Data-Driven Based Battery Health Prognosis with Diagnosis Uncertainties and Insufficient Training Data Sets

Zhimin Xi[1], Xiangxue Zhao[2], and Chao Hu[3]

[1]*University of Tennessee – Knoxville, Knoxville, TN, 37996, USA*
zxi1@utk.edu

[2]*University of Maryland – College Park, College Park, MD, 20742, USA*
xiangxuezhao@gmail.com

[3]*Iowa State University, Ames, IA, 50011, USA*
chaohu@iastate.edu

ABSTRACT

This paper investigates data-driven based battery prognosis with diagnosis uncertainties and insufficient training data sets. Four types of data-driven prognosis methods are investigated including the neural network, similarity-based approach, relevance vector machine, and a recently developed copula-based approach. The remaining useful life (RUL) predictions of lithium-ion battery capacity are compared with capacity estimation error due to the fact that onboard lithium-ion battery capacity estimation is difficult and almost always contains estimation errors. Thus, robustness of each prognosis methods can be studied for real time capacity RUL estimation. Furthermore, collection of sufficient run-to-failure training data sets for lithium-ion batteries is almost impossible even though it is desirable for all data-driven based methods. Therefore, robustness of these methods in terms of the insufficient training data sets is also studied. These insightful results will help designers choose appropriate prognosis algorithms in designing battery management systems (BMS) for lithium-ion batteries.

1. INTRODUCTION

A Copula-based sampling method was recently proposed for the RUL prediction, in which the statistical relationship between the failure time and the time realizations at specified degradation levels is constructed using various Copulas [1]. Generally, it is difficult to know which data-driven prognostic algorithm performs the best in a specific application because of the implicit relationship between the RUL and sensory signals. Typical data-driven prognostic approaches intend to build such an empirical functional relationship with some assumptions. For example, regression-based approaches assume certain functional structures and the similarity-based approach treats each training data set as an individual degradation model. The novelty of the Copula-based sampling approach is to eliminate the assumptions for functional relationship between the RUL and sensory signals, but to build a general statistical relationship between them purely driven from the available training data sets.

Despite abundant research in data-driven prognostics [1-3], little attention was paid for the data-driven prognosis with lack of offline training datasets. Generally, accuracy of the RUL prediction can be well maintained if sufficient offline training datasets were employed for building the degradation model. However, it is typically difficult to collect abundant run-to-failure training datasets even with the aid of the accelerated life testing. Hence, motivation of this paper is to study the performance of typical data-driven approaches when only limited number of training datasets are available so that pros and cons of these methodologies can be better revealed in a more practical manner.

2. REVIEW OF COPULA-BASED SAMPLING METHOD

This method uses a generic health index system that is composed of two distinguished health indices: physics health index (PHI) and virtual health index (VHI). In general, the PHI uses a dominant physical signal as a direct health metric and is thus applicable only if sensory signals are directly related to physics-of-failures. In contrast, the virtual health index (VHI) is applicable even if sensory signals are not directly related to system physics-of-failures. For example, typical battery systems only measure terminal voltage, current, and temperature which are not directly related to the battery health indicator. After the extraction of the PHI or VHI, the health index is further smoothed out using available tools. The purpose is to de-noise the health index originated from the noisy sensor signals. Furthermore, it is assumed in the Copula-based sampling method that the processed health index is non-decreasing over time or, in other words, health condition of the system becomes worse from healthy to failed states. It was proposed not to model the local non-monotonicity caused by noisy sensor signals.

The health index is discretized into a certain number of degradation levels to obtain a time realization matrix characterized by random variables that represent the random time realizations at the corresponding degradation levels. The matrix is then used for the statistical dependence modeling of the time realizations at different degradation levels using the Bayesian copula approach. Copula modeling is only performed between the time realizations at the i^{th} and the N^{th} degradation level (or the defined failure time) because the objective is to predict the failure time (or the RUL) of the engineering system provided that we know some actual time realizations at a certain number of degradation levels. Therefore, N-1 times of bivariate Copula modeling are required in the approach. It should be noted that the time realization at the N^{th} degradation level must be larger than or equal to the time at the i^{th} degradation level (i.e. $T_N \geq T_i$ for $i < N$) because it takes time for the system to decay. This property, however, is not guaranteed in the Copula modeling. Hence, a semi-Copula model was proposed to impose such a requirement by eliminating samples of $T_N < T_i$.

