Segmentation Based Feature Evaluation and Fusion for Prognostics

Feature Selection based on Segment Evaluation

Vepa Atamuradov[1], Fatih Camci[2]

[1]*Production Engineering Laboratory (LGP), INP-ENIT 47 Av. d'Azereix, 65000 Tarbes, France*
vepa.atamuradov@enit.fr

[2]*Advanced Micro Devices (AMD), USA*
fatih.camci@amd.com

ABSTRACT

Quantification of feature goodness, called feature evaluation, is crucial in the identification of best features and achieving high accuracy in diagnostics and prognostics. Even though feature evaluation for diagnostics is a mature area, it is a developing research area for prognostics. The feature goodness for prognostics is measured by change in degradation. Most, if not all, of existing methods, analyze the feature change in the whole failure degradation. In other words, features collected throughout the failure degradation are analyzed to create a goodness value for the feature. In reality, the goodness of the features may change during the failure progression. A feature may be a good representative of failure progression in the initial phase but not in the final phases, or vice versa. This paper presents a methodology that divides the features into segments, each of which may have different goodness for prognostics. Thus, some part of the feature may be good, whereas the others not. The presented approach leads to extract more value from the features' changing properties during the failure degradation. The method has been applied to simulated and real datasets obtained from Li-ion batteries aging tests. State of health (SoH) estimation accuracy is enhanced with the presented approach.

1. INTRODUCTION

Diagnostics and prognostics are the major steps in Prognostics and Health Management (PHM) (Faith Camci & Chinnam, 2005; Zhang & Lee, 2011). Diagnostics is the process of identification of existing failures with its severity and/or location. Diagnostics is a classification or clustering problem in nature depending on the availability of labeled data. On the other hand, prognostics is the process of identification of Remaining Useful Life (RUL) of the system or component under observation given its current health status. Prognostics is a forecasting problem that makes it more challenging due to many uncertainties involved in failure progression.

The raw sensory data is mostly large and include many useless data. Features are extracted from raw sensory data using statistical or computational methods such as mean, standard deviation, kurtosis, skewness, and Fourier transform etc. Features give more clear information about the system with less size compared to the raw data. Many features can be extracted from raw data and each feature gives different types of information about the system.

The goodness of features varies depending on the purpose. The features are evaluated and selected based on their goodness for the defined purpose. Feature selection and evaluation in diagnostics have been studied in the literature extensively (Liang, Liu, Li, He, & Xu, 2016; Senoussi & Chebel-Morello, 2011). However, feature evaluation for prognostics is a relatively new problem and researchers have started publishing articles about the evaluation of features for prognostics in recent years (J. Coble & Hines, 2009). However, none of these works, to the best of our knowledge, addresses the features' changing goodness during failure progression. A feature that does not represent the failure progression in the initial phase of the failure may represent the progression fully in the final phases. In contrary, a feature that represents the failure degradation perfectly in the initial phase may not represent the degradation in the final phases. In other words, the sensitivity of feature to the failure degradation may change in the life of the system due to environmental effects or operational profiles. Thus, feature evaluation based on the full life of the component or system may be misleading.

The feature with highest goodness values may change during the failure degradation. It is possible to use different features to measure the failure degradation in different phases of the

component or system life. To the best of our knowledge, feature evaluation considering the features' subsequence goodness change in the failure progression has not been addressed in the literature yet. This paper aims to fill this gap in the literature by presenting a segmentation based feature evaluation and fusion algorithm for prognostics.

The organization of the paper is as follows: Section 2 gives the literature review; Section 3 discusses the presented method for dynamic feature evaluation and fusion. Section 4 presents the results obtained by the presented method on simulated and real data obtained from Li-ion batteries. The paper is concluded with Section 5.

2. LITERATURE REVIEW

Diagnostics is the process of identification of the failure mode, its location, and severity. Measuring the effectiveness of features for diagnostics has been studied extensively (Cecille, Dana, & B;, 2015; Guana, Yuana, Leea, Najeebullaha, & Rasela, 2014; Hannah Inbarani, Bagyamathi, & T;, 2015; Lamraoui, Barakat;, Thomas, & Badaoui, 2015; Mwangi, Tian, & Soares, 2014). These studies can be categorized in two groups: non-transformed and transformed analysis (Mwangi et al., 2014). In the former one, the features are evaluated as they are without converting them into another form. Evaluation can be performed by an individual or combined analysis of features. The major disadvantage of this approach is the requirement of ignoring unselected features and not benefiting from the features that received a low grade in the evaluation. In the latter group, the features are converted into a different form in such a way that all features contribute to the new formatted features. The selection is performed based on the transformed features. In such a transformation, low graded features may also be contributed partially in the selected transformed features (Tianzhen, Hao, & Jingang, 2015). Principal component analysis (PCA) and independent component analysis (ICA) are two examples of this approach (Tianzhen et al., 2015; Z, Jin, & Guangming, 2011). PCA transforms the features into the lower dimension, which is obtained with the contribution of old features based on their variance. The new features are ranked from the ones that hold most information to least information. Thus, the features with least information can be dropped from the analysis leading to reduced number of features. ICA, on the other hand, performs the transformation based on the independence of features.

Changing features' goodness has been studied for diagnostics in the literature. A feature may not be good in the beginning for diagnostics, but become good after the failure reaches to a point. Online PCA has been developed to be able to handle the changing goodness of the features (Honeine, 2012).

Even though there has been extensive work on feature evaluation for diagnostics, this is not true for prognostics (F. Camci, Medjaher, Zerhouni, & Nectoux, 2013). The nature of the problem in prognostics is totally different from diagnostics. Thus the methods used in diagnostics for feature evaluation cannot be used for prognostics.

Prognostics involves two major steps: health state estimation and trending (Omer F. Eker, Camci, & Jennions, 2015). Health state estimation can also be considered as severity detection in diagnostics. Health state can be continuous or discrete. If health state is continuous, then health state is trended towards future until it reaches a predefined threshold. When the trended health state reaches the threshold, the time is named as the expected failure time. In the basic approach, a feature such as average vibration, maximum temperature, the median force applied, etc. is assumed to be the health state. In most cases, one feature is not sufficient for effective prognostics. Several features can be processed and fused to obtain health state. Physical or data-driven methods are used to obtain the health state using various sensors. Even though these methods are effective, features to be used as input to these methods should be evaluated and selected. One can extract a large number of features, all of which cannot be processed to obtain the health status. Thus, it is important to evaluate and select the features before processing them.

If the health state is discrete, then stochastic process models can be used to estimate the time to move from the current state to the failure state (Faith Camci & Chinnam, 2005). Transition probabilities between states are used for estimating the remaining useful life in these models (O F; Eker & Camci, 2013). State space models can also be used for discrete or continues health states. The parameters of stochastic and state space models should be learned using the features selected. Thus, it is important to evaluate features and select the good ones for further processing. The first study in prognostics feature evaluation was presented in (J. Coble, 2009; J. B. Coble, 2010). The authors presented prognosability, trendability and monotonicity feature evaluation metrics to identify the most suitable features for better RUL estimations. Feature evaluation methodology was also presented in (F. Camci et al., 2013), that quantifies the monotonicity of the trend in the features by dividing them into windows. Statistical difference of data in consecutive windows is employed to quantify the monotonicity.

Genetic algorithm (GA) has been used to generate a formula to calculate a new feature that represents the failure degradation using existing features (Linxia, 2014). GA selects features among a feature pool and operators from math operations pool. The resultant formula has been evaluated using its effectiveness in representing the failure degradation.

Entropy-based sensor selection method is proposed in (Liu, Wang, Liu, Zhang, & Peng, 2015) for prognostics. This method quantifies the trend representing the failure degradation for given sensory dataset and entropy is used to represent the uncertainty within the data. Trigonometric functions and their cumulative transformation have been used to extract monotonic features. The goodness of the features

for prognostics has been quantified by analysis of monotonicity and trendability (Javed, Gouriveau, Zerhouni, & Nectoux, 2015). Monotonicity is the continuous increasing or decreasing nature of the feature and quantified as the sum of positive and negative derivatives. Trendability is basically defined as the correlation of a feature with time.

