_t \sim \mathcal{N}(Gd_{(\theta_G, \theta_h)}(\mathbf{Q}_{t-1}, \kappa_{t-1}), \xi^2), \quad (3-a)$$

$$S_t = S_{t-1} + \kappa_t + \int_{t-1}^t \varphi_\tau d\Lambda_\tau, \quad (3-b)$$

where φ_t is the time-invariant diffusion parameter for battery degradation, Λ represents standard Brownian motion. Due to the stability of φ_t , it can be reduced to a constant φ .

Compared with traditional methods proposed by Zhai, et al. [19], the proposed model can not only describe the degradation dynamics due to the adaptive degradation rate in the degradation process but also facilitate updating based on the entire historical degradation information thanks to the strong dependence-learning ability of GRU.

3. Parameter Estimation

Due to non-Markovian state transition in the proposed model, a joint-learning sampling-based expectation maximization algorithm is proposed to estimate model parameters. The model parameter vector $\Theta = (\Theta_\kappa, \varphi)$, where $\Theta_\kappa = (\xi, \theta_G, \theta_h)$, can be learnt by maximizing the log-likelihood of Θ , i.e., $\Theta^* = \text{argmax}_\Theta \ell(\Theta)$, under EM framework.

Sampling based E-step:

The variational lower bound of the log probability $\ell(\Theta)$, L_Θ , can be expressed as $L_\Theta = E_{\kappa_{1:T} \sim q(\kappa_{1:T})} \left\{ \log \left[\frac{p(S_{1:T}, \kappa_{1:T}; \Theta)}{q(\kappa_{1:T})} \right] \right\}$. Instead of maximizing $\ell(\Theta)$, we can just maximize L_Θ . Furthermore, optimal Θ^* can be obtained by $\Theta^* = \text{argmax}_\Theta E_{\kappa_{1:T} \sim q(\kappa_{1:T})} [\log p(S_{1:T}, \kappa_{1:T}; \Theta)]$.

Because the state transition is non-Markovian, the closed form of optimal Θ^* cannot be found by following the conventional EM algorithm. Thus, $\kappa_{1:T}$ need to be sampled from the posterior $p(\kappa_{1:T} | S_{1:T}, \Theta^{old})$.

Using the variational lower bound, posterior samples $\kappa_{1:T}$ can be obtained from $\kappa_{1:T} \sim q(\kappa_{1:T})$ rather than the original intractable posterior $p(\kappa_{1:T} | S_{1:T}, \Theta^{old})$. Here, SMC is adopted to implement the sampling. Specifically, $p(S_{1:t}, \kappa_{1:t})$ is the unnormalized target distribution at cycle t . For notational convenience, Θ is omitted in the remaining E-step. Then posterior samples $\kappa_{1:T}$ can be given by Algorithm 1.

Algorithm 1: Posterior sampling algorithm.

Initialization: Set $\kappa_0^n = \kappa_0$ and $\omega_0^n = 1/N$ for $n = 1, \dots, N$.

- 1 For $t = 1, \dots, T$ do
- 2 Draw ancestor $a_{t-1}^n \sim \mathcal{F}(\cdot | \mathbf{W}_{t-1})$ for $n = 1, \dots, N$, where $\mathcal{F}(\cdot | \mathbf{W}_{t-1})$ is the discrete probability distribution;
- 3 Draw rate $\kappa_t^n \sim p(\cdot | \kappa_{1:t-1}^n, S_t)$ for $n = 1, \dots, N$;
- 4 Set $\kappa_{1:t}^n = (\kappa_{1:t-1}^n, \kappa_t^n)$ for $n = 1, \dots, N$;
- 5 Compute the weight ω_t^n for $n = 1, \dots, N$;
- 6 Compute the normalized the weight W_t^n for $n = 1, \dots, N$;
- 7 End for

8 Sample $\chi \sim \mathcal{F}(\cdot | \mathbf{W}_T)$, and trace backward to obtain the particle path $\kappa_{1:T}^\chi$.

Joint-learning M-step:

Based on the previous step, Θ^* can be given by $\Theta^* = \operatorname{argmax}_{\Theta} \left[\log p(\kappa_{1:T}^\chi) + \sum_{t=1}^T \log p(S_t | \kappa_t^{B^\chi}) \right]$. Due to independence between Θ_κ and φ , parameter estimation for Θ can be divided into two subproblems: $\Theta_\kappa^* = \operatorname{argmax}_{\Theta_\kappa} \log p(\kappa_{1:T}^\chi; \Theta_\kappa)$, and $\varphi^* = \operatorname{argmax}_{\varphi} \sum_{t=1}^T \log p(S_t | \kappa_t^{B^\chi}; \varphi)$, which represents the maximum likelihood estimations of a GRU network and (3-b), respectively.

