

**In-Situ early anomaly detection and remaining useful lifetime prediction for high-power white LEDs with distance and entropy-based long short-term memory recurrent neural networks**

Wen, Minzhen; Ibrahim, Mesfin Seid; Meda, Abdulmelik Husen; Zhang, Guoqi; Fan, Jiajie

**DOI**

[10.1016/j.eswa.2023.121832](https://doi.org/10.1016/j.eswa.2023.121832)

**Publication date**

2024

**Document Version**

Final published version

**Published in**

Expert Systems with Applications

**Citation (APA)**

Wen, M., Ibrahim, M. S., Meda, A. H., Zhang, G., & Fan, J. (2024). In-Situ early anomaly detection and remaining useful lifetime prediction for high-power white LEDs with distance and entropy-based long short-term memory recurrent neural networks. *Expert Systems with Applications*, 238, Article 121832. <https://doi.org/10.1016/j.eswa.2023.121832>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

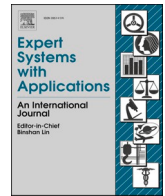
Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

***Green Open Access added to TU Delft Institutional Repository***

***'You share, we take care!' - Taverne project***

**<https://www.openaccess.nl/en/you-share-we-take-care>**

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.



# *In-Situ* early anomaly detection and remaining useful lifetime prediction for high-power white LEDs with distance and entropy-based long short-term memory recurrent neural networks

Minzhen Wen<sup>a,1</sup>, Mesfin Seid Ibrahim<sup>c,d,1</sup>, Abdulmelik Husen Meda<sup>e</sup>, Guoqi Zhang<sup>f</sup>, Jiajie Fan<sup>a,b,f,g,\*</sup>

<sup>a</sup> Institute of Future Lighting, Academy for Engineering and Technology, Fudan University, Shanghai 200433, China

<sup>b</sup> State Key Laboratory of Applied Optics, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

<sup>c</sup> Centre for Advances in Reliability and Safety, New Territories, Hong Kong

<sup>d</sup> College of Engineering, Kombolcha Institute of Technology, Wollo University, Kombolcha 208, Ethiopia

<sup>e</sup> The Hong Kong Polytechnic University, Hung Hom, Hong Kong

<sup>f</sup> EEMCS Faculty, Delft University of Technology, Delft 2628, The Netherlands

<sup>g</sup> Fudan Zhangjiang Institute, Shanghai 201203, China

## ARTICLE INFO

### Keywords:

Light-emitting diodes (LEDs)  
Mahalanobis distance (MD)  
Entropy generation rate (EGR)  
Deep Learning Algorithms  
Anomaly detection  
Remaining Useful Lifetime (RUL) Prediction

## ABSTRACT

High-power white light-emitting diodes (LEDs) have demonstrated superior efficiency and reliability compared to traditional white light sources. However, ensuring maximum performance for a prolonged lifetime use presents a significant challenge for manufacturers and end users, especially in safety-critical applications. Thus, identifying functional anomalies and predicting the remaining useful lifetime (RUL) is of enormous importance in the operational longevity of the device. To address such challenges, this study proposes a combination of distance-based Mahalanobis distance (MD), entropy generation rate (EGR), and deep learning models for improved anomaly detection and RUL prediction accuracy. Unlike conventional health indicators based on luminous flux data that are challenging to monitor relevant optical performance, the MD and EGR methods are employed to extract in-situ monitored thermal and electrical data as new health indicators. Long short-term memory recurrent neural networks (LSTM-RNN) and convolutional neural networks (CNN) are established to detect anomalies and predict the RUL. The accelerated degradation tests of 3 W high-power white LED have been conducted, and the online and offline collected experimental data are deployed for model development and performance evaluation. The performance of the proposed methods is compared against the Illuminating Engineering Society of North America (IESNA) TM-21 method. The results indicate that LSTM-RNN, when combined with either MD or EGR, can detect anomalies with significantly fewer data (70 %) than is typically required. Furthermore, a significant improvement in prediction accuracy in RUL prediction based on MD and EGR-constructed time series health indicators and employed with the LSTM-RNN model demonstrates the effectiveness of the proposed methods.

## 1. Introduction

Coupled with their superior efficiency, reliability, and versatility, the high-power white light-emitting diodes (LEDs) have shown an increased market demand and a wide range of applications, including general lighting, aerospace lighting, automotive lighting, medical applications,

and communication devices (Ibrahim et al., 2020; Pode, 2020; Schubert & Kim, 2005). However, ensuring maximum performance over a long lifetime of use presented a significant challenge for manufacturers and end users, especially in safety-critical applications. Thus, identifying operational anomalies and predicting the remaining useful lifetime (RUL) is of enormous importance in the functional longevity of the

\* Corresponding author at: Institute of Future Lighting, Academy for Engineering and Technology, Fudan University, Shanghai 200433, China.

E-mail addresses: [20210860092@fudan.edu.cn](mailto:20210860092@fudan.edu.cn) (M. Wen), [mesfin.ibrahim@cairs.hk](mailto:mesfin.ibrahim@cairs.hk) (M.S. Ibrahim), [abdulmelik.meda@connect.polyu.hk](mailto:abdulmelik.meda@connect.polyu.hk) (A.H. Meda), [G.Q.Zhang@tudelft.nl](mailto:G.Q.Zhang@tudelft.nl) (G. Zhang), [jiajie\\_fan@fudan.edu.cn](mailto:jiajie_fan@fudan.edu.cn) (J. Fan).

<sup>1</sup> They contributed equally.

<https://doi.org/10.1016/j.eswa.2023.121832>

Received 3 January 2023; Received in revised form 24 September 2023; Accepted 24 September 2023

Available online 2 October 2023

0957-4174/© 2023 Elsevier Ltd. All rights reserved.

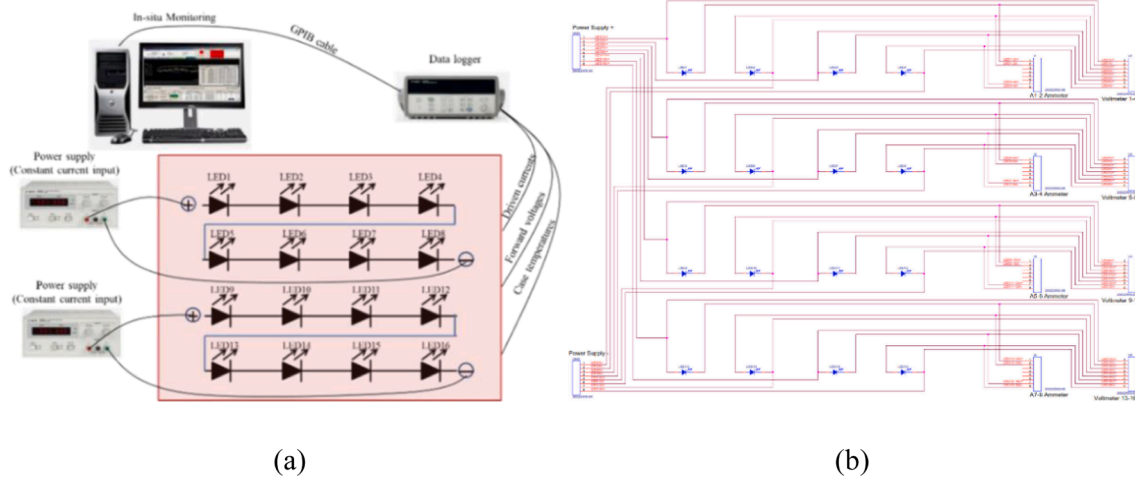


Fig. 1. The accelerated degradation test design with (a) an in-situ monitoring system and (b) a PCB test board.

device and has become one of the critical issues in the field of solid-state lighting (Driel & Fan, 2013; Ibrahim et al., 2021a; Wang, 2021).

Prognostics and health management (PHM) has become an emerging research topic getting much attention. It permits the health state assessment (reliability) of a component or complex system in its actual working condition and helps make the right decision in maintenance (Hu et al., 2022). Specifically, the purpose of prognostics is to monitor the state of health of a product, detect anomalies, and predict RUL based on the extent of deviation from its expected state of health under the conditions it is expected to be used (Wiley & Leong Gan, 2020). Existing literature categorizes the prognostics approach into physical model-based, data-driven, and hybrid/fusion methods (Chang et al., 2018; Ibrahim et al., 2022; Sun et al., 2017). The physical model-based approach involves understanding the physics of the failure mechanism and using this knowledge to formulate a mathematical model that captures the degradation trajectory of components or systems. However, this approach may not be practical when the physical processes are poorly understood or when multiple failure modes are present.

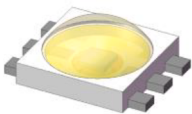
On the other hand, a data-driven approach uses operational and non-operational historical data and information collected from condition monitoring installed sensors to determine anomalies and predict future health without using any particular physical model (Si et al., 2011; Tsui et al., 2015). The performance of a data-driven approach in detecting anomalies and predicting RUL depends on having sufficient high-quality data and selecting the appropriate algorithms and techniques for data processing and modelling. Thus, data-driven approach can be a valuable alternative where physical models are difficult to develop or validate. Anomaly detection is always an essential and inevitable task for critical components in complex engineering systems. Various studies have reported the applications of conventional machine learning and deep learning models to detect anomalies in multiple domains (Li et al., 2019; Moallemi et al., 2021; Zhao et al., 2010). In the case of anomaly detection of LED products, Fan et al. (2012) used the Weibull statistical model to detect chromaticity shift anomalies. A similar study using forward voltage, drive current, and case temperature to detect anomalies in high-power white LEDs has also been reported (Fan et al., 2015a). Park and Ko (2021) also proposed an unsupervised learning-based inspection method for anomaly detection that requires unlabeled data with a micro-LED chip defect inspection dataset.