Feature extraction and selection processes affect the data structure. The change due to feature selection process has been controlled through preserving the local and global structure of the data in order to achieve effective feature selection (Peng, Xu, Liu, & Li, 2015; Wang, Zhang, Yin, & Liu, 2014). A feature selection method that aims to preserve the local and global structure has been presented in (Wang et al., 2014), and its application in prognostics has been discussed in (Peng et al., 2015). A novel diagnostics feature extraction methodology was proposed in (Gelman, Patel, Persin, Murray, & Thomson, 2013) for bearing fault detection based on combination of frequency-domain based techniques. The authors in (Qu, Bechhoefer, He, & Zhu, 2013) proposed a new physics-based feature extraction approach for gearbox fault detection using acoustic emission (AE) signals, which outperformed other frequency-domain features in gear fault detection.

Even though these feature evaluation studies on prognostics have led to some level of success, they still miss an important aspect of the features. The effectiveness of a feature or sensory data may not be static throughout the life of the component or system under observation. A feature may be a good representative of the failure progression in the initial failure phase but not in the later part of the failure. In contrary, a feature may be effective close to the final phase of the failure but not in the initial phase. Thus, a single evaluation of a feature for whole failure degradation duration may not be effective. The bad failure progression representation in one part may negate the good failure progression representation in another part. This different goodness of a feature during the failure degradation is handled with online PCA approaches for diagnostics and other classification problems as discussed above. However, evaluation of features' changing goodness in its degradation path has not been addressed in the literature. This paper presents a feature selection and fusion algorithm based on subsequence evaluation by enhancing and improving our previous work (Atamuradov & Camci, 2016). The main contributions of the paper can be summarized as follows:

- The change of feature goodness in different phases of failure degradation is demonstrated.

- The inefficiency of feature evaluation based on its single analysis in the whole failure degradation is validated.

- The segmentation-based feature evaluation and selection methodology is presented to improve battery SoH estimation.

3. PROBLEM DEFINITION

State of Health (SoH) indicates the health of the system at a time indicating its location in the failure degradation pathway. The term 'real SoH' will be used as the ground truth of failure degradation. Features react differently to the failure progression. The failure degradation representation capability (FDRC) of a feature can be defined as the feature's sensitivity to the real SoH progression and its real SoH identification ability.

Monotonically decreasing or increasing features with failure degradation have been accepted as good representatives. However, the good representation may be partial in the life of the component or system. Figure 1 displays a linearly degrading system. The solid linear line in the middle gives the real SoH by time. FDRC of feature 1 is very high in the first phase of the life. However, the feature stays constant with some noise in the second phase of the component life with low FDRC. In contrary, feature 2 does not represent the failure progression in the initial phase. It then becomes a good representative in the second part with continuously increasing value. Note that y-axis for features shows feature values that will be used to extract SoH, not SoH itself.

A single analysis of a feature based on the whole failure degradation may mislead the feature evaluation. It is important to take the most value from good representative phases of the features and avoid the effects of bad representative phases. Feature evaluation method that leads to high goodness value in the first (second) phase of the first (second) feature and low goodness value for the remaining phases for both features is needed to extract the most value from the features.

4. METHODOLOGY

The proposed methodology has four major steps: feature extraction and normalization, segmentation and selection, feature fusion, and prognostics.
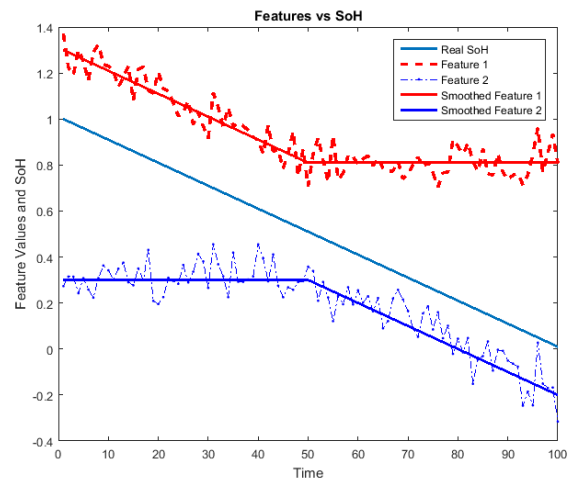

Figure 1. Features with different failure progression.

4.1. Feature Extraction and Normalization

Different statistical features are extracted using descriptive statistics as shown in Table 6. Only the features with better failure representability (e.g. monotonicity etc.) are selected and used for battery SoH estimation among time-domain based features. Features may be decreasing or increasing or may be in different scales. It is important to normalize these features for evaluation and fusion. Equation (1) shows the normalization for decreasing features with the failure progression, whereas equation (2) gives the normalization for increasing features with the failure progression. As a result of the normalization process, the features will be decreasing from 1 to 0 as the failure propagates. Please note that maximum ($F_{max(1:T)}$) and minimum ($F_{min(1:T)}$) of first T values of the features are selected in order to handle potential noise within the features.

$$NF_{i,t} = \frac{F_{i,t}}{\max(F_{i,1:T})}, \forall\ F_i \text{that decreases with failure progression (1)}$$

$$NF_{i,t} = \frac{\min(F_{i,1:T})}{F_{i,t}}, \forall\ F_i \text{that increases with failure progression (2)}$$

After the normalization, the degradation samples have been separated into two groups as training and testing. Training samples are used to identify the model parameters, whereas the testing samples are used to check the model accuracy with unforeseen samples.

4.2. Feature Segmentation and Selection

Time series segmentation has been defined as the decomposition of a time series data into a series of discrete segments. Time series segmentation is a mature area and has been applied in many engineering fields (Glezakos, Tsiligiridis, & Yialouris, 2014). Sliding window algorithm is used in this study for feature segmentation.

The sliding window algorithm is initiated with *n* data points (called window) for a given feature (T {t_i: $i=1…n$}). A linear model has been fitted to the data points within the window. If curve fitting error does not exceed a predefined threshold, the size of the window is increased to include the next data point outside of the window. This process continues until a point where the curve fitting error exceeds the threshold or end of the time series has been reached. When the curve fitting error exceeds the threshold, the last data point added to the window is defined as the start of the second segment. The predefined threshold values were optimized by the user to split a feature into two homogeneous segments due to features' degradation behaviors. Table 1 gives the pseudo-code of the general segmentation algorithm. The features will be split into multiple segments depending on the change in the progression. In our case, the features consist of two segments. In a different system, the number may be different. However, a number of segments should not be too many, which may lead to instability. The detection of the transition point in various samples is a challenge. Thus, a high number of segments may lead to an increased challenge in the determination of the segment. The window size initially was set to 2. The optimized segmentation threshold values for the simulated and the real features are given in Table 2 and Table 3, respectively.

Feature selection occupies an important role in many engineering problems after feature extraction step which helps to classify either redundant or irrelevant feature from useful ones without losing much information before model construction. In literature, mostly, selection of useful features is carried as a whole selection to be used in latter steps of aimed application. But in this current work, extracted features are partitioned into two homogeneous segments by segmentation algorithm and then subsequence feature selection is carried out. As stated in problem definition section, prognostics features do not degrade until to the end of life with same increasing/decreasing monotonicity

```
sliding_window (data, max_err):
anchor=1;
while not segmented data
w=2; % w: window size. err: estimated MSE
 if err (data(anchor:anchor+w)) <max_err
   w=w+1;
 else %convert into segment
segments <- data (anchor: anchor+(w-1));
anchor= anchor+w; %update anchor
 end
end
```

Table 1. Sliding Window time series segmentation.

Feature	**F1**	**F2**	**F3**	**F4**
Threshold	0.0036	0.0078	0.004	0.007
Feature	**F5**	**F6**	**F7**	**F8**
Threshold	0.0125	0.0034	0.0038	0.003
Feature	**F9**	**F10**	**F11**	**F12**
Threshold	0.0043	0.0034	0.00521	0.0045
Feature	**F13**	**F14**	**F15**	**F16**
Threshold	0.0041	0.0045	0.0044	0.0088
Feature	**F17**	**F18**	**F19**	**F20**
Threshold	0.007	0.0058	0.0041	0.0041

Table 2. Segmentation threshold values for simulated features.

Feature	**F1**	**F2**	**F3**	**F4**
Threshold	0.0065	0.008	0.038	0.017
Feature	**F5**	**F6**	**F7**	**F8**
Threshold	0.0129	0.00375	0.1	0.0067
Feature	**F9**	**F10**	**F11**	**F12**
Threshold	0.06	0.0077	0.011	0.0067
Feature	**F13**	**F14**	**F15**	**F16**
Threshold	0.008	0.0152	0.00891	0.019
Feature	**F17**	**F18**	x	x
Threshold	0.0075	0.012	x	x

Table 3. Segmentation threshold values for real features.

statistics. Hence, prior feature degradation behavior analysis and evaluation are very important before feature fusion process. The feature selection process was carried out in two steps; as a whole and segment based selection to emphasize the importance of proposed approach and then used in feature fusion for prognostics. An adopted feature selection criterion is the mean square error value between real SOH value and extracted features.