Based on the statistical properties of Brownian motion, (3-a) can be rewritten as a Gaussian distribution. Thus, Θ_κ and φ can be jointly estimated by adopting the negative log-likelihood of (3-a) to be the loss function of the neural network. Specifically, the expectation of (3-a) (if ignoring the constant term S_{t-1}) can be expressed as $E(S_t) \approx Gd_{(\Theta_G, \Theta_h)}(\mathbf{Q}_{t-1}, \kappa_{t-1})$. Furthermore, φ and ξ can be estimated by an extended Gd neural network. By Bayesian theorem, the variance of S_t can be expressed as $\operatorname{Var}(S_t) = \xi^2 + \varphi^2$. When constant terms are ignored, the simplified log-likelihood of the Gaussian distribution $L \propto \ln(\xi^2 + \varphi^2) + (\xi^2 + \varphi^2)^{-1} (S_t - Gd_{(\Theta_G, \Theta_h)}(\mathbf{Q}_{t-1}, \kappa_{t-1}))^2$ can be used as the loss function of the model training. Thus, parameter estimation problem is transformed into a neural network training problem as shown in Fig. 1.

Fig. 1. Parameter estimation of the proposed model.

4. EoL Forecasting

After obtaining all parameters, the proposed model can be used for battery degradation forecasting. Specifically, battery EoL refers to the number of cycles required for the battery capacity to fall below a specific malfunction threshold (usually 70-80%, 80% is used in this work) under a certain charge and discharge setting. Obviously, the initial SoH is always 1 based on the definition of SoH. Then battery EoL can be predicted based on the generated SoH series following Algorithm 2.

Algorithm 2: Battery EoL forecasting.

- 1 Initialize $i = 0$
- 2 While $S_{t_0+i} < 1 - \zeta$ do
- 3 Perform pattern transition: (1);
- 4 Predict degradation rate: (2);
- 5 Predict SoH: (3-b);

```

6       $i = i + 1;$ 
7      End while
8      Calculate  $EoL = i - 1.$ 

```

Based on parameters learned from historical SoH series, the proposed model can be formulated and EoL can be predicted correspondingly.

5. Case Study

To validate the applicability of the proposed model, we applied it to two battery datasets (Series A: CS2_33 and Series B: CS2_34) [20] and compared its performance with GRU model and adaptive wiener model [19]. These two datasets underwent the charging profile which was a standard constant current and constant voltage protocol with a constant current rate of 0.5C until the voltage reached 4.2V and then 4.2V was sustained until the charging current dropped to below 0.05A.

Further, SoH series extracted from raw data can be separated into a training set and a test set. The training set is used to train the model and estimate the model parameters. The first part of the train set is used to estimate model parameters, while the second part is used to tune hyperparameters. We call the first cycle of the test data to be the starting cycle of degradation forecasting. Following the starting cycle, SoH series are generated based on Algorithm 2 to estimate corresponding EoL.

By setting 70% and 80% of the degradation series to be the starting cycles, we can predict EoL using the proposed model and have concluded absolute errors of EoL prediction results in Table 1. Additionally, Figs. 2 and 3 illustrate the differences between predicted degradation trajectories and real ones. We can find that the proposed model outperforms the other two methods with respect of the accuracy of EoL prediction.

Fig. 2. Degradation trajectories of series A: Predicted SoH V.S. Real SoH, where starting cycle is 300.

Fig. 3. Degradation trajectories of series A: Predicted SoH V.S. Real SoH, where starting cycle is 388.

Table 1. Absolute error of EoL prediction from different starting cycles.
(Absolute error = |predicted EoL – real EoL|)

Dataset	Series A		Series B	
Starting Cycle	300	388	334	386
Real EoL	184	96	150	98
Method				
Proposed	72	17	41	13
GRU	77	26	72	42
Adaptive Wiener	121	67	110	67

6. Conclusion

In this paper, we proposed a GRU-driven degradation process to capture long-term dependencies among the historical degradation rates. To address challenges from non-Markovian state transitions, a joint-learning sampling-based expectation maximization algorithm is proposed to estimate model parameters. Then, a forecasting algorithm is developed to predict EoL based on the generative process of the proposed model. Finally, the case study using two battery datasets shows that the proposed model outperforms both GRU model and adaptive Wiener model in terms of EoL prediction accuracy.

In the future work, EoL distribution can be obtained according to the property of generative process of the proposed model. Additionally, the proposed model can be extended to high-dimensional degradation modelling.