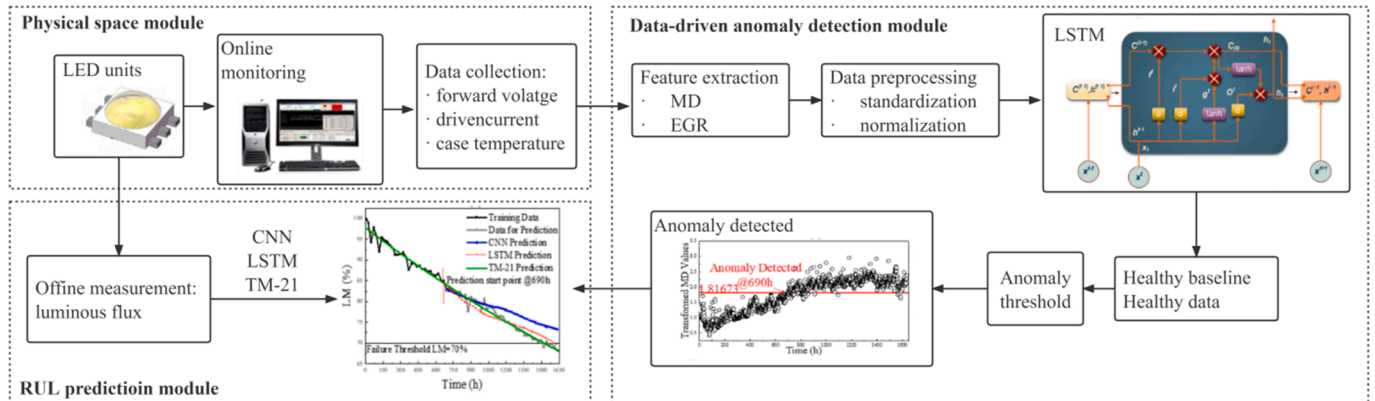
The prognostics and health management process typically involves fault diagnosis or RUL predictions after anomaly detection. Similar to anomaly detection, several studies have reported using conventional machine learning and deep learning models to predict RUL across different domains, including batteries, LEDs, power devices, semiconductor lasers and aero-propulsion systems (Abdelli et al., 2022b;

Abdelli et al., 2022a; Chen et al., 2019; Chen et al., 2020; Ibrahim et al., 2023; Jing et al., 2020; Khelif et al., 2017; Qian et al., 2021). In the RUL prediction domain of LED products, Ibrahim et al. (2019) and Fan et al. (2021) proposed the Gamma process model, Ibrahim et al. (2021b) employed a Bayesian methods combined with the Monte Carlo Markov Chain and Fan et al. (2015b) proposed a particle filter method for high-power white LEDs. Liu et al. (2019) employed two artificial neural networks to get the temperature distribution and the lifetime for a multi-chip high-power LED light source. Jing et al. (2020) fitted the radiation flux data of ultraviolet LEDs with the LSTM-RNN algorithm. These proposed methods have been compared with the Illuminating Engineering Society of North America (IESNA) TM-21 method, which demonstrated advantages in accuracy and robustness.

Various parameters have been used as LED health indicators, including photometric, colorimetric, radiometric, geometric, electrical, thermal, and others (Li et al., 2020). Photometric and colorimetric parameters are widely used for LED health indicators, where the most critical parameter is the lumen maintenance (LM) data. In addition, electrical and thermal parameters such as forward voltage, drive current, and case and junction temperatures can be used in PHM. Parameters such as those mentioned above are often used to characterize LED degradation and are synthesized from actual field data or accelerated aging test data. Photometric and colorimetric data are generally the most intuitive indicators of LED degradation (Li et al., 2020). They are typically obtained through integrating spheres, which is a great deal of inconvenience for manufacturers and users to achieve real-time monitoring. On the other hand, sensors can monitor electrical and thermal parameters in situ but are less direct. Overall, achieving efficient in-situ monitoring and establishing effective health indicators are essential to anomaly detection for high-power white LEDs. However, literature shows that most studies involving anomaly detection and RUL prediction of high-power white LED rarely consider establishing proper health indicators, which is critical for data-driven diagnostics and prognostics. Cuadras et al. (2017) proposed that optical degradation is related to thermodynamic entropy, which simultaneously considers the input electrical power, the temperature, and the optical efficiency and compared the thermoelectrical results with the optical light emission evolution during degradation. However, entropy generation rate (EGR) is a suggestive approach because LEDs' performance highly depends on operating temperature. Fan et al. (2015a) stated that the Mahalanobis distance (MD)-based anomaly detection approach could provide an early anomaly warning at around 45 % of the lifetime before actual failure happens in all test LEDs evaluated. As a preliminary attempt, our research team proposed (Wen et al., 2021) a method of EGR with Weibull distribution and LM with Back Propagation Neural Network for

**Table 1**  
Experimental Test Protocols.

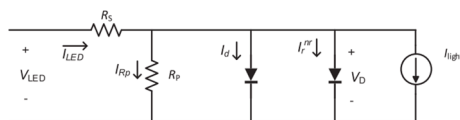
| Test samples  | Number of samples | Data collection frequency  | Online data   | Offline data  |
|---|-------------------|--|---|---|
|  | 16                | 1 h/cycle for online data<br>23 h/cycle for offline data a total of 1633 h | Forward voltage,<br>Drive current, Case temperature | Lumen maintenance, Spectra power distribution and Color coordinates |



**Fig. 2.** The flowchart of the proposed methodology implemented in this study.

**Table 2**  
The steps of the MD approach.

| <b>Algorithm 1:</b> Calculating MD values.   |
|--|
| <b>Input:</b> the in-situ monitored data during the accelerated degradation tests in the matrix $X$ . /* $p, m$ is the number of row and column size. */ |
| <b>Output:</b> MD values $MD_j$  |
| 1: <b>when</b> $i < p$ , <b>do</b>   |
| Calculating the overall mean $\bar{X}_i$ and the standard deviation $s_i$ . (-> $\bar{X}_i$ and $s_i$ )  |
| 2: <b>end when</b>   |
| 3: <b>for</b> $X_{ij}$ in matrix $X$ , <b>do</b>   |
| Performing z-score normalization. (-> $Z_j^i$ and $Z_j$ )  |
| 4: <b>end for</b>  |
| 7: <b>if</b> $p > m$ , <b>then</b>   |
| 8: <b>for</b> $X_i$ in matrix $X$ , <b>do</b>  |
| Calculating the covariance matrix $C$ . (-> $C$ and $C^{-1}$ )   |
| $Z_j^i, C^{-1}$ and $Z_j$ do multiplication to obtain the MD values. ( $(Z_j^i, C^{-1}, Z_j) \rightarrow MD$ )   |
| 9: <b>end for</b>  |
| 10: <b>end if</b>  |
| 11: <b>return MD</b>   |



**Fig. 3.** LED electrical model.

high-power white LED anomaly detection and RUL prediction. Yet, the prediction accuracy achieved was not significant.

This study proposes distance-based MD and entropy-based EGR models coupled with LSTM-RNN and CNN to improve anomaly detection accuracy and RUL prediction of high-power white LEDs. Compared to the previous study, our proposed approach's most significant contribution is detecting anomalies with considerably less data and achieving equivalent RUL prediction accuracy with 15 % less training data. This demonstrates the effectiveness of our proposed approach in providing an early estimate RUL with comparable prediction accuracy.

The procedure implemented is described as follows: first, the significance of MD and EGR in extracting in-situ monitored thermal and electrical data as new health indicators were demonstrated. A new anomaly threshold point was then defined using the LSTM-RNN model. Using the established threshold as a starting point for RUL prediction, both LSTM-RNN and CNN models were used to estimate RUL. Finally, the accuracy of the two proposed methods' prediction was compared against the IESNA TM-21 standard recommendation. The remaining parts of this paper are organized as follows: [Section 2](#) introduces the accelerated degradation test setup and data collection for a 3-W high-power white LED. [Section 3](#) presents the proposed methodology and models for anomaly detection and RUL prediction. [Section 4](#) reveals all results of the proposed methods and detailed discussion, followed by the concluding remark drawn in [section 5](#).

## 2. Accelerated degradation test and data collection

This section introduces the test samples, experimental setup, and data collection strategy employed in the accelerated degradation tests. The type of LED used in this study is a 3 W high-power white LED light source from Avago Technologies (ASMT-JN31-NTVO) with rated driven current  $I_f = 350$  mA and forward voltage  $V_F = 3.2$  V ([Broadcom, 2009](#)). The white LEDs used are InGaN-based and phosphor-converted. The accelerated degradation test design is described as follows: A DC power supply (Agilent E3611A) provides a constant current ( $I_c = 200$  mA) to the test samples. Long-term high-temperature accelerated aging was achieved by maintaining a constant aging temperature ( $T_a = 90$  °C). Gigahertz-Optik BTS256-LED tester was used to collect photometric and colorimetric data every 23 h for a total of 71 cycles (1633 h) to calculate lumen maintenance (LM) and color shift for the LEDs ([Fan et al., 2015a](#)).

As shown in [Fig. 1\(a\)](#), the electrical and thermal data, including forward voltage, drive current, and case temperature collected from the solder joint, were in-situ monitored by an Agilent-34970A data acquisition. The PCB test board was designed as shown in [Fig. 1\(b\)](#). The settings related to the experimental test protocol are shown in [Table 1](#). Another thing to keep in mind is that, even though a total of 16 samples were tested, only the data from 9 of them met the criteria for the study in

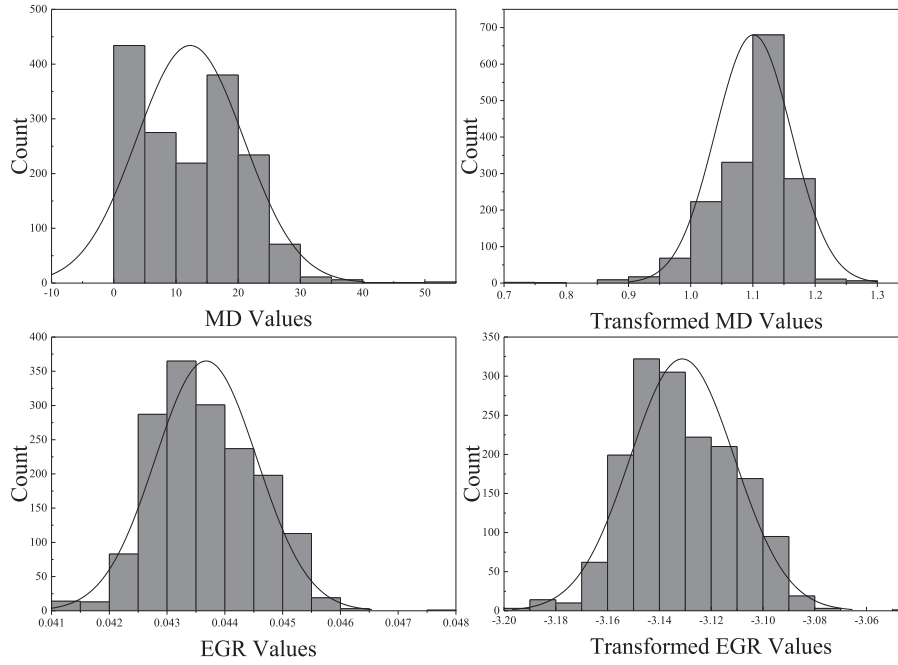


Fig. 4. Histograms of healthy MD and EGR for test LED 8# before and after the transformation.