The failure time or RUL prediction is essentially a process to identify possible time realizations at the N^{th} degradation level provided that we know some true time realizations at a certain number of degradation levels. Mathematically, it is a process to identify a conditional PDF of T_N, $f(T_N|T_i)$, given that $T_i = a_i$. The health index of a new test unit may have experienced several defined degradation levels such that multiple true time realizations are known for the T_i where i ranges from 1 to j. Theoretically, it is feasible to obtain such a conditional PDF of T_N if a joint PDF $f(T_1, \ldots T_j, T_N)$ is available. In reality, however, only bivariate joint PDFs of T_i and T_N are constructed using Copulas. Hence, Eq. (1) was used to approximate the conditional PDF of T_N.

$$f\left(T_N \mid T_1 = a_1, T_2 = a_2, \ldots, \mathrm{T}_j = a_j\right) \cong \beta \prod_{i=1}^{j} f\left(T_N \mid T_i = a_i\right) \quad (1)$$

where β is a normalization parameter such that the integration of the PDF over the whole domain equals to one. Since semi-Copula model was used for modeling the time realization matrix by eliminating infeasible samples from the original Copula model, it is not feasible to obtain the analytical solution in Eq. (1) considering multiple conditional semi-Copula models. Hence, a sampling approach was proposed to estimate the mean, standard deviation, and empirical PDF of T_N.

3. ENHANCED COPULA-BASED PROGNOSIS

The copula-based sampling approach may fail to predict the RUL using Eq. (1) especially when the amount of training datasets is insufficient, which can be schematically shown in Fig.1. Due to the lack of training datasets, the copula modeling between T_i and T_N may not well represent the time realization correlation at the i^{th} and the N^{th} (i.e., the failure time) degradation levels for the population. Hence, it is very likely that the test unit reaches the i^{th} degradation level either too early (e.g., t_1 in Fig.1) or too late (e.g., t_2 in Fig.2) resulting in the failure calculation of the conditional PDF as shown in Eq. (1).

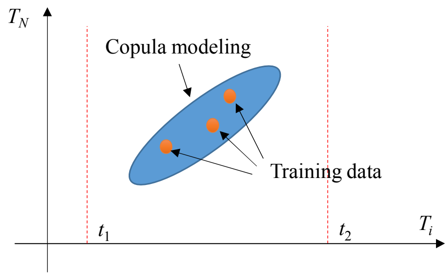


Figure 1. Failure of the RUL prediction using the copula-based approach

This issue can be addressed with simple handling of the copula model as shown in Fig. 2. For the abnormal degradation (as compared to the training datasets) reaches the i^{th} degradation level too early, the center of the copula model is shifted to t_1 so that the conditional PDF can be obtained. However, the identified possible failure time T_N should subtract the shifted time (i.e., t_m-t_1) as adjustment because the degradation speed is too fast (i.e., t_m-t_1 earlier than expected) reaching the i^{th} level for obtaining the equivalent conditional PDF of T_N. Similarly, the identified possible failure time T_N should add the shifted time (i.e., t_2-t_m) as adjustment because the degradation speed is too slow (i.e., t_2-t_m later than expected) reaching the i^{th} level for obtaining the equivalent conditional PDF of T_N.

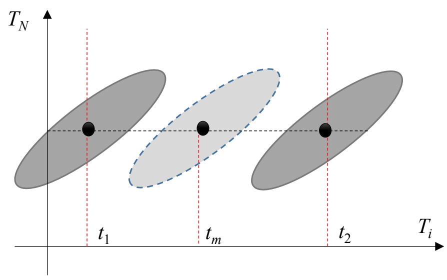


Figure 2. Copula-based approach for RUL prediction with abnormal degradation speed

For both scenarios, accuracy of the RUL prediction could still be maintained if the correlation pattern between T_i and T_N does not change significantly as if extra training datasets were available for the abnormal degradation units. Despite the wishful thinking due to the lack of training datasets, a warning message could be issued for both scenarios. For example, if the test unit reaches the i^{th} degradation level too fast, it is reasonable to judge that the true RUL may be less than the predicted RUL on the basis of available training datasets. Therefore, an early failure warning may be issued. On the other hand, a late failure message may be delivered if the test unit reaches the i^{th} degradation level too slow such that the true RUL may be higher than the predicted RUL. Since the marginal PDF of each T_i is available in the copula modeling, the criteria for issuing the warning message can be easily determined if the actual time realization of t_i exceeds a certain level of confidence bounds (e.g., 99%) of T_i.