4.3. Feature Fusion

Feature fusion is the process of combining different features to enhance the SoH estimation of the electro-mechanical system. There are different methods used in the literature for fusion. The weighted average is one of the most widely used approaches in fusion (Williard, He, Osterman, & Pecht, 2013). Fusion is commonly used in diagnostics and prognostics. A novel fusion approach in RUL prediction based on superstatistic and information fusion has been presented in (J. Lui, Zhang, Zuo, & Xie, 2014). Composite health index is obtained through fusion in (K. Lui, Gabraeel, & Shi, 2013). A three-step multi-sensor fusion approach for prognostics of partially degraded systems was proposed in (Fang, Paynabar, & Gebraeel, 2017), based on functional principal components analysis (FPCA) algorithm and penalized (log)-location-scale regression. The proposed approach was effectively validated on simulated aircraft engine degradations. In (Yan, Liu, Zhang, & Shi, 2016) authors proposed a fusion based health index (HI) construction methodology for better asset prognostics which operates under multiple operational conditions. The data-level HI fusion methodology was validated on simulated data effectively. It is very difficult, if not impossible, to extract a single feature that perfectly represents the failure progression. Thus, it is important to extract value from different features for better SoH estimation.

Fusion is important for effective usage of segmented features since the segments of features with low goodness should be supported by other features with high goodness in the same phase of the failure degradation. The fusion algorithm to be employed should incorporate the changing goodness values of the features. The fusion process at any time t requires analysis of the goodness of the features (or FDRC). The feature with high FDRC should be given more importance in SoH estimation compared to the feature with low FDRC. FDRC of feature i at time t is represented by weight parameter as $w_{i,t}$. The fusion process bases on weighted average calculation as shown in equation (3):

$$SoH_{f,t} = \frac{\sum_{i=1}^{N} w_{i,t} NF_{i,t}}{\sum_{i=1}^{N} w_{i,t}} \qquad (3)$$

Initially, the feature weight values are initialized using feature samples obtained from similar degraded systems, if available. If previous data for the degraded system is not available, the weights of all features start with the same number giving equal importance to all. Since the system

under observation may degrade differently compared to other similar systems, the weight values are updated with newly available data. The update is based on two major parameters: Time-based SoH similarity degree ($SD_{i,t}$) and overall segment estimation accuracy (EA_i). Time based similarity degree ($SD_{i,t}$) quantifies the effectiveness of the normalized feature of sample i at a given time (t) calculating the error at previous time ($t-1$). Overall segment estimation accuracy (EA_j) quantifies the effectiveness of a segment calculating errors in SoH calculation within the segment using similar samples. Note that $SD_{i,t}$ bases on the degrading sample under observation, however, EA_j uses historical data (previously degraded systems) to measure the overall segment effectiveness. $SD_{i,t}$ is calculated using the difference between real SoH and normalized feature value at time $t-1$ as shown in (4), where i is the sample number, t is the time.

$$SD_{i,t} = exp\left(-|SoH_{i,t-1} - NF_{i,t-1}|\right) \qquad (4)$$

Overall segment estimation accuracy (EA_i) is calculated using the sum of square of differences between real SoH and normalized feature values as shown in equation (5), where j is the segment number, S is the sample size, i is the sample number, $T_{i,j}$ is the duration sample i spent in segment j.

$$EA_j = 1 - \frac{1}{S}\left(\sum_{i=1}^{S}\left(\frac{1}{T_{i,j}}\sum_{t=1}^{T_{i,j}}\left(SoH_{i,t}^{real} - NF_{i,t}\right)^2\right)\right) \qquad (5)$$

High estimation accuracy of a feature segment and similarity degree should increase the weight of the feature. In contrary, low estimation accuracy and similarity score should decrease the weight. The update of weight at time $t+1$ is performed by using the weight at time t, similarity degree at time t, and segment estimation accuracy as shown in equation (6). As seen from the equation, weight values increase in any case. Increase amount depends on the values of segment estimation accuracy and similarity degree. High segment estimation accuracy and similarity degree lead to high increase in the weight value, whereas their low values lead to low increase.

$$w_{i,t+1} = w_{i,t} + \left(EA_{i,t} * SD_i\right) \qquad (6)$$

Note that the weight values of a feature will converge to a number if FDRC of the feature stays constant. If FDRC of feature changes during the failure progression, the weight value will change and converge again in the new segment with new FDRC. Once the weight is updated, the SoH estimation is calculated based on equation (3).

4.4. Prognostics: Remaining Useful Life Prediction (RUL)

Selection of prognostics algorithms plays a great importance in condition monitoring to increase system reliability, availability, and safety (Aizpurua & Catterson, 2015; Saxena, Celaya, Saha, Saha, & Goebel, 2010). In this paper, two

statistical trending algorithms Holt's double exponential smoothing (HDES) (Holt, 2004) and Autoregressive Integrated Moving Average (ARIMA) (Saha, Goebel, & Christophersen, 2009) have been used for one-step RUL prediction due to their implementation simplicity.

4.4.1. Holt's Double Exponential Smoothing Method

Exponential smoothing methods have been used widely in literature for time series forecasting (Maia & de Carvalho, 2011). Holt's method is extended version of simple exponential smoothing with trend component. This method consists of level, trend estimation and forecast equations as shown in (7), (8), and (9).

$$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \quad (7)$$

$$b_t = \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1} \quad (8)$$

$$\hat{y}_{t+h} = \ell_t + hb_t \quad (9)$$

Where ℓ_t is level estimation at time t, derived by weighted average of observation y_t, b_t denotes trend estimation at time t, derived by weighted average of estimated trend, where α and β are smoothing parameters for level and trend equations to be estimated. Intervals for β is, $0 \le \beta \le 1$ and for α is $0 \le \alpha \le 1$. Initial values for $\ell_0 = y_1$ and $b_0 = y_2 - y_1$ can be assigned.

4.4.2. Autoregressive Integrated Moving Average

ARIMA is another widely-used forecasting approach integrating autoregressive (AR) with a degree of p and moving average (MA) with a degree of q model based on d times differencing. Given a time series y_t, and ARIMA (p,d,q) model is formulated in (10):

$$(1 - \varphi_1 B - \cdots - \varphi_p B^p)(1 - B)^d y_t = c + (1 + \theta_1 B + \cdots + \theta_1 B^q)e_t \quad (10)$$

B is lag operator, φ's and θ's are parameters of AR and MA part of the model and e_t is error term assumed to be independent and identically distributed (normal distribution with zero mean. The proposed methodology scheme is illustrated in Figure 3.

5. RESULTS & DISCUSSIONS

The presented approach has been implemented in two types of datasets: simulated data and Li-ion battery degradation data.

5.1. Simulation Results

The model is trained using some selected samples first and then the trained model is used for testing unused samples. In training, all extracted features from the selected samples went through segmentation, selection and fusion steps. In testing, the same selected features in training step are also selected