Table 3

The steps of IESNA TM-21 standard.

| Algorithm 2: The IESNA TM-21 standard   |
|---|
| <b>Input:</b> the offline LM data of LED light sources and project long-term lifetimes.                                 |
| <b>Output:</b> the fitting curves and $L_{70}$ or $L_{80}$  |
| 1: Data pre-processing: calculating the average values of LM data for all samples, $LM = \{LM_1, LM_2, \dots, LM_t\}$ . |
| 2: Assuming $LM = \beta \exp(-at)$ , that is, a linear fit to $\ln LM$ and $t, LM' = mt + b$                            |
| 3: <b>for</b> $i = 1, \dots, t$ , <b>do</b>   |
| $LM' = \ln LM$  |
| 5: <b>end for</b>   |
| 6: Nonlinear least squares regression   |
| <b>for</b> $i = 1, \dots, t$ , <b>do</b>  |
| $sumt = sumt(\text{the initial value is } 0) + t_i$   |
| $sum' = sum'(\text{the initial value is } 0) + LM'_i$   |
| $sumt' = sumt'(\text{the initial value is } 0) + t^*LM'_i$  |
| $sumt2 = sumt2(\text{the initial value is } 0) + t^*t$  |
| <b>end for</b>  |
| $m = (t^*sumt' - sum'^2sumt) / (t^*sumt2 - sumt^2sumt)$   |
| $b = sumt' / t - m^*sumt / t$   |
| 7: $\beta = \exp(b)$ , $\alpha = -m$  |
| 8: Calculating $L_{70}$ and <b>return</b> $L_{70}$  |

this paper because the lumen maintenance data from 7 of them did not degrade as anticipated.

### 3. Methodology and model implementation

The flowchart for the proposed methods and implementation process is described in Fig. 2. First, we conducted an accelerated degradation test for high-power white LED with an in-situ monitoring design. The electrical and thermal parameters were monitored online every hour, and the optical and colorimetric data were collected offline. The data collected are pre-processed to remove noise and handle missing data. Sequential health indicators are constructed using MD and EGR approaches and are normalized and standardized to ensure the relevancy of data points for prediction. Secondly, the time series of the health indicators were modeled using LSTM-RNN models with different amounts of training data. The results of the LSTM-RNN model were evaluated with RMSE and training time to determine the minimum

amount of data required for anomaly detection, thereby establishing a threshold point for the prediction of RUL. Finally, the RUL of the devices was estimated using the LSTM-RNN and CNN models. The performances of the proposed methods were compared against the conventional IESNA TM-21 industrial standard.

#### 3.1. Health indicator extraction with in-situ monitored data

##### (I) Distance-based approach

Since MD has the advantage of condensing a multivariate system into a monolithic system and detecting anomalies based on the correlation of several parameters, they have been widely employed in monitoring the gradual degradation of electronic machinery and identifying anomalies before their occurrence (Ji, 2021). Mathematics and statistics have long used MD for dimension reduction. It provides an efficient method of calculating similarity among sample sets while considering variable correlation without regard to the measurement scale. As Fan et al. (2015a) indicated, MD could detect anomalies in all test LEDs at around 45 % of the lifetime before they failed.

The pseudo-code for the MD process is shown in Table 2, and a detailed procedure of the MD process used in this study is described as follows:

- (i) The in-situ monitored data (forward voltage  $X_1$ , driven current  $X_2$ , and case temperature  $X_3$ ) during the accelerated degradation tests are represented by the matrix  $X$ :

$$X = \begin{bmatrix} X_{11} & \dots & X_{1m} \\ \dots & X_{ij} & \dots \\ X_{p1} & \dots & X_{pm} \end{bmatrix} \quad (1)$$

The row vectors represent the data collected at the current time, and the column vectors represent the data for one performance parameter collected from the beginning to the end. Where,  $p$  denotes time spent for in-situ monitoring and  $i = 1, 2, \dots, p$ , and  $m$  denotes the number of performance parameters and  $j = 1, 2, \dots, m$ . In this work,  $p = 1633$  and  $m = 3$ .

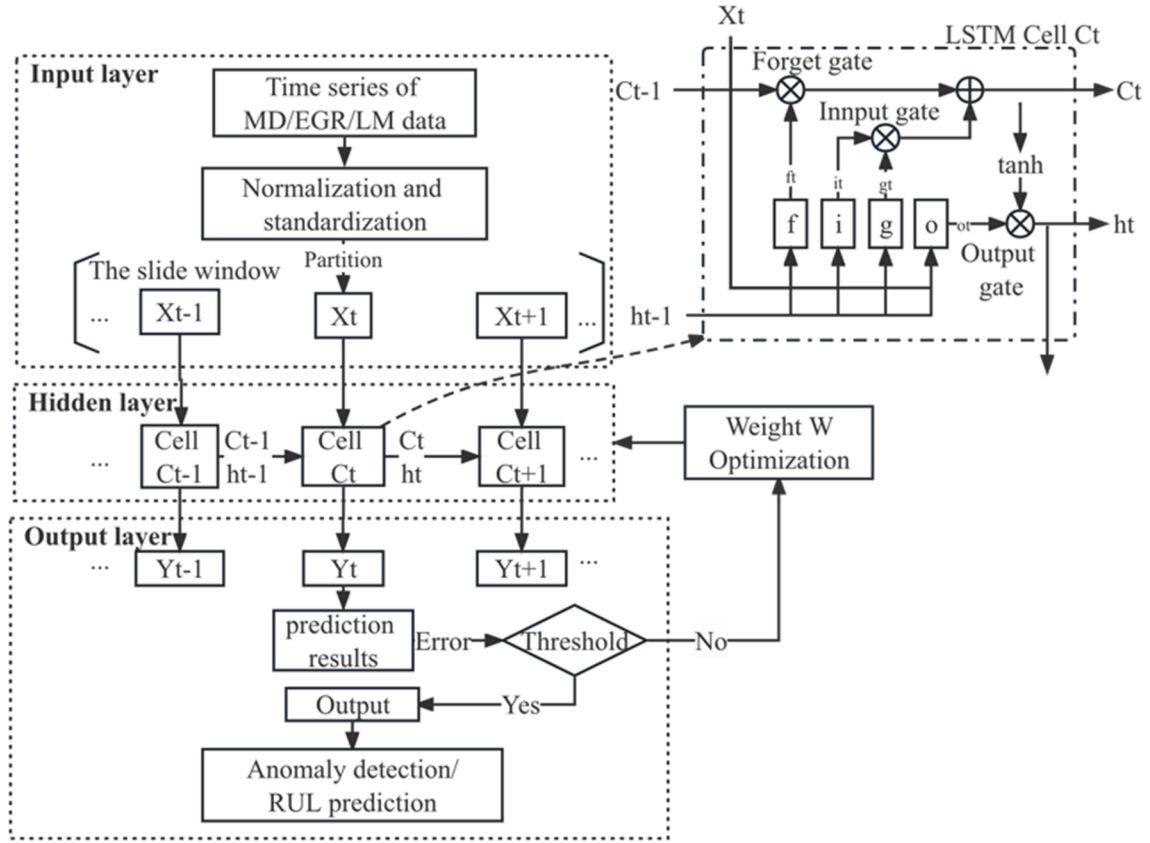


Fig. 5. The process of LSTM-RNN for anomaly detection and RUL prediction.

Table 4

The steps of LSTM-RNN model training.

**Algorithm 3:** LSTM-RNN model training algorithm

- Input:** the health indicators with in-situ monitored data during the accelerated degradation tests.
- Output:** the prediction values of the health indicators after the setting start point for prediction.
- 1: Data pre-processing: standardization and normalization.
  - 2: Data segmentation to convert time series into training and test sets for supervised learning.
  - 3: **Stage 1: Setting LSTM-RNN model to train datasets.**
    - Add data to the neuron cell in order.
    - Add the batch and slide window size to the model.
    - Set the loss and optimizer functions. This model uses the RMSE loss function and the Adam optimizer.
    - Delete the internal state and reset the epoch end to ensure the combination between each epoch.
    - Add the output of each cell and restore the optimal model parameters.
  - 4: **Stage 2: Predicting RUL by LSTM-RNN mode.**
  - 5: **for**  $t = 1, m$  **do:**
    - Callback the previously trained model.
    - Calculate the probability of forgetting the  $t$ -th hiding cell state  $f^{(t)}$ .
    - Calculate the current input  $i^{(t)}$  and update the cell state  $C^{(t)}$ .
    - Calculate the output  $O^{(t)}$  according to window size. In this model, use 0-60th data to prediction 61st values.
    - Restore the output data and convert the results into the array.
    - Invert the predicted data to keep the same form as the original data.
  - 6: **end for**
  - 7: Evaluate the training time and RMSE
  - 8: **return the LSTM-RNN model**

- (ii) The overall mean  $\bar{X}_i$  of each sample and the standard deviation  $s_i$  are as given in Equation (2) and (3):

$$\bar{X}_i = \frac{1}{m-1} \sum_{j=1}^m X_{ij} \quad (2)$$

$$S_i = \sqrt{\frac{\sum_{j=1}^m (X_{ij} - \bar{X}_i)^2}{m-1}} \quad (3)$$

- (iii) Perform z-score normalization and calculate the covariance matrix  $C$  according to Equation (4) and (5):

$$Z_{ij} = \frac{(X_{ij} - \bar{X}_i)}{S_i} \quad (4)$$

$$C = \frac{1}{m-1} \sum_{j=1}^m Z_j Z_j^T \quad (5)$$

- (iv) MD values could be calculated with Equation (6):

$$MD_j = \frac{1}{p} Z_j^T C^{-1} Z_j \quad (6)$$

In this work, we calculate the results from each sample point to the center of the data.

## (II) Entropy-based approach

The existing LED reliability model has not established a clear relationship between degradation parameters (Cuadras et al., 2016). Entropy from thermodynamics is considering the input electric power,

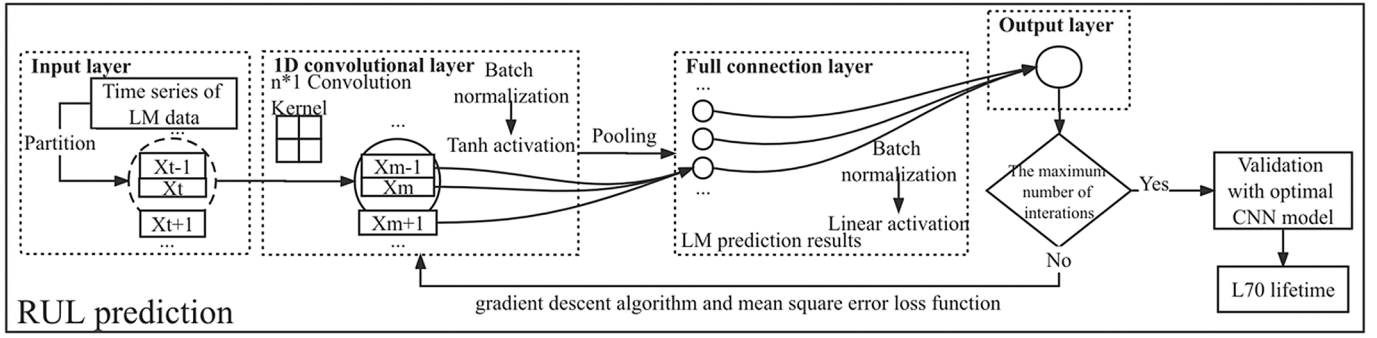


Fig. 6. The 1D CNN model architecture.