To further enhance the efficiency and stability of the RUL prediction particularly for the lack of training datasets, Eq. (1) is proposed to be simplified as Eq. (2) where only the latest time realization is used for the RUL prediction. One of the major reason is schematically shown in Fig. 3. The conditional PDF is much wider at the early degradation stages (i.e., when i is a small value) than the latest stage (i.e., i=j) because stronger and stronger statistical correlation between T_i and T_N will be presented with the increase of i. As a result, the interaction of all available conditional PDFs is very similar to the narrowest one.

$$\beta\prod_{i=1}^{j} f\left(T_N \mid T_i = a_i\right) \cong f\left(T_N \mid T_j = a_j\right) \tag{2}$$

Figure 3. Typical conditional PDFs of failure time at different degradation levels

3. CASE STUDY

A set of four Li-ion batteries were run through three different operational profiles (i.e., charge, discharge and impedance) at room temperature [4]. Charging was carried out in a constant current (CC) mode at 1.5A until the battery voltage reached 4.2V and then continued in a constant voltage (CV) mode until the charge current dropped to 20mA. Discharge was carried out at a constant current (CC) level of 2A until the battery voltage fell to the cut-off voltage for four batteries. Impedance measurement was carried out through an electrochemical impedance spectroscopy (EIS) frequency sweep from 0.1 Hz to 5 kHz. Repeated charge and discharge cycles resulted in accelerated aging of the batteries while impedance measurements provide insight into the internal battery parameters that change as aging progresses. The experiments were stopped when the batteries reached end-of-life (EOL) criteria, which was a 30% fade in rated capacity (i.e., from 2Ahr to 1.4Ahr). In this example, capacity fade as shown in Fig. 4 was employed for the demonstration. Each battery cell was used once as test unit while the other three as training units.

For the back-propagation NN, the input was the battery capacity and the output was corresponding RULs. The network architecture employed a single node input layer, a single node output layer, and one hidden layer with tan-sigmoid transfer function. The number of nodes for hidden layer was optimized as ten nodes based on the least average error using the training datasets. The RUL was predicted each time with a provided capacity value. For the copula-based sampling approach, 50 degradation levels were evenly defined from healthy to failure states and 49 copula models were constructed based on three training datasets. RUL prediction results are shown in Fig. 5, where the solid line is the true RUL versus the operation time in days, dots, star and cross symbols are RULs prediction from NN, copula-based approach, and similarity-based approach, respectively. The RUL was predicted where actual capacity measurement was available. The copula-based approach consistently shows accurate RUL prediction for all cells and the predicted RUL converges to the true RUL when close to the failure time. The similarity-based approach shows overall good accuracy except for cell #3 which has the longest life time. The fundamental reason is that the training units are not similar to the test unit by predicting relative short RULs at any operation time. The back-propagation NN could provide very accurate RUL prediction such as cell #4 but also could be not accurate such as cell #1 and cell #2. In addition, convergence to the true RUL may not be ensured even when close to the failure time as shown in cell #3. Due to the black box nature of NN, it is also difficult to analyze the reason for inaccurate RUL prediction.

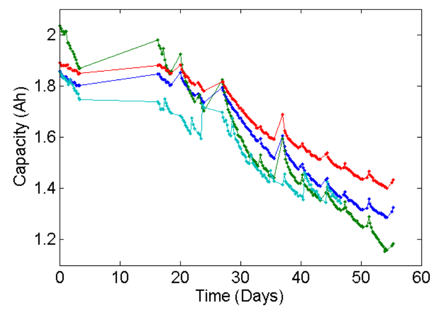
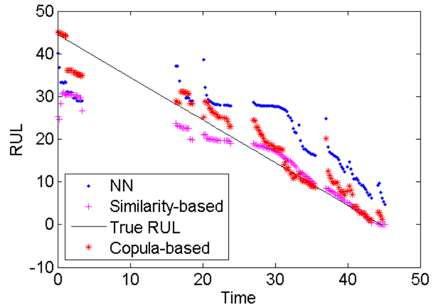