from testing samples and are used in the fusion algorithm to estimate the SoH. In the simulated dataset, SoH degradation has been simulated first. Figure 2 displays the simulated SoH. Then, features that are functions of the simulated SoH degradation have been generated. 20 features have been simulated based on this simulated SoH degradation. Each feature is created based on one or two functions of SoH with some noise. When two functions are used, the first function represents some part of the SoH degradation, whereas the second one represents the remaining part of the SoH degradation. Totally 20 features for training the system and 20 for testing have been simulated using the same a methodology based on simulated SoH value. Figure 4 displays some of the simulated SoH degradation. In order to evaluate the value of the presented approach, SoH estimation has been performed under two scenarios. Scenario 1 involves utilizing all selected features without segmentation. Scenario 2 involves using segmentation based feature evaluation as proposed in this paper. Feature selection process has been applied without segmentation in Scenario 1. The best features selected are displayed in the first column of Table 4. Feature numbers 18, 20, 17, 15, and 9 have been selected in scenario1. In scenario 2, proposed segmentation based feature evaluation has been performed. Segmentation process has led to two segments in some features. Thus the evaluation process is reported for two segments. As seen in the second and third columns in Table 4, features (2, 8, 11, 14, 12) and (20, 1, 17, 8, 6) have been selected for the segments respectively. As a result, one can observe the value of segmentation through better evaluation of features in different phases of life with the changing weight values that are used in the fusion process. Different phases contribute differently in the fusion process avoiding the negative effect of a phase to the evaluation of a feature. The change of weight values depends on the similarity degree of the features. High similarity degree leads to high weight values. It is important to identify the phase that is highly correlated with the failure progression and give high importance to this feature in this phase as well as ignoring the uncorrelated phase. Figure 5 and Figure 6 display the SoH estimation results for both scenarios
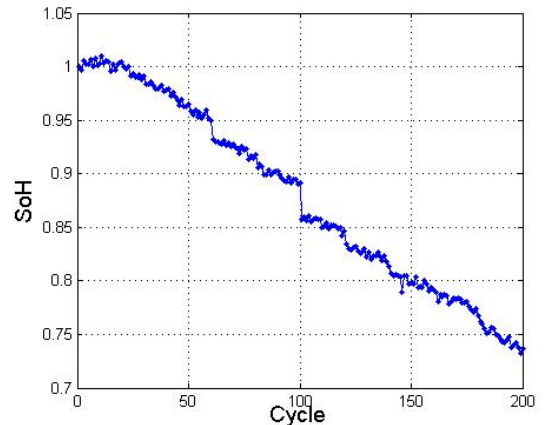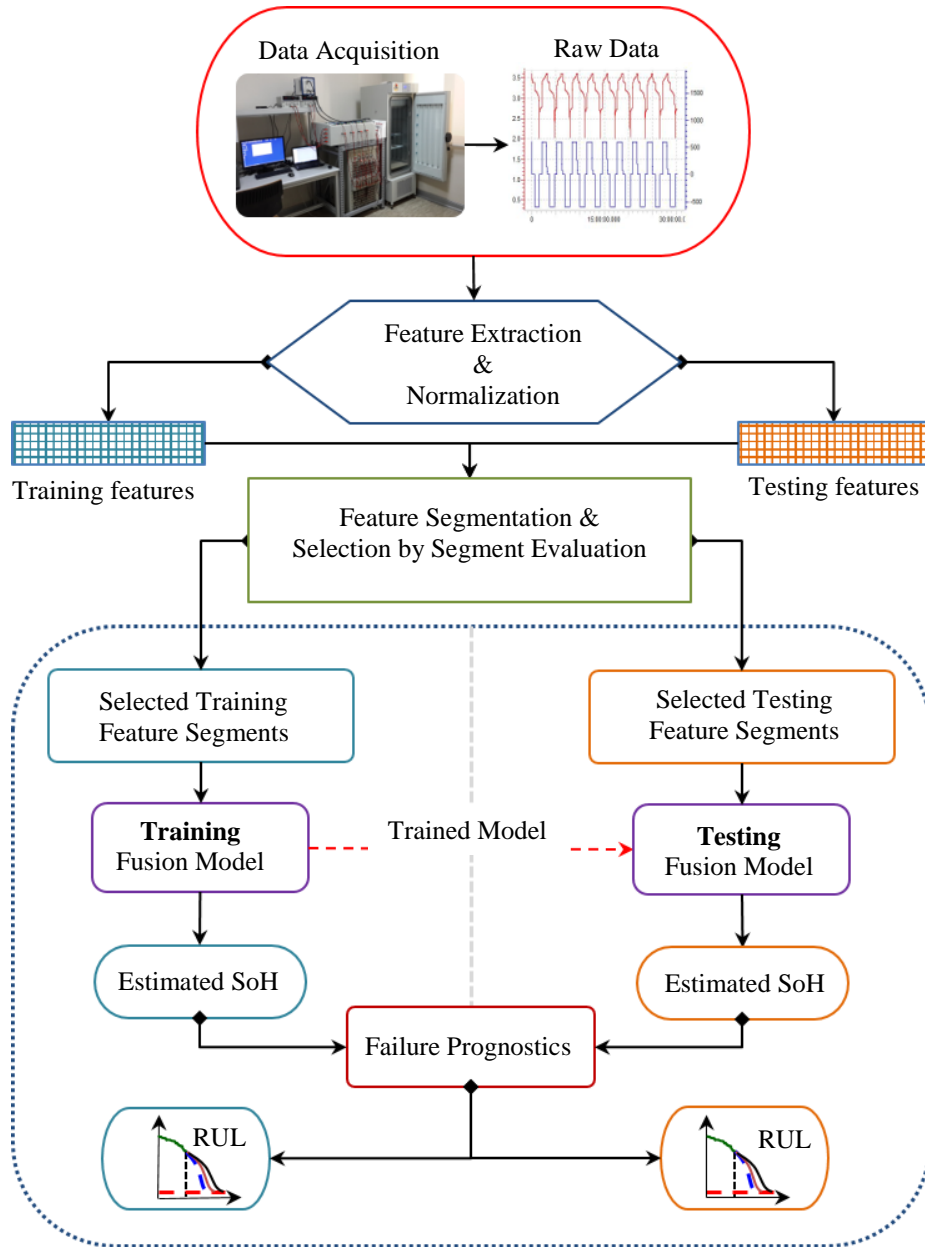

Figure 2. Simulated SoH feature.

Figure 3. Proposed methodology scheme.

for training and testing, respectively. Figure 5.a and Figure 6.a (Figure 5.b and Figure 6.b) display the simulated SoH degradation and SoH estimation result for scenario 1 (scenario 2). As seen from segmented approach leads to better SoH estimation. Improvement in SoH estimation in the first phase is better observed. The SoH estimation errors for both training and testing results are given in Table 5 as Mean Square Error (MSE). As seen from the table, segmented feature selection fusion reduces the error. Results show that segmentation based feature evaluation and fusion improves SoH estimation.

Selected F_i	1st segment selection from F_i	2nd segment selection from F_i
18 20 17 15 9	2 18 11 14 12	20 1 17 8 6

Table 4. List of selected features.