Table 5

The steps of 1D CNN model training.

| Algorithm 4: CNN model training algorithm   |
|---|
| <b>Input:</b> the health indicators with in-situ monitored data during the accelerated degradation tests.   |
| <b>Output:</b> the prediction values of the health indicators after the setting start point for prediction.   |
| 1: Data pre-processing: standardization and normalization.  |
| 2: Convert 1D time series into more samples with training and test sets.  |
| 3: <b>Stage 1: Setting 1D CNN model to train datasets.</b>  |
| Add data after segmenting to the input layer in order.  |
| Set timesteps and features. This model's timestep size is set to 60, and the feature size is 1.   |
| Add the 1D convolutional, pooling, and full-connection layers. In the 1D convolutional layer, window length*1 kernel and tanh activation are used, and the parameter is calculated: |
| In the full connection layer, linear activation is used.  |
| 7: Set the loss and optimizer function. In this model, the loss function is set to RMSE and the optimizer is adam with its parameter tuned.   |
| 8: Add the output of each cell and restore the optimal model parameters.  |
| 9: <b>Stage 2: Predicting time series by 1D CNN model.</b>  |
| 10: Obtain the vector $a^1$ according to P and the edge of the time series.   |
| 11: Initial the parameters of hiding layers W and b.  |
| 12: <b>for</b> $l = 2$ to $L - 1$ , <b>do</b> :   |
| Callback the previously trained model.  |
| Calculate the output of the current layer.  |
| 13: <b>end for</b>  |
| 18: Store the output of the whole 1D CNN model  |
| 19: Evaluate the training time and RMSE   |
| 20: <b>return the 1D CNN model</b>  |

temperature, and optical efficiency. It has been applied to mechanical (Basaran et al., 2003) and electrical (Amiri & Modarres, 2014) damage characterization. Cuadras et al. (2016) suggested that LED degradation is related to thermodynamic entropy. A typical model of an LED is illustrated in Fig. 3. The ideal diodes represent diffusion ( $I_d$ ) and non-radiative recombination ( $I_r^{nr}$ ).  $R_s$  and  $R_p$  are the dissipative elements, and  $I_{light}$  stands for light emission. A thermodynamic approach to the LEDs' failure can be described by introducing the first and second laws of thermodynamics as follows (Grundmann, 2006):

$$dE_{in} = dW + dQ + dE_{irr} \quad (7)$$

$$\dot{S} = \dot{S}_e + \dot{S}_i \quad (8)$$

With

$$\dot{S}_i \geq 0 \quad (9)$$

The dot over the variable represents a time derivative, i.e.,  $\dot{S}_i = dS/dt$ . In Equation (10),  $E_{in}$  is the electrical input energy delivered to the LED;  $W$  is the light emitted energy;  $Q$  is the dissipated heat; and  $E_{irr}$  is the energy devoted to causing irreversible damage. In Equation (8), sub-indices  $e$  and  $i$  refer to external (entropy exchange) and internal entropy (entropy generation), respectively.

Input energy  $E_{in}$  is usually expressed in terms of the total input injected power  $P$

$$E_{in} = \int P dt = \int V_{LED} I_{LED} dt \quad (10)$$

where  $I_{LED}$  and  $V_{LED}$  are the current and voltage drop between the external terminals, as shown in Fig. 3. Input power is split into light emission and dissipative terms.

Heat dissipation about the resistances due to the Joule effect is depicted in Fig. 3:

$$Q_d = R_s I_{LED}^2 + \frac{V_D^2}{R_p} \quad (11)$$

For commercial display LEDs, wall-plug efficiency (ratio of input energy (in W) to output energy (in W)) is small) is less than 0.1 %, and  $E_{irr}$  form is unknown at present. In terms of entropy generation rate,  $\dot{S}_i = dS/dt$ , for electrical systems, it is commonly written as:

$$\dot{S}_i = \frac{P}{T} = \frac{V_{LED} I_{LED}}{T_j} \quad (12)$$

After collecting the original in-situ monitored data, we calculated their MDs and EGRs values according to Equations (6) and (12). However, as demonstrated in Fig. 4 with sample LED#8, the histogram of all data does not follow a normal distribution. Therefore, a typical transformation of the original data is needed. Exponential, logarithmic, and power transformations are usually used to deal with kurtosis and skewness to achieve data normalization.

In the preliminary preparation, we performed multiple standard normal transformations (BOX-COX, logarithmic, square root, square root inverse sine, square, and inverse transformation) of MD and EGR data. It is found that the BOX-COX transformation for MD data makes the MD data compliant with the standard and normal distribution, and the logarithmic conversion for EGR data is more compliant. Therefore, we perform a power transformation (Box-Cox transformation) on the MD data (Fan et al., 2015a) and a logarithmic transformation on the EGR data. The Box-Cox transformation is calculated as follows:

$$y(\lambda) = \begin{cases} (y^{\lambda} - 1)/\lambda; & \lambda \neq 0 \\ \log y; & \lambda = 0 \end{cases} \quad (13)$$

Where  $y(\lambda)$  is the new variable obtained after Box-Cox transformation;  $y$  is the original continuous dependent variable;  $\lambda$  is a transformation parameter.

The logarithmic transformation is calculated as follows:

$$y' = \ln y, y > 0 \quad (14)$$

Where  $y'$  and  $y$  are the values before and after logarithmic transformation.



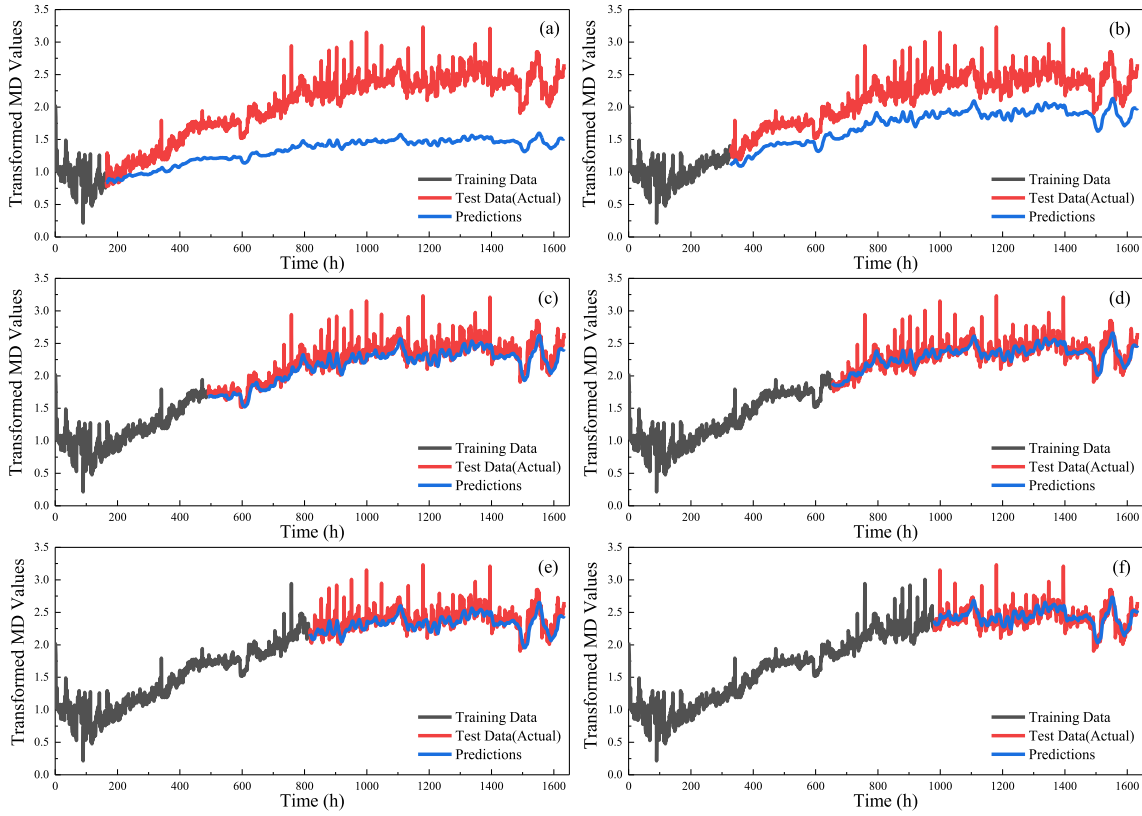


Fig. 7. The LSTM-RNN prediction on transformed MD values of test sample LED#8 with different training datasets.

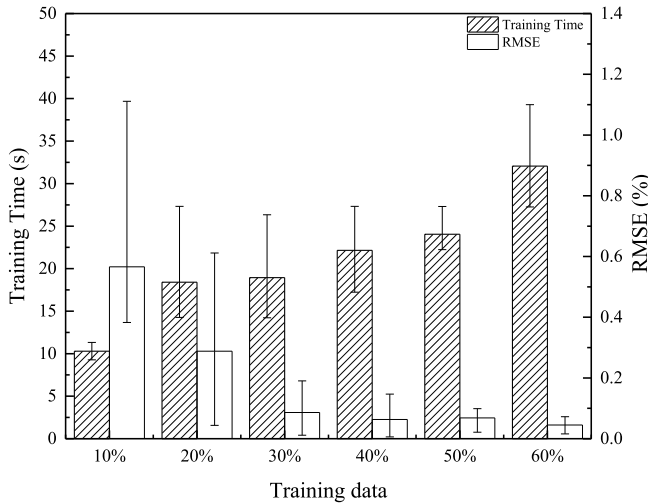


Fig. 8. The MD-LSTM-RNN training results with training time and RMSE.

### 3.2. Anomaly detection and RUL prediction

#### (I) The IESNA TM-21 standard

The IESNA released the standard IES-21-11 to predict the lumen lifetime for LED light sources based on the lumen maintenance data collected from the IES LM-80-08 test report. The TM-21 standard is commonly used in the lighting industry to deal with the LM of LED light sources and project long-term lifetimes. Among the two failure thresholds, LEDs LM decay to 80 % and 70 % of the initial value; 70 % is adopted for this study. The pseudo-code for the TM-21 standard is shown in Table 3, and the specific process of the IESNA TM-21 method is

explained as follows (IESNA, 2011):

- (i) Normalize the luminous flux data of LEDs and take the average value to obtain the mean LM of all samples.
- (ii) Use the nonlinear least squares regression method to curve-fit the mean LM data according to Equation (15):

$$LM(t) = \beta \exp(-\alpha t) \tag{15}$$

where  $LM(t)$  is the LM data;  $\beta$  is the initial coefficient of curve fitting, and  $\alpha$  is the derived coefficient of curve fitting,  $\alpha > 0$ ;  $t$  is operation time in the aging process. Here, parameters  $\alpha$  and  $\beta$  are estimated from historical (or experimental data) using the least-squares regression method.