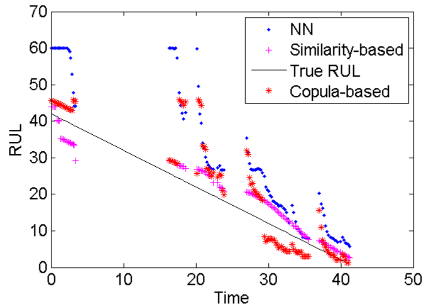
Figure 4. Training data of capacity degradation

To further quantify the RUL prediction accuracy, the error of RUL prediction was calculated for four cells, where statistics such as mean and standard deviation

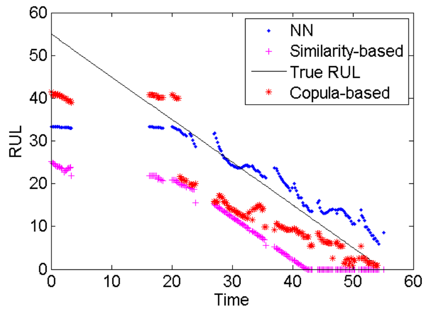
(STD) are listed in Table 1. With respect to the mean of the error for each cell, the copula-based approach shows the overall best accuracy, followed by the similarity-based approach and the NN. In terms of the STD of the error, the copula-based approach again shows the smallest value indicating its overall best prediction robustness.



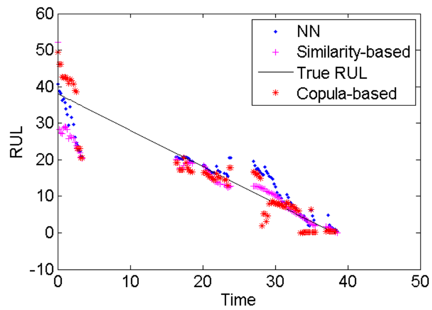
a. Cell #1



b. Cell #2



c. Cell #3



d. Cell #4

Figure 5. RUL comparison using three approaches

Table 1. RUL accuracy comparison of four battery cells

	NN		Similarity-based		Copula-based	
RUL error	Mean	STD	Mean	STD	Mean	STD
Cell #1	8.55	3.29	3.61	4.78	2.86	2.27
Cell #2	12.74	8.51	4.39	1.58	5.88	5.48
Cell #3	5.39	5.75	13.02	7.09	6.93	4.15
Cell #4	3.09	3.53	3.29	4.18	3.29	3.19
Average	**7.44**	**5.27**	**6.08**	**4.41**	**4.74**	**3.77**

4. CONCLUSION

This paper investigated performances of three data-driven prognosis approaches under the condition that offline training datasets are insufficient. The similarity-based approach could present large RUL prediction error if the training units are not similar to the online test units, which is very likely to occur when training datasets are insufficient. The accuracy of NN is not robust and it is difficult to further improve the NN due to the black box nature. The enhanced copula-based approach shows very good accuracy, efficiency and robustness even with the lack of offline training datasets.

5. ACKNOWLEDGMENTS

The research was supported by U.S. National Science Foundation.

6. REFERENCES

[1]. Xi, Z., Jing, R., Wang, P., and Hu, C., A Copula-based sampling method for data-driven prognostics, Reliability Engineering and System Safety, Vol.132: 72-82, 2014.

[2]. Wang T, Yu J, Siegel D, Lee J. A Similarity-Based Prognostics Approach for Remaining Useful Life Estimation of Engineered Systems. IEEE, International Conference on Prognostics and Health Management 2008, Oct 6-9.

[3]. Khan, Muhammad Adnan. Bearing prognostics using neural network under time varying conditions. Diss. Concordia University, 2010.

[4]. http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/

BIOGRAPHIES

Zhimin Xi is an assistant professor in the Department of Industrial and Systems Engineering at University of Tennessee – Knoxville.

Xiangxue Zhao is a Ph.D. student in the Department of Mechanical Engineering at University of Maryland – College Park.

Chao Hu is an assistant professor in the Department of Mechanical Engineering at Iowa State University.