	Training	Testing
Scenario 1	0.0004	0.001
Scenario 2	0.0002	0.0004

Table 5. Estimated SoH feature MSE.
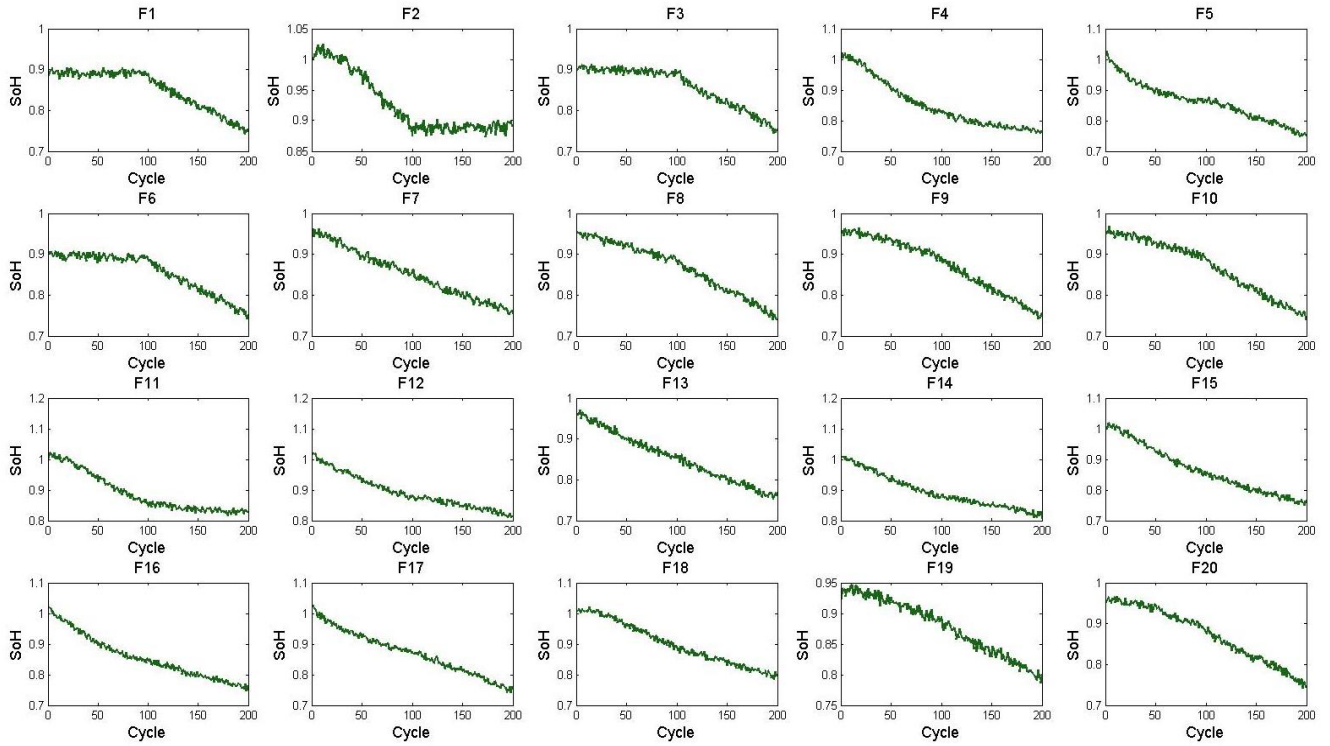
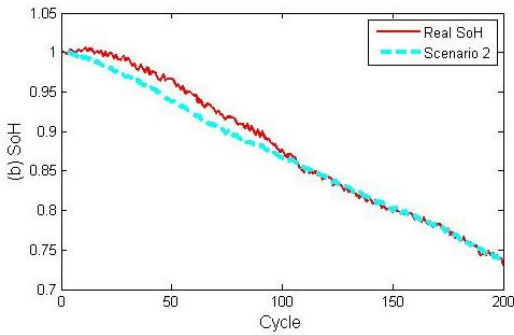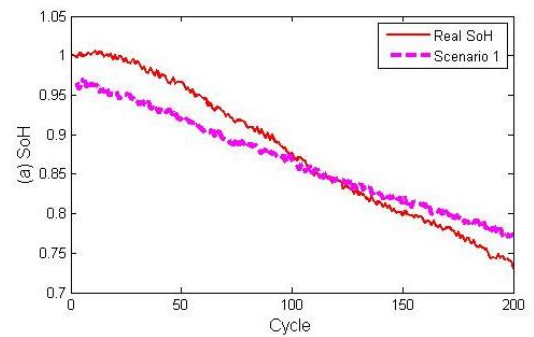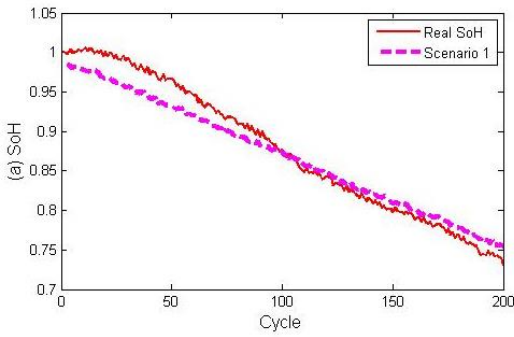Figure 4. Simulated features.
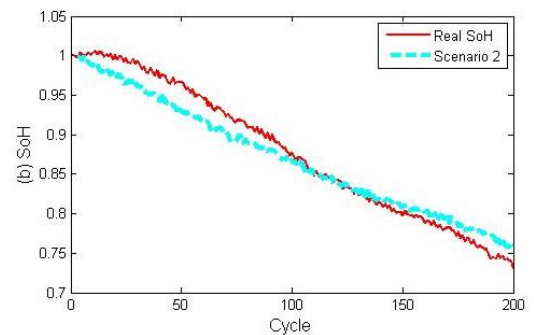
Figure 5. Fusion results for training.

Figure 6. Fusion results for testing.

5.2. Li-ion Degradation

Li-ion batteries have been used in many areas in today's world. SoH estimation and prognostics play a crucial role in reliability, safety, and cost of lithium-ion batteries (F Camci, Ozkurt, Toker, & V, 2015). This section discusses the application of the presented methodology on the degradation data obtained from Li-ion batteries in the lab environment.

Two LiFePO4 14505 0.6Ah batteries with capacity and a nominal voltage of 3.2V are used in this paper, one for training the proposed methodology and second one for testing. Figure 7 shows the experimental setup used for the accelerated degradation tests. An accelerated test consists of three main phases: cycling, test measurement, and characterization. Prior to cycling process, the cell is kept at 45 ℃ for two hours to stabilize cell temperature. Cycling process for Li-ion cells is carried out by charging cell up to 3.6V with a constant current of 0.6Ah and discharging cell up to cut-off voltage 2V, SoH and internal resistance features were tested to make sure whether aging threshold is met or not. Constant current constant voltage (CCCV) charging protocol was used throughout the entire life of batteries. In constant current charging (galvanostatic mode) step, batteries were charged until it reached the maximum voltage threshold and in constant voltage (potentiostat mode) step, the voltage was applied to the batteries until current dropped down to predefined threshold value. A final characterization test takes place right after if the cells have met its predefined SoH threshold, where EIS and other characterization tests were applied to extract SoH indicatory features of cells.

Several measurements have been collected during the accelerated tests such as cell capacity, charge (Chg)/ discharge (DChg) voltage curves, Chg/ DChg OCV curves at rest time and internal resistances through Electrochemical Impedance Spectroscopy (EIS). As seen from Figure 8, degradation of the battery is easily detected from changing Chg/DChg voltages and OCV curves. There are totally 18 features (shown in Figure 9) extracted from aged battery measurements. These measurements have been further processed to obtain different features such as different types of resistances within the battery. Figure 9 and Figure 10 displays the examples of the extracted and normalized features through the lifetime of batteries for both cells; y-axis displays values of the features, whereas the x-axis gives the time within a cycle. The progression of the features as the battery degrades is shown as different lines in the figure. The change in the shape of the line in the figure indicates the failure progression. Current discharge capacity is obtained through integration of discharge current using Coulomb Counting.

In segmentation, the features have been segmented at most two parts. A number of segments depend on the error threshold in the segmentation process. Due to variance in features, the threshold of each feature should be analyzed and decided distinctly. Table 7 displays the features selected in both scenarios. As seen from the table, segmentation process leads to the selection of different features. In scenario 1, the best features are features 16, 4, 18, 17, 13, whereas features 3, 4, 5, 10, 16 and feature 13, 16, 18, 17, 4 are selected as best features in the first and second segment in scenario 2. Since Li-ion batteries are electrochemical substances, degradation of batteries is highly dependent on the environmental changes such as; temperature and aging cycle profiles. Thus, accurate SoH estimation of Li-ion batteries should involve analysis of multiple features. The fusion process has been performed using two scenarios to evaluate the value of the presented approach. Scenario 1 involves feature evaluation and SoH estimation without segmentation process, whereas Scenario 2 involves the same processes with a segmentation based selection. In training step, the battery cell 1 is used for training for both scenarios. In testing, the battery cell 2 is used in the model trained with battery cell 1.

After features are selected, SoH estimation through the fusion of selected features. In order to identify the best number of features to be used, SoH estimation is performed with the best
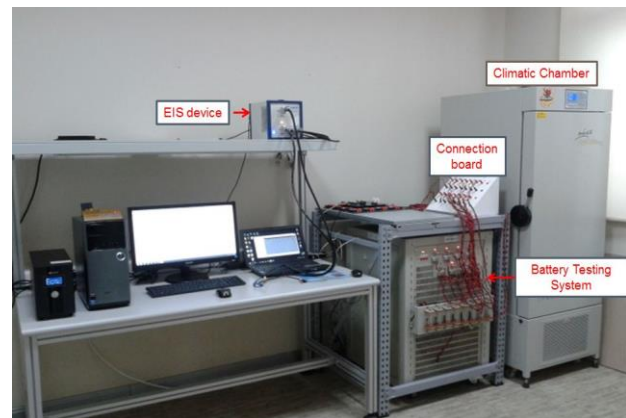

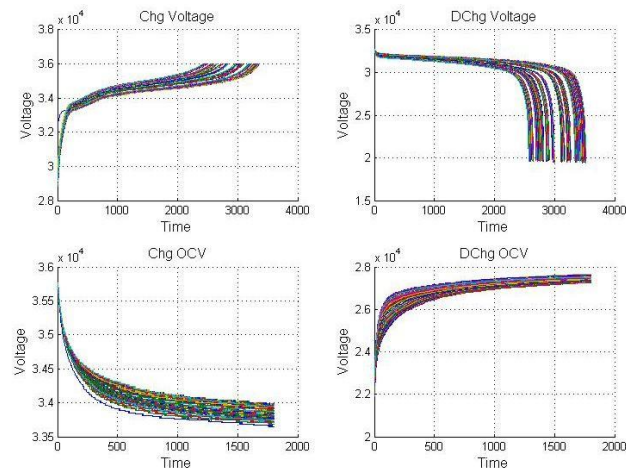Figure 7. Experimental rig setup.