- (iii) Perform the logarithmic transformation on Equation (15) to obtain the lifetime prediction model. The results of  $L_{70}$  in Equation (16) is the lifetime when the white LED's LM decays to 70 % of the initial value.

$$L_{70} = \ln(\beta/0.7)/\alpha \tag{16}$$

#### (II) LSTM-RNN Model

Standard RNNs are inefficient for learning long-term dependencies along degradation patterns due to the network vanishing gradient. Hochreiter and Schmidhuber (1997) introduced the LSTM-RNN model to overcome the shortcomings of long-term dependencies in RNN. This model has been applied to various sequence learning tasks, including handwriting recognition, speech recognition, and sentiment analysis. As described earlier, the application of LSTM-RNN is also extended to PHM of batteries (Catelani et al., 2021; Wong et al., 2021), bearings (Hinchi & Tkiouat, 2018), traffic prediction (Abduljabbar et al., 2021), and other products. LSTM is a special RNN with a gate structure, which is the key to updating or discarding the output data through the logic control of gate units. It overcomes the shortcomings of traditional RNN, such as the

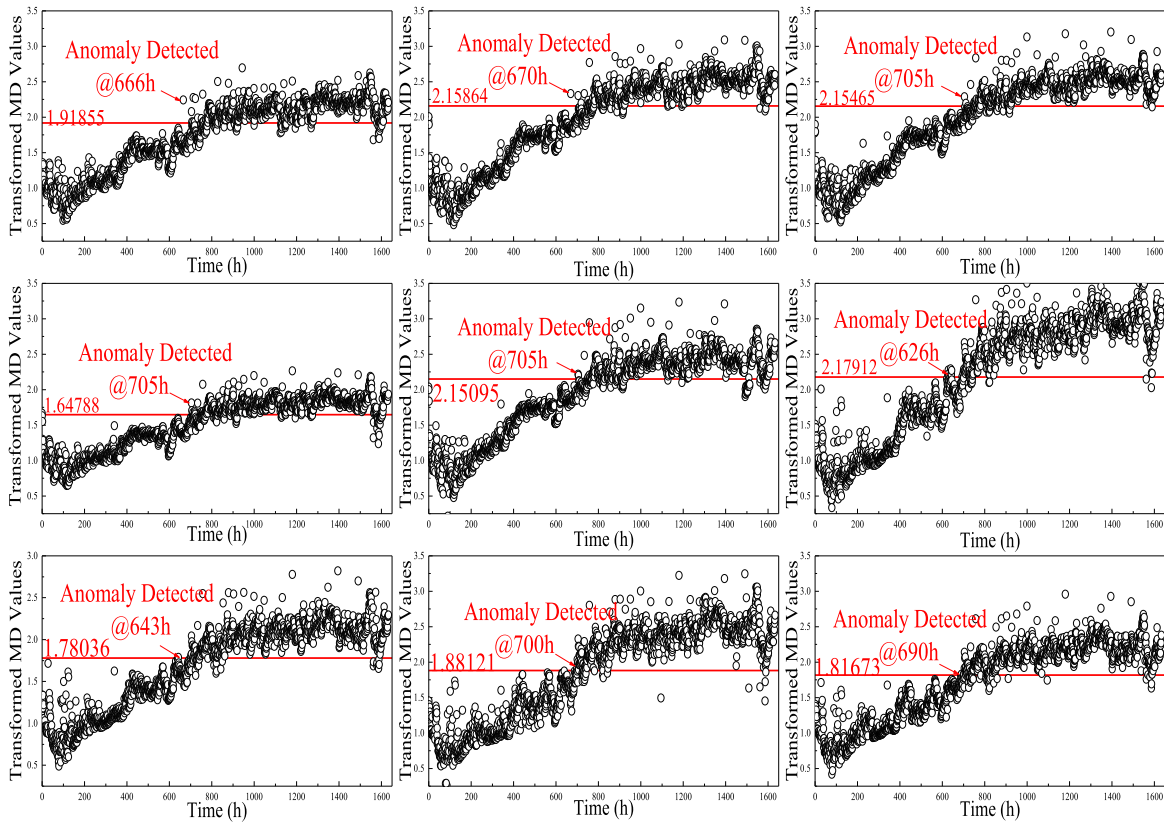


Fig. 9. Anomaly detection results with transformed MD values of all test samples.

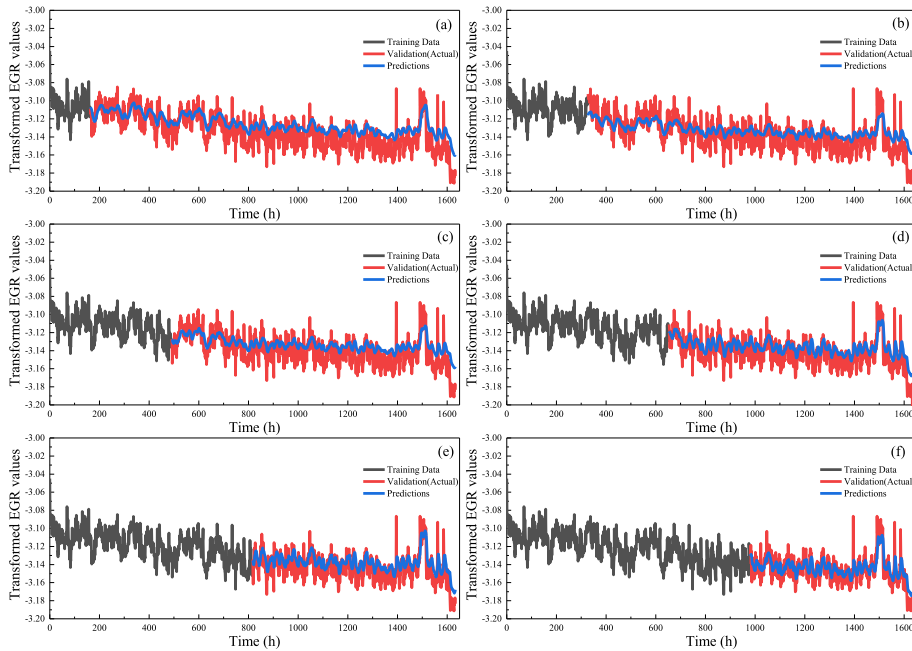


Fig. 10. The LSTM-RNN prediction on transformed EGR values of test sample LED#8 with different training datasets.

excessive influence of weights, gradient vanishing, and explosion.

Furthermore, it allows the network to converge faster to help effectively improve prediction accuracy. Based on the historical health indicator data of LEDs, the relationship between the effect of the LED performance degradation and the health indicators should be continuously explored, and the corresponding dynamic indexes should be established based on this effect. The process of the LSTM-RNN model for

anomaly detection and RUL prediction proposed in this study and the unit network structure is shown in Fig. 5. Internal state (C) is a key to LSTM-RNN, which is considered as the heart of each time series sequence and as a self-connected recurrent edge with fixed unit weight. It can be regarded as a carrier to which information has been added or from which it has been removed. A distinctive structure of this approach, known as gates, carefully regulates the flow of information. In doing so,

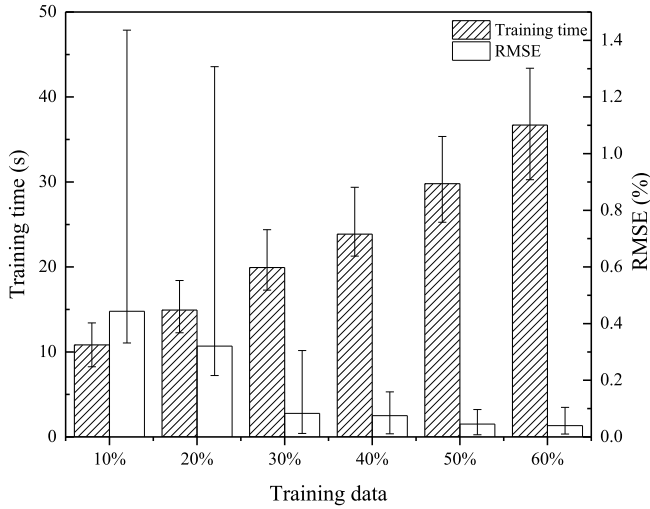


Fig. 11. The EGR-LSTM-RNN training results with training time and RMSE.

it decides which information should pass through and which should not. It is a sigmoidal unit that is activated from the current input layer  $x^{(t)}$  as well as from the hidden layer at the previous time step. The forget gate ( $f_c$ ) is a sigmoid layer that decides the information needs to be discarded. It takes health indicator inputs of  $x^{(t)}$  and  $h^{(t-1)}$  and outputs a number between 0 representing the value to be discarded and 1 representing the value to be retained for each internal state  $C$ . The discarded information can be the outliers, noises, and redundant information of degradation data among adjacent cycles. The forget gate is calculated as Equation (17):

$$f^{(t)} = \sigma(W^{fx}x^{(t)} + W^{fh}h^{(t-1)} + b_f) \quad (17)$$

In this work, the input data are all one-dimension time series. The input gate and input node decide what new information will be stored in the internal state. This step has two parts: first, a sigmoid layer called the “input gate”,  $i_c$ , decides which values to update; then, a  $\tanh$  layer called the “input node”,  $g_c$ , creates a vector of the new candidate state  $\hat{c}^{(t)}$  which could be added to the state.

$$i^{(t)} = \sigma(W^{ix}x^{(t)} + W^{ih}h^{(t-1)} + b_i) \quad (18)$$

$$g^{(t)} = \tanh(W^{gx}x^{(t)} + W^{gh}h^{(t-1)} + b_g) \quad (19)$$

Combining (17)–(19) to update the previous internal state  $C^{(t-1)}$  into the current state  $c^{(t)}$  as shown in:

$$C^{(t)} = g^{(t)} * x^{(t)} + f^{(t)} C^{(t-1)} \quad (20)$$

Finally, there is a sigmoid layer called the “output gate”,  $O^{(t)}$ , which determines what information to output. After putting the internal state  $C$  through a  $\tanh$  layer, it is multiplied by the output of the sigmoid gate to obtain the remaining state value. This can be implemented as:

$$o^{(t)} = \sigma(W^{ox}x^{(t)} + W^{oh}h^{(t-1)} + b_o) \quad (21)$$

$$h^{(t)} = \tanh(C^{(t)}) * O^{(t)} \quad (22)$$

Where  $W$  and  $b$  values are learnable parameters corresponding to the layer weights and biases, respectively. After the model has been trained, weight optimization is required to reduce loss and error, and the optimal parameters are then stored and used to validate the new model. The test dataset is fed into the optimal parameter model to obtain optimized data.