References

- [1] Mehmet Ugras Cuma, Tahsin Koroglu. A comprehensive review on estimation strategies used in hybrid and battery electric vehicles. *Renewable and Sustainable Energy Reviews*. 2015, 42: 517-531.
- [2] Bhaskar Saha¹, Kai Goebel, Jon Christophersen. Comparison of prognostic algorithms for estimating remaining useful life of batteries. 2009, 31 (3/4): 293-308.
- [3] Changfu Zou, Lei Zhang, Xiaosong Hu, Zhenpo Wang, Torsten Wik, Michael Pecht. A review of fractional-order techniques applied to lithium-ion batteries, lead-acid batteries, and supercapacitors. 2018, 390: 286-296.
- [4] Chao Lyu, Qingzhi Lai, Tengfei Ge, Honghai Yu, Lixin Wang, Na Ma. A lead-acid battery's remaining useful life prediction by using electrochemical model in the Particle Filtering framework. 2017, 120: 975-984.
- [5] Jingliang Zhang, Jay Lee. A review on prognostics and health monitoring of Li-ion battery. 2011, 196: 6007-6014.
- [6] Chao Hu, Hui Ye, Gaurav Jain, Craig Schmidt. Remaining useful life assessment of lithium-ion batteries in implantable medical devices. *Journal of Power Sources*. 2018, 375: 118-130.
- [7] Lijun. Zhang, Zhongqiang Mu, Changyan Sun. Remaining useful life prediction for Lithium-ion batteries based on exponential model and particle filter. *IEEE Access*. 2018, 6: 17729-17740.
- [8] Guangzhong Dong, Zonghai Chen, Jingwen Wei, Qiang Ling. Battery health prognosis using Brownian motion modeling and particle filtering. *IEEE Transactions on Industrial Electronics*. 2018, 65(11): 8646-8655.

- [9] Dong Wang, Yang Zhao, Fangfang Yang, Kwok-Leung Tsui. Nonlinear-drifted Brownian motion with multiple hidden states for remaining useful life prediction of rechargeable batteries. *Mechanical Systems and Signal Processing*. 2017, 93: 531-544.
- [10] Xiaosheng Si. An adaptive prognostic approach via nonlinear degradation modeling: application to battery data. *IEEE Transactions on Industrial Electronics*. 2015, 62(8): 5082-5096.
- [11] Liming Deng, Yucheng Hsu, Hanxiong Li. An improved model for remaining useful life prediction on capacity degradation and regeneration of lithium-ion battery. *Annual Conference of the Prognostics and Health Management Society*. 2017.
- [12] Meru A. Patil, Piyush Tagade, Krishnan S. Hariharan, Subramanya M. Kolake, Taewon Song, Taejung Yeo, Seokgwang Doo. A novel multistage Support Vector Machine based approach for Li ion battery remaining useful life estimation. *Applied Energy*. 2015, 159: 285-297.
- [13] Yongzhi Zhang, Rui Xiong, Hongwen He, Michael Pecht. Validation and verification of a hybrid method for remaining useful life prediction of lithium-ion batteries. *Journal of Cleaner Production*. 2019, 212: 240-249.
- [14] Gozde O. Sahinoglu, Milutin Pajovic, Zafer Sahinoglu, Yebin Wang, Philip V. Orlik, Toshihiro Wada. Battery state-of-charge estimation based on regular/recurrent Gaussian process regression. *IEEE Transactions on Industrial Electronics*. 2018, 65 (5): 4311-4321.
- [15] Jianbo Yu. State of health prediction of lithium-ion batteries: Multiscale logic regression and Gaussian process regression ensemble. *Reliability Engineering and System Safety*. 2018, 174: 82-95.
- [16] Xiaoyu Li, Lei Zhang, Zhenpo Wang, Peng Dong. Remaining useful life prediction for lithium-ion batteries based on a hybrid model combining the long short-term memory and Elman neural networks. *Journal of Energy Storage*. 2019, 21: 510-518.
- [17] Yongzhi Zhang, Rui Xiong, Hongwen He, Michael G. Pecht. Long short-term memory recurrent neural network for remaining useful life prediction of Lithium-ion batteries. *IEEE Transactions on Vehicular Technology*. 2018, 67(7): 5695-5705.
- [18] Wenxian Duan, Chuanxue Song, Silun Peng, Feng Xiao, Yulong Shao, Shixin Song. An improved gated-recurrent unit network model for state-of-charge estimation of Lithium-ion battery. *Energies*. 2020, 13(23): 6366.
- [19] Qingqing Zhai, Zhisheng Ye. RUL prediction of deteriorating products using an adaptive wiener process model. *IEEE Transactions on Industrial Informatics*. 2017, 13(6): 2911-2921.
- [20] Wei He, Nicholas Williard, Michael Osterman, Michael Pecht. Prognostics of Lithium-ion batteries on Dempster-Shafer theory and the Bayesian Monte Carlo method. *Journal of Power Sources*. 2011, 196(23): 10314-10321.