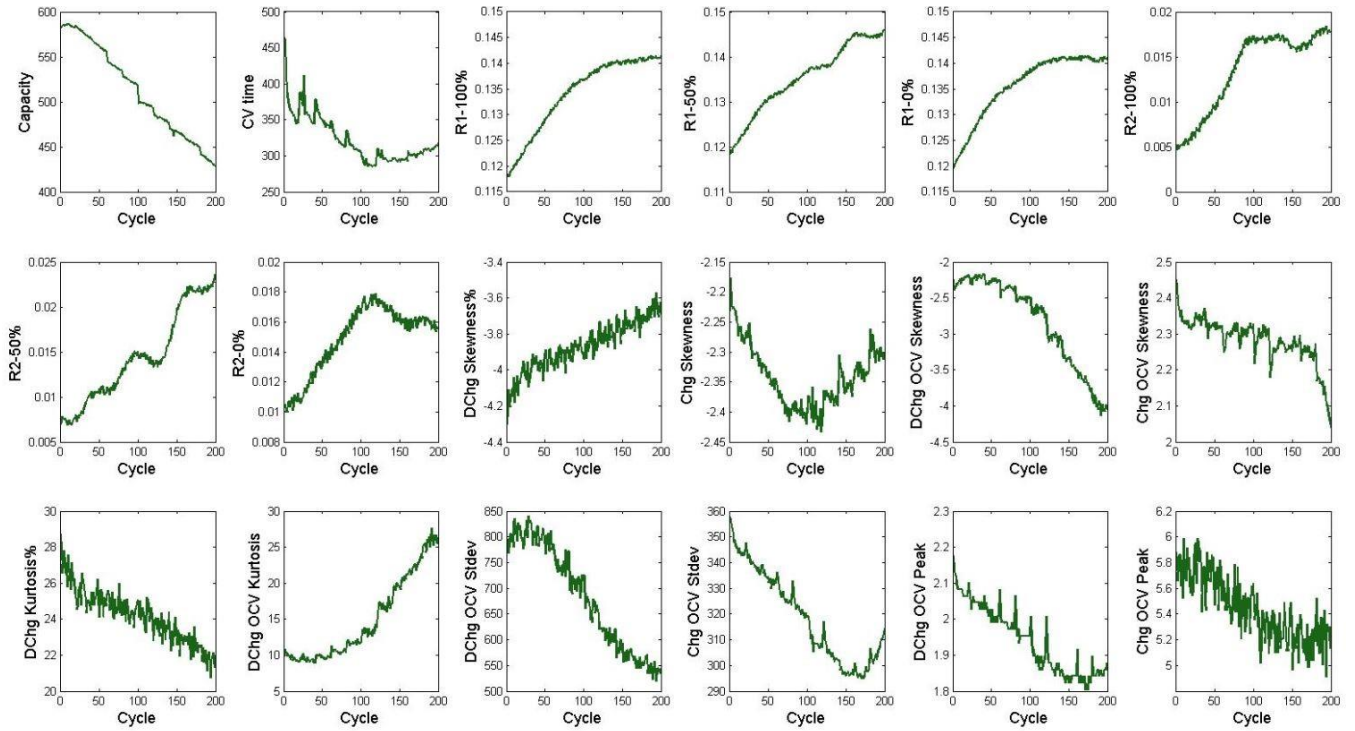Figure 8. Chg/DChg voltage and OCV curves.
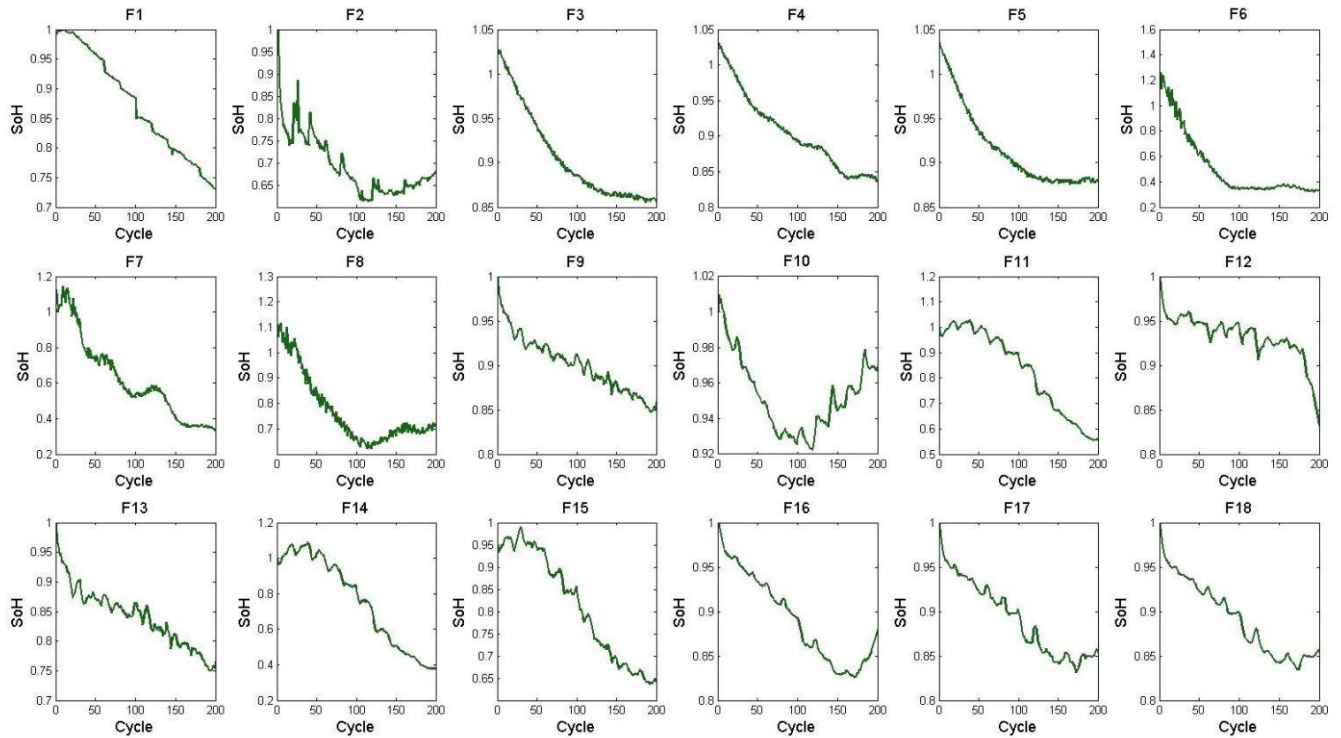
Figure 9. Extracted battery features.

Figure 10. Normalized battery features.

feature and the next best feature is added in the next iteration. In each iteration mean square error is calculated. Figure 11 displays the error on y-axis and number of features used on the x-axis. As seen from the figure, the minimum error is obtained with 8 features. SoH estimation results for both scenario 1 and scenario 2 are given in Figure 12 for training and Figure 13 for testing. As seen from the Figure 12 and Figure 13, the presented method leads to better SoH estimation both in training and testing. SoH estimation curve is much closer to scenario 2 to the real SoH curve compared to scenario 1. In testing, SoH of the battery cell 2 was estimated based on the trained model. Table 8 shows the error terms for training and testing for both scenarios. As seen from the table, the scenario 2, which is the proposed approach, leads to less error.

Index	Measurement	Features	Index	Measurement	Features
1	DChg cycle	Capacity	11	Chg Voltage	Skewness
2	Chg cycle	CV time	12	DChg OCV	Skewness
3	EIS	Resistance (R1-100% SoC)	13	DChg OCV	Kurtosis
4	EIS	Resistance (R1-50% SoC)	14	DChg OCV	Standard Deviation
5	EIS	Resistance (R1-0% SoC)	15	DChg OCV	Peak value
6	EIS	Resistance (R2-100% SoC)	16	Chg OCV	Skewness
7	EIS	Resistance (R2-50% SoC)	17	Chg OCV	Standart Deviation
8	EIS	Resistance (R2-0% SoC)	18	Chg OCV	Peak value
9	DChg Voltage	Skewness			
10	DChg Voltage	Kurtosis			

Table 6. Extracted features from the measurements.

The RUL estimation has been performed using SoH estimation based on both scenarios. Two data-driven algorithms (i.e., Holt's method and ARIMA) have been used for trending the SoH.