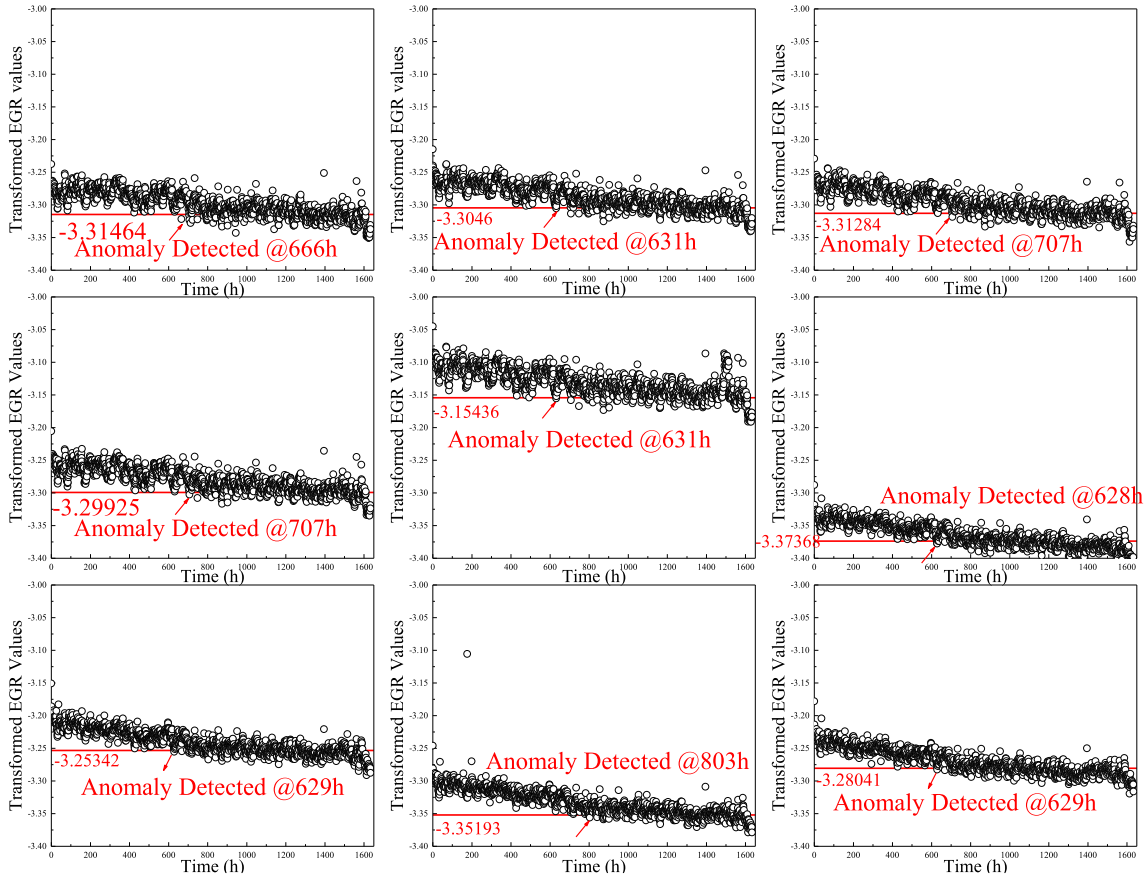


Fig. 12. Anomaly detection results with transformed EGR values of all test samples.

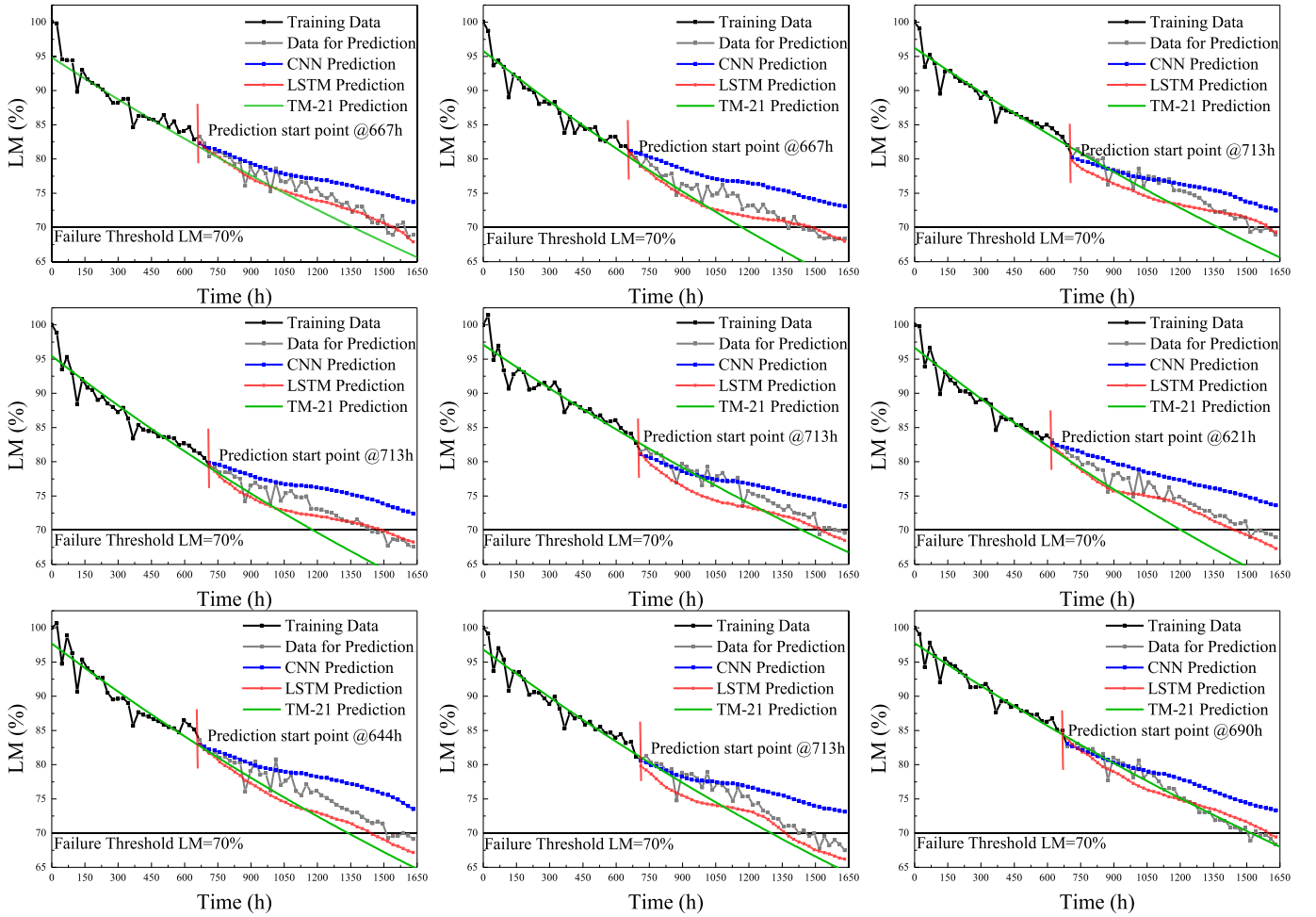


Fig. 13. The LM RUL prediction results of all test samples with different MD-LSTM-RNN anomaly detections.

In this study, all model configuration and simulation are carried out using Python 3.8 version on Windows 10 personal computer built with an Intel Core i5-6200U processor (4 MB cache, up to 2.40 GHz), and a graphic card of Intel HD Graphics 520 at 4 GB. Accordingly the parameters are used in developing the set for the model are listed: (i) the window size = 60; (ii) the activation function is tanh, and when in output fully-connected layers, the activation function is linear; (iii) the regularization rate (dropout) for each layer is set 0.2 to avoid overfitting; (iv) RMSE is selected to calculate prediction errors; (v) batch size = 100; (vi) the optimization method is Adam, and the learning rate is set the default values 0.001; (vii) the RMSE and training time are used as the indicators of model performance evaluation. The pseudo-code for LSTM recurrent neural network model is shown in Table 4.

### (III) 1D CNN model

CNN can be seen as a modification of traditional neural networks, which use a hierarchical network structure capable of learning many mapping relations. As illustrated in Fig. 6, 1D CNN for RUL prediction comprises the input (LM time series) layer, convolution layer, pooling, fully connected layer, and output layer. We opted to use the  $\tanh$  function as the activation function in convolutional layers and the linear function in the fully connected layer. Convolutional neural networks are mainly used for image recognition tasks, which usually involve  $2 \times 2$  convolution to extract meaningful features from the input images or videos. In this work, the input data is one-dimensional. After each convolution layer, a batch normalization technique is applied to improve the performance and stability of the CNN mode. The pooling

layer reduces the model dimensionality by scaling the data in the upper layer, keeps the data features scaled invariant, and prevents the data from overfitting. Finally, the output layer is fully connected.

Moreover, the stride sizes in the convolutional and pooling layers are artificially set to 1. We used the adaptive moment estimation (Adam) optimizer as a preferred choice of optimizer as it has a combined advantage of root mean square propagation (RMSProp) and adaptive gradient algorithm (AdaGrad) techniques. The Adam optimizer is applied with a learning rate of 0.001. Then, in the training process, these parameters are updated based on a gradient descent algorithm and mean square error loss function. A pseudo code and flowchart of the 1D CNN model for RUL prediction is shown in Table 5 and Fig. 6 respectively. The parameter in convolutional layers is calculated based on Equation (23):

$$N = \frac{W - F + 2P}{S} + 1 \quad (23)$$

Where  $N$  is output size,  $W$  is the number of inputs,  $F$  is the number of convolutional kernels,  $S$  is the number of steps (stride), and  $P$  is the padding value; i.e., in our dataset,  $W = 1633$ ,  $F = 2$ ,  $P = 1$ , and  $S = 60$ ,  $N = 28$ .

The following three cases should exist for the output of the current layer: when the layer is convolutional layer, the output is obtained according to Equation (24); when is pooling or full connection, it is by Equation (25) and (26), respectively.