Implementation of Holt's method starts with the estimation of α, β parameters, and initialization of l_0 and b_0. The initial value for level equation (7) is set to the first value of fused SoH feature, 0.9925 and for trend equation (8), was set to the difference of second and first values of fused SoH feature which is -0.0016. Smoothing parameters α was set to 0.45 and β was set to 1 which gave minimum MSE error value in estimation. For ARIMA(p,d,q) method, orders of p, q was chosen as 1, and for differencing parameter d was chosen as 2 with minimum model fit estimation error. Two RUL prediction results for both Holt's method and ARIMA are given in Table 9 for fused SoH feature including both training and testing steps. End of Life (EOL) for estimated prognostics SoH from training step is 163 cycles, where for testing step is 189 cycles. The predicted RUL for either Holt's or ARIMA is calculated by differencing the estimated EOL (eEOL) value with the given prediction point (see Table 9), with the confidence interval of 95%. For example, the estimated EOL cycle for Holt's at point 125 was 157, from this given information the RUL prediction can be calculated as eEOL – 125, which is 32. As seen from Table 9 ARIMA gives better prediction results when compared to Holt's

Selected F_i	1st segment selection from F_i	2nd segment selection from F_i
16	3	13
16, 4	3, 4	13, 16
16, 4, 18	3, 4, 5	13, 16, 18
16, 4, 18, 17	3, 4, 5, 10	13, 16, 18, 17
16, 4, 18,17, 13	3, 4, 5, 10, 16	13, 16, 18, 17, 4

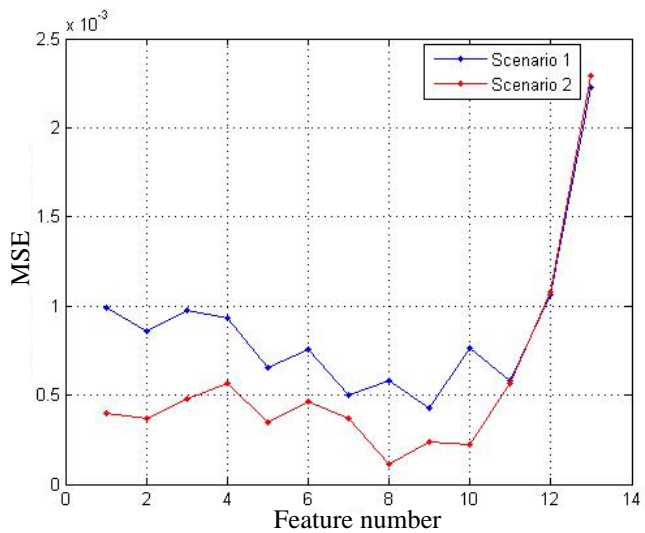Table 7. List of selected sample features.

Figure 11. Fusion MSE values.

method for both training and testing SoH estimations. It's important to note that prognostics algorithm selection is one of the essential steps to perform accurate RUL predictions, especially for systems where component degradation varies under different operational conditions.

	Training	Testing
Scenario 1	0.0004	0.0012
Scenario 2	0.0001	0.0007

Table 8. Estimated SoH feature MSE.

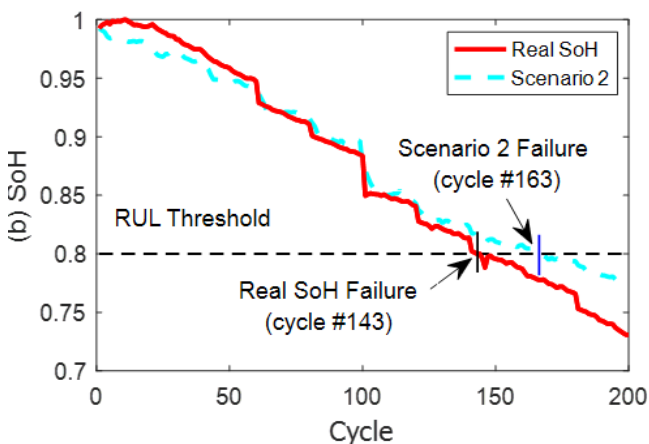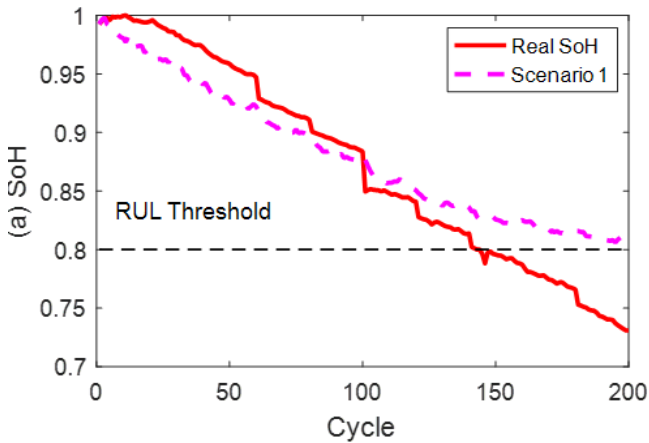| Point	Training		Real RUL	Testing		Real RUL
	Holt's	ARIMA		Holt's	ARIMA	
105	49	38	59	80	82	84
125	32	38	39	87	62	67

Table 9. RUL prediction results.

Figure 12. a) Scenario 1, and b) scenario 2 fusion results for training.
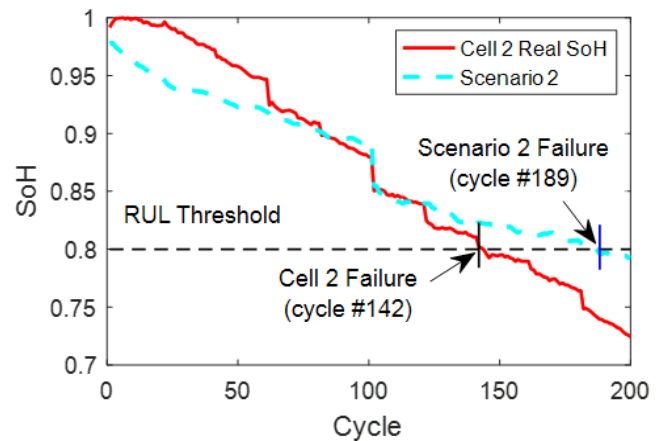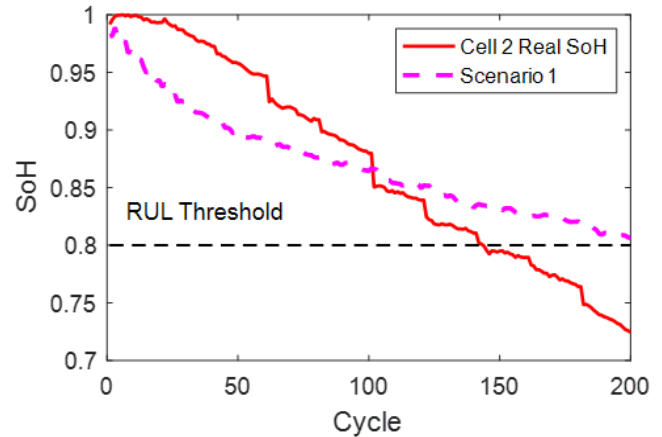

Figure 13. a) Scenario 1, and b) scenario 2 fusion results for testing.

5.3. Discussions

As seen from Figure 12 and Figure 13, the Scenario 2 estimation results of the training and testing steps overestimated the real EOL, which may result in a late RUL predictions affecting maintenance actions negatively. However, despite the SoH overestimation, the proposed segmentation-based feature evaluation methodology improved the SoH estimation by decreasing the estimation error when compared to the Scenario 1. It's also important to note that the battery capacity, which was assumed as the best FDRC feature, may also have nonmonotonic degradation under different operational conditions (Williard et al., 2013). In this case, a filter based methods can be used to select the least sensitive feature with the better prognostics parameter (J. Coble & Hines, 2009) as a ground truth feature in component condition monitoring where the features have dynamic degradation. Consequently, it's believed that the proposed feature evaluation methodology can improve a battery SoH estimation under different operational profiles for failure prognostics.

6. CONCLUSION

Analysis of a feature for the whole life cycle may mislead the feature evaluation. Features may reflect the SoH differently in different phases of the life of the electro-mechanical system. It is important to take the most value from good representative phases of the features and avoid the negative effects of bad representative phases. A methodology for segmentation-based feature evaluation of features has been presented. SoH estimation based on segmented evaluation has been developed. The presented approach has been demonstrated in simulated and Li-ion battery degradation data. The results show that evaluation of features with segmentation improves SoH and RUL estimation results. Optimization of a number of segments and handling variance in segmentation points from different samples are the future research topics.

REFERENCES

Aizpurua, J. I., & Catterson, V. M. (2015). Towards a Methodology for Design of Prognostic Systems. *Annual Conference of the Prognostics and Health Management Society*, (October), 1–13.

Atamuradov, V., & Camci, F. (2016). Evaluation of Features with Changing Effectiveness for Prognostics. *Annual Annual Conference of the Prognostics and Health Management Society*.

Camci, F., & Chinnam, R. B. (2005). Dynamic Bayesian networks for machine diagnostics: Hierarchical Hidden Markov models vs. competitive learning. *Proceedings of the International Joint Conference on Neural Networks*, *3*, 1752–1757.

Camci, F., Medjaher, K., Zerhouni, N., & Nectoux, P. (2013). Feature Evaluation for Effective Bearing Prognostics. *Quality and Reliability Engineering International*, *29*(4), 477–486.

Camci, F., Ozkurt, C., Toker, O., & Atamuradov, V. (2015). Sampling based State of Health estimation methodology for Li-ion batteries. *Journal of Power Sources*, *278*, 668–674.

Cecille, F., Dana, K., & B;, O. (2015). An evaluation of classifier-specific filter measure performance for feature selection. *Pattern Recognition*, *48*(5), 1812–1826.

Coble, J. (2009). An Automated Approach for Fusing Data Sources to Identify Optimal Prognostic Parameters. *Annual Conference of the Prognostics and Health Management Society*.

Coble, J. B. (2010). *Merging data sources to predict remaining useful life--an automated method to identify prognostic parameters*. Retrieved from http://trace.tennessee.edu/utk_graddiss/683

Coble, J., & Hines, J. W. (2009). Identifying optimal prognostic parameters from data: a genetic algorithms approach. *Proceedings of the Annual Conference of the Prognostics and Health Management Society*, 1–11.

Eker, O. F., & Camci, F. (2013). State Based Prognostics with State Duration Information,. *Quality Reliability Engineering International*, *29*(4), 465–476.

Eker, O. F., Camci, F., & Jennions, I. K. (2015). Physics-based prognostic modelling of filter clogging phenomena. *Mechanical Systems and Signal Processing*, *75*, 395–412.

Fang, X., Paynabar, K., & Gebraeel, N. (2017). Multistream sensor fusion-based prognostics model for systems with single failure modes. *Reliability Engineering and System Safety*, *159*(November 2016), 322–331.

Gelman, L., Patel, T. H., Persin, G., Murray, B., & Thomson, A. (2013). Novel Technology Based on the Spectral Kurtosis and Wavelet Transform for Rolling Bearing Diagnosis. *Int. J. Progn. Health Manag*, *4*(2), 1–7.

Glezakos, T., Tsiligiridis, T. A., & Yialouris, C. P. (2014). Piecewise evolutionary segmentation for feature extraction in time series models. *Neural Computing and Applications*, *24*(2), 243–257.

Guana, D., Yuana, W., Leea, Y. K., Najeebullaha, K., & Rasela, M. K. (2014). A Review of Ensemble Learning Based Feature Selection. *IETE Technical Review*, *31*(3), 190–198.

Hannah Inbarani, H., Bagyamathi, M., & T;, A. A. (2015). A novel hybrid feature selection method based on rough set and improved harmony search. *Neural Computing and Applications*, *26*(8), 1859–1880.

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, *20*(1), 5–10.

Honeine, P. (2012). Online Kernel Principal Component Analysis: A Reduced-Order Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(9), 1814–1826.

Javed, K., Gouriveau, R., Zerhouni, N., & Nectoux, P. (2015). Enabling Health Monitoring Approach Based on Vibration Data for Accurate Prognostics. *IEEE Transactions on Industrial Electronics*, *62*(c), 647–656.

Lamraoui, M., Barakat;, M., Thomas, M., & Badaoui, M. El. (2015). Chatter detection in milling machines by neural network classification and feature selection. *Journal of Vibration and Control*, *21*(7), 1251–1266.

Liang, L., Liu, F., Li, M., He, K., & Xu, G. (2016). Feature selection for machine fault diagnosis using clustering of non-negation matrix factorization. *Measurement*, *94*, 295–305.

Linxia, L. (2014). Discovering Prognostic Features Using Genetic Programming in Remaining Useful Life Prediction. *IEEE Transactions on Industrial Electronics,* , *61*(5), 2464–2472.

Liu, L., Wang, S., Liu, D., Zhang, Y., & Peng, Y. (2015). Entropy-based sensor selection for condition monitoring and prognostics of aircraft engine. *Microelectronics Reliability*, *55*(9–10), 2092–2096.

Lui, J., Zhang, M., Zuo, H., & Xie, J. (2014). Remaining useful life prognostics for aeroengine based on superstatistics and information fusion. *Chinese Journal of Aeronoutics*, 1086–1096.

Lui, K., Gabraeel, N. Z., & Shi, J. (2013). A data-level fusioin model for developing composite health indices for degradation modeling and prognostics analysis. *Automation Science and Engineering, IEEE Transactions*, *10*(3), 652–664.

Maia, A. L. S., & de Carvalho, F. de A. T. (2011). Holt's exponential smoothing and neural network models for forecasting interval-valued time series. *International Journal of Forecasting*, *27*(3), 740–759.

Mwangi, B., Tian, T. V., & Soares, J. C. (2014). A Review of Feature Reduction Techniques in Neuroimaging. *Neuroinformatics*, *12*(2), 229–244.

Peng, Y., Xu, Y., Liu, D., & Li, J. (2015). Locality structure preserving based feature selection for prognostics. *Intelligent Data Analysis*, *19*(3), 659–682.

Qu, Y., Bechhoefer, E., He, D., & Zhu, J. (2013). A New Acoustic Emission Sensor Based Gear Fault Detection Approach. *Int. J. Progn. Health Manag*, *4*, 1–14.

Saha, B., Goebel, K., & Christophersen, J. (2009). Comparison of prognostic algorithms for estimating remaining useful life of batteries. *Transactions of the Institute of Measurement and Control*, *31*(3–4), 293–308.

Saxena, A., Celaya, J., Saha, B., Saha, S., & Goebel, K. (2010). Metrics for Offline Evaluation of Prognostic Performance. *Int. J. Progn. Health Manag.*

Senoussi, H., & Chebel-Morello, B. (2011). Feature selection and categorization to design reliable fault detection systems. *Control*.

Tianzhen, W., Hao, X., & Jingang, H. (2015). Cascaded H-Bridge Multilevel Inverter System Fault Diagnosis Using a PCA and Multiclass Relevance Vector Machine Approach. *IEEE Transactions on Power Electronics*, *30*(12), 7006–7018.

Wang, L., Zhang, J., Yin, J., & Liu, H. (2014). Global and Local Structure Preservation for Feature Selection. *IEEE Transactions on Neural Networks and Learning Systems*, *25*(6), 1083–1095.

Williard, N., He, W., Osterman, M., & Pecht, M. (2013). Comparative analysis of features for determining state of health in lithium-ion batteries. *Int. J. Progn. Health Manag*, *2013*(4), 1–7.

Yan, H., Liu, K., Zhang, X., & Shi, J. (2016). Multiple Sensor Data Fusion for Degradation Modeling and Prognostics under Multiple Operational Conditions. *IEEE Transactions on Reliability*, *65*(3), 1416–1426.

Z, W., Jin, C., & Guangming, D. (2011). Constrained independent component analysis and its application to machine fault diagnosis. *Mechanical Systems and Signal Processing*, *25*(7), 2501–2512.

Zhang, J., & Lee, J. (2011). A review on prognostics and health monitoring of Li-ion battery. *Journal of Power Sources*, *196*(15), 6007–6014.

BIOGRAPHIES

Vepa Atamuradov was born in Turkmenistan in 1986. He received his Ph.D. in Electrical-Computer Engineering from Selçuk University (2016), MS in Computer Engineering from Fatih University (2009) in Turkey. He has worked as RA in "Development of Failure Prognostics and Maintenance Planning System for Point Mechanisms in Railway (2008-2009)" and "Development of Design based State-of-Health and Remaining Useful Life Estimation Techniques for Battery Management Systems and Its Application to Rechargeable Batteries (2014-2016)" projects, granted by The Scientific and Technological Research Council (TUBITAK) of Turkey. Currently, he is working as a Postdoc research fellow at Tarbes National School of Engineering (ENIT), France, since July 2016. His research interests include failure diagnostics and prognostics of Industrial systems using Machine Learning and Statistical Methods.

Fatih Camci's main research areas include Decision Support Systems in Industrial Engineering field focusing on Prognostics and Health Management (PHM). He has worked on research projects related to decision support systems on PHM and Energy in the USA, Turkey, UK, and France. He currently works as Manager of IT Architecture at AMD.