$$a^l = \sigma(a^{l-1} * W^l + b^l) \quad (24)$$

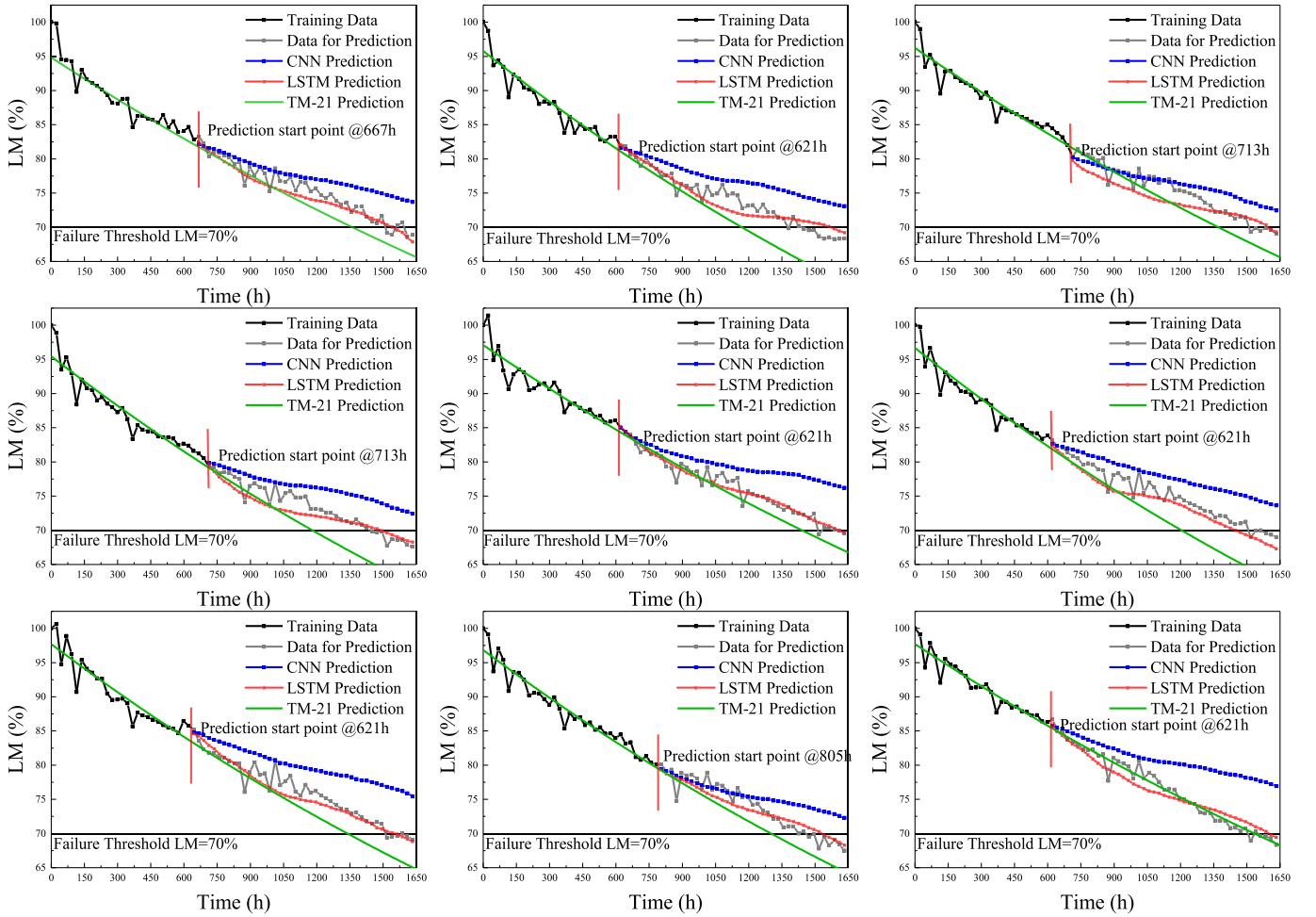


Fig. 14. The LM RUL prediction results of all test samples with different EGR-LSTM-RNN anomaly detections.

$$a^l = \text{pool}(a^{l-1}) \quad (25)$$

$$a^l = \sigma(W^l a^{l-1} + b^l) \quad (26)$$

where  $\sigma$  is the activation function.

Finally, the output of the whole model output layer is given in Equation (27):

$$a^l = \text{softmax}(W^l a^{l-1} + b^l) \quad (27)$$

## 4. Results and discussion

### 4.1. In-situ anomaly detection results and discussion

To reduce the computational overhead and improve the accuracy of the network model, we normalized and standardized the input data. Then, we defined the same window length of 60 for the input data points in both MD and EGR approaches. Accordingly, 0-60th input data sequences were used to predict the 61st value. Finally, the predicted results are compared against the actual values.

#### (I) MD-LSTM-RNN

The training data was selected as gradually decreased from 60 % to 10 % of the total 1633 h. The LSTM-RNN prediction results of the LED 8# MD values are displayed in Fig. 7, in which the black and red lines represent the transformed data, and the blue lines are the predicted results. The MD values generally increase but fluctuate slightly. A

randomly selected nine sets of test samples were modeled with training data that was increased from 10 % to 60 %.

Fig. 8 shows the MD LSTM-RNN model training results for all test samples. As the amount of training data gradually decreases to 30 %, the RMSE tends to stabilize. Thus, it was determined that 30 % of the overall data should be used for MD-based anomaly detection. Only 0-490 h of data is sufficient for the anomaly detection threshold. The anomaly detection thresholds defined as  $(\mu - 3\sigma$  or  $\mu + 3\sigma)$  (Fan et al., 2015a) were calculated according to the mean ( $\mu$ ) and standard deviation ( $\sigma$ ) values from the 0-490 h (0-30 %) MD values. Finally, anomalies were detected, and the starting point of the RUL prediction was determined for each of the nine test samples, as shown in Fig. 9. Although all the anomalies detected are concentrated between 600 and 700 h, there are rise and fall in the time anomaly noticed between samples. This issue is most likely due to sample variation.

#### (II) EGR-LSTM-RNN

Instead of BOX-COX transformation, the logarithmical transformation was used in EGR data preprocessing. We followed a similar procedure as described in the above section. The LSTM-RNN prediction on transformed EGR values of test sample LED#8 with different training datasets is depicted in Fig. 10. EGR values generally exhibit a downward trend on the negative axis. As a result of substantial thermal energy being wasted during accelerated degradation tests, the EGR data demonstrates a non-monotonic increase.

Similarly, The EGR LSTM-RNN model training results for all test samples are shown in Fig. 11. The RMSE tends to stabilize as the amount

**Table 6**

The LM prediction results and prediction error with MD and EGR-based health indicators.

| Test Samples | Method | $L_{70}$ Prediction/h |                       | Prediction Error/%   |                       |
|--------------|--------|-----------------------|-----------------------|----------------------|-----------------------|
|              |        | MD Health indicators  | EGR Health indicators | MD Health indicators | EGR Health indicators |
| LED#1        | Actual | 1518                  | 1518                  | –                    | –                     |
|              | IESNA  | 1363                  | 1363                  | –10.2108             | –10.2108              |
|              | TM-21  |                       |                       |                      |                       |
|              | CNN    | $L_{74} = 1610$       | $L_{74} = 1610$       | 6.0606               | 6.0606                |
| LED#2        | Actual | 1380                  | 1380                  | –                    | –                     |
|              | IESNA  | 1170                  | 1167                  | –15.2174             | –15.2174              |
|              | TM-21  |                       |                       |                      |                       |
|              | CNN    | $L_{74} = 1518$       | $L_{74} = 1518$       | 10                   | 10                    |
| LED#3        | Actual | 1518                  | 1518                  | –                    | –                     |
|              | IESNA  | 1373                  | 1373                  | –9.5520              | –9.5520               |
|              | TM-21  |                       |                       |                      |                       |
|              | CNN    | $L_{73} = 1587$       | $L_{73} = 1587$       | 4.5455               | 4.5455                |
| LED#4        | Actual | 1449                  | 1449                  | –                    | –                     |
|              | IESNA  | 1181                  | 1181                  | –18.4955             | –18.4955              |
|              | TM-21  |                       |                       |                      |                       |
|              | CNN    | $L_{73} = 1587$       | $L_{73} = 1587$       | 9.5238               | 9.5238                |
| LED#5        | Actual | 1518                  | 1518                  | –                    | –                     |
|              | IESNA  | 1442                  | 1442                  | –5.0066              | –5.0066               |
|              | TM-21  |                       |                       |                      |                       |
|              | CNN    | $L_{74} = 1587$       | $L_{73} = 1587$       | 4.5455               | 4.5455                |
| LED#6        | Actual | 1518                  | 1518                  | –                    | –                     |
|              | IESNA  | 1206                  | 1206                  | –20.5534             | –20.5534              |
|              | TM-21  |                       |                       |                      |                       |
|              | CNN    | $L_{74} = 1610$       | $L_{74} = 1610$       | 6.0606               | 6.0606                |
| LED#7        | Actual | 1518                  | 1518                  | –                    | –                     |
|              | IESNA  | 1339                  | 1340                  | –11.7918             | –11.7260              |
|              | TM-21  |                       |                       |                      |                       |
|              | CNN    | $L_{74} = 1610$       | $L_{76} = 1610$       | 6.0606               | 6.0606                |
| LED#8        | Actual | 1426                  | 1426                  | –                    | –                     |
|              | IESNA  | 1302                  | 1302                  | –8.6957              | –8.6957               |
|              | TM-21  |                       |                       |                      |                       |
|              | CNN    | $L_{74} = 1518$       | $L_{73} = 1587$       | 6.4516               | 11.2903               |
| LED#9        | Actual | 1518                  | 1518                  | –                    | –                     |
|              | IESNA  | 1520                  | 1534                  | 0.1318               | 1.0540                |
|              | TM-21  |                       |                       |                      |                       |
|              | CNN    | $L_{74} = 1564$       | $L_{75} = 1633$       | 3.0303               | 7.5758                |

of training data gradually decreases to 30 %. Therefore, 30 percent of the overall data should also be used for detecting anomalies using EGR. The anomaly detection thresholds were calculated according to  $\mu - 3\sigma$  or  $\mu + 3\sigma$  from the 0–490 h (0–30 %) EGR values. Finally, we integrated the times the anomalies were detected of all 9 test samples into Fig. 12 correspondingly. Compared with the results of the times when the MD-LSTM-RNN and EGR-LSTM-RNN detected the anomalies, it was found that most samples did not have many peaks and troughs, except for LED 8#. With that, it was found that before the anomaly was detected, there were many discrete points and noisy values in LED #8, nevertheless we have not handled it when calculating, which might essentially affect the final anomaly threshold results.

The above results indicate that using EGR and MD for LED anomaly

detection significantly reduces the amount of accelerated degradation test data required. As demonstrated for high-power white LEDs, the EGR and MD health indicators can be used to detect LED degradation anomalies. However, the EGR-based approach detected anomalies earlier than the MD-based approach in a significant proportion of samples. This phenomenon does not indicate that EGR is more suitable than MD as a health indicator to characterize white LED degradation. It is widely known that the physical failure model of LEDs is so complex that there is no clear relationship between the performance degradation parameters. The EGR is simply the product of the relevant factors.

#### 4.2. RUL prediction results and discussion

We used the anomaly detection time mentioned in Section 4.1 for RUL prediction. IESNA TM-21 considers lumen degradation as the most common failure mode of LED light sources (IESNA, 2011). The lumen maintenance  $L_{70}$  lifetime is defined as the operating time that 70 % of the luminous flux from its initial light output for general applications. Based on the anomaly detected times obtained from MD and EGR values of each sample, the LM data were imported into the CNN and LSTM-RNN models for RUL ( $L_{70}$ ) prediction, and the prediction errors were then calculated as compared to the actual  $L_{70}$  lifetimes. The RUL prediction results of the two models MD-LSTM-RNN and EGR-LSTM-RNN anomaly detections, are shown in Fig. 13 and Fig. 14, respectively. The results are compared against the CNN-based model and the IESNA TM-21 method. It can be noted that the CNN model has the least prediction accuracy (higher prediction error) results, followed by the IESNA TM-21 method, and the LSTM-RNN model is the best, regardless of the starting point of RUL prediction with MD or EGR.

A summary of the results obtained by the IESNA TM-21 and other methods are listed in Table 6. Since the  $L_{70}$  lifetimes could not be obtained from the prediction curves of the CNN, only  $L_{73}$ ,  $L_{74}$ , or  $L_{75}$  lifetimes are recorded here. A starting point with more data points for the prediction will also result in a lower prediction error. The results depicted in Fig. 13, Fig. 14, and Table 6 demonstrated that RUL prediction based on LSTM-RNN provided better prediction accuracy than its counterpart CNN and IESNA TM-21. The essential difference between CNN and LSTM-RNN is the presence or absence of a memory function and the ability to associate. In LSTM-RNN, each time series sequence selectively forgets and discards the information from the previous moment to keep updating. Although CNN can recognize spatial correlation in data, it does not perform as well with sequential data. In convolutional and fully connected layers of CNN, the weighted summation method is suitable for extracting local information, lacking in the before-and-after correlation of time series sequences. The IESNA TM-21 standard is a nonlinear least squares regression fitting with the data before the prediction start point. In this case, the LSTM-RNN model is more suitable for large-scale data. Another significant achievement of the LSTM-RNN model over the IESNA TM-21 methods is that it only needs some sequences with fixed lengths to make better predictions than IESNA TM-21 using all available data points. The proposed methods achieved equivalent prediction accuracy with 15 % less training data when compared to related works by Ibrahim et al. (2021b) and Fan et al. (2015b). In this work, to achieve a prediction error of 10 % or less, the minimum training data required for MD-LSTM-RNN or EGR-LSTM-RNN is 30 % of the experimental data. The comparison showed that the starting point of RUL prediction is about 20 % earlier than other studies. In other words, this study achieves a better prediction error with the premise that the RUL prediction is performed as early as possible.

## 5. Conclusions

In this study, the performance degradation of high-power white LEDs was evaluated by the in-situ monitored electro-thermal data from accelerated degradation tests. Since the most obvious indicator of LED degradation is lumen flux data, two new health indicator metrics, MD

and EGR, were proposed and feature-extracted by LSTM-RNN to achieve in-situ early anomaly detection. After the timely anomaly detection, the threshold was used as the starting point to perform RUL prediction. With only 30 % of the experimental data, the proposed methods achieved a 10 % prediction accuracy, an improvement of 15 % over the related works. The EGR-based and MD-based approaches as health indicators constructors for LED anomaly detection have been shown to significantly reduce the amount of accelerated degradation test data required, providing a viable solution to achieve dynamic LED failure diagnosis.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgments

The work described in this paper was supported by the National Natural Science Foundation of China (52275559, 51805147), State Key Laboratory of Applied Optics (SKLAO2022001A01), Shanghai Science and Technology Development Foundation (21DZ2205200), Shanghai Pujiang Program (2021PJD002) and Centre for Advances in Reliability and Safety (CAiRS) admitted under AIR@InnoHK Research Cluster.

## References

- Abdelli, K., Grießer, H., Neumeier, C., Hohenleitner, R., & Pachnicke, S. (2022). Degradation Prediction of Semiconductor Lasers Using Conditional Variational Autoencoder. *Journal of Lightwave Technology*, 40(18), 6213–6221. <https://doi.org/10.1109/JLT.2022.3188831>
- Abdelli, K., Grießer, H., & Pachnicke, S. (2022). A Machine Learning-Based Framework for Predictive Maintenance of Semiconductor Laser for Optical Communication. *Journal of Lightwave Technology*, 40(14), 4698–4708. <https://doi.org/10.1109/JLT.2022.3163579>
- Abduljabbar, R. L., Dia, H., & Tsai, P. W. (2021). Unidirectional and bidirectional LSTM models for short-term traffic prediction. *Journal of Advanced Transportation*, 2021. <https://doi.org/10.1155/2021/5589075>
- Amiri, M., & Modarres, M. (2014). An Entropy-Based Damage Characterization. *Entropy*, 16(12), 6434–6463. <https://doi.org/10.3390/e16126434>
- Basaran, C., Lin, M., & Ye, H. (2003). A thermodynamic model for electrical current induced damage. *International Journal of Solids and Structures*, 40(26), 7315–7327. <https://doi.org/10.1016/j.jislsr.2003.08.018>
- Broadcom. (2009). ASMT-Jx3x 3 W Mini Power LED Light Source. <https://www.mouser.com/new/broadcom/broadcom-asmt-jx3x-leds/>.
- Catelani, M., Ciani, L., Fantacci, R., Patrizi, G., & Picano, B. (2021). Remaining Useful Life Estimation for Prognostics of Lithium-Ion Batteries Based on Recurrent Neural Network. *IEEE Transactions on Instrumentation and Measurement*, 70. <https://doi.org/10.1109/TIM.2021.3111009>
- Chang, M.-H., Fan, J., Qian, C., & Sun, B. (2018). PHM of Light-Emitting Diodes. *Prognostics and Health Management of Electronics*. <https://doi.org/10.1002/9781119515326.CH14>
- Chen, J., Jing, H., Chang, Y., & Liu, Q. (2019). Gated recurrent unit based recurrent neural network for remaining useful life prediction of nonlinear deterioration process. *Reliability Engineering & System Safety*, 185, 372–382. <https://doi.org/10.1016/j.ress.2019.01.006>
- Chen, X., Jin, G., Qiu, S., Lu, M., & Yu, D. (2020). Direct Remaining Useful Life Estimation Based on Random Forest Regression. *2020 Global Reliability and Prognostics and Health Management, PHM-Shanghai 2020*. <https://doi.org/10.1109/PHM-SHANGHAI49105.2020.9281004>.
- Cuadras, A., Romero, R., & Ovejias, V. J. (2016). Entropy characterisation of overstressed capacitors for lifetime prediction. *Journal of Power Sources*, 336. <https://doi.org/10.1016/j.jpowsour.2016.10.077>
- Cuadras, A., Yao, J., & Quilez, M. (2017). Determination of LEDs degradation with entropy generation rate. *Journal of Applied Physics*, 122(14), 145702. <https://doi.org/10.1063/1.4996629>
- Driel, W., & Fan, X. (2013). Solid State Lighting Reliability. *Springer, New York*. <https://doi.org/10.1007/978-1-4614-3067-4>
- Fan, J., Chen, Y., Jing, Z., Ibrahim, M. S., & Cai, M. (2021). A Gamma process-based degradation testing of silicone encapsulant used in LED packaging. *Polymer Testing*, 96, 107090. <https://doi.org/10.1016/j.polymeresting.2021.107090>.
- Fan, J., Qian, C., Fan, X., Zhang, G., & Pecht, M. (2015a). In-situ monitoring and anomaly detection for LED packages using a Mahalanobis distance approach. *Proceedings of 2015 the 1st International Conference on Reliability Systems Engineering, ICRSE 2015*.
- Fan, J., Yung, K. C., & Pecht, M. (2012). Anomaly detection for chromaticity shift of high power white LED with mahalanobis distance approach. *14th International Conference on Electronic Materials and Packaging, EMAP 2012*. <https://doi.org/10.1109/EMAP.2012.6507916>.
- Fan, J., Yung, K. C., & Pecht, M. (2015). Predicting long-term lumen maintenance life of LED light sources using a particle filter-based prognostic approach. *Expert Systems with Applications*, 42(5), 2411–2420. <https://doi.org/10.1016/j.eswa.2014.10.021>
- Grundmann, M. (2006). *The physics of semiconductors: An introduction including devices and nanophysics*. <https://doi.org/10.1007/3-540-34661-9>
- Hinchi, A. Z., & Tkiouat, M. (2018). Rolling element bearing remaining useful life estimation based on a convolutional long-short-Term memory network. *Procedia Computer Science*.
- Hocheiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/NECO.1997.9.8.1735>
- Hu, Y., Miao, X., Si, Y., Pan, E., & Zio, E. (2022). Prognostics and health management: A review from the perspectives of design, development and decision. *Reliability Engineering & System Safety*, 217, 108063. <https://doi.org/10.1016/j.ress.2021.108063>
- Ibrahim, M. S., Abbas, W., Waseem, M., Lu, C., Lee, H. H., Fan, J., & Loo, K.-H. (2023). Long-Term Lifetime Prediction of Power MOSFET Devices Based on LSTM and GRU Algorithms. *Mathematics*, 11(15), 3283.
- Ibrahim, M. S., Fan, J., Yung, W. K., Priscaru, A., van Driel, W., Fan, X., & Zhang, G. (2020). Machine learning and digital twin driven diagnostics and prognostics of light-emitting diodes. *Laser & Photonics Reviews*, 14(12), 2000254.
- Ibrahim, M. S., Fan, J., Yung, W. K. C., Jing, Z., Fan, X., van Driel, W., & Zhang, G. (2021). System level reliability assessment for high power light-emitting diode lamp based on a Bayesian network method. *Measurement*, 176, Article 109191.
- Ibrahim, M. S., Fan, J., Yung, W. K. C., Wu, Z., & Sun, B. (2019). Lumen Degradation Lifetime Prediction for High-Power White LEDs Based on the Gamma Process Model. *IEEE Photonics Journal*, 11(6), 1–16. <https://doi.org/10.1109/JPHOT.2019.2950472>
- Ibrahim, M. S., Jing, Z., & Fan, J. (2022). Health Monitoring, Machine Learning, and Digital Twin for LED Degradation Analysis. In *Reliability of Organic Compounds in Microelectronics and Optoelectronics: From Physics-of-Failure to Physics-of-Degradation* (pp. 151–210). Springer.
- Ibrahim, M. S., Jing, Z., Yung, W. K. C., & Fan, J. (2021). Bayesian based lifetime prediction for high-power white LEDs. *Expert Systems with Applications*, 185, 115627. <https://doi.org/10.1016/j.eswa.2021.115627>
- IESNA. (2011). Projecting long term lumen maintenance of LED light sources. Illuminating Engineering Society of North America.
- Ji, H. (2021). Statistics Mahalanobis distance for incipient sensor fault detection and diagnosis. *Chemical Engineering Science*, 230. <https://doi.org/10.1016/j.ces.2020.116233>
- Jing, Z., Liu, J., Ibrahim, M. S., Fan, J., Fan, X., & Zhang, G. (2020). Lifetime Prediction of Ultraviolet Light-Emitting Diodes Using a Long Short-Term Memory Recurrent Neural Network. *IEEE Electron Device Letters*, 41(12). <https://doi.org/10.1109/LED.2020.3034567>
- Khelif, R., Chebel-Morello, B., Malinowski, S., Laajili, E., Fnaiech, F., & Zerhouni, N. (2017). Direct Remaining Useful Life Estimation Based on Support Vector Regression. *IEEE Transactions on Industrial Electronics*, 64(3), 2276–2285. <https://doi.org/10.1109/TIE.2016.2623260>
- Li, S., Chen, Z., Liu, Q., Shi, W.-K., & Li, K. (2020). Modeling and Analysis of Performance Degradation Data for Reliability Assessment: A Review. *IEEE Access*, 8, 74648–74678. <https://doi.org/10.1109/ACCESS.2020.2987332>
- Li, Z., Li, J., Wang, Y., & Wang, K. (2019). A deep learning approach for anomaly detection based on SAE and LSTM in mechanical equipment. *The International Journal of Advanced Manufacturing Technology* 2019 103:1, 103(1), 499–510. <https://doi.org/10.1007/S00170-019-03557-W>.
- Liu, H., Yu, D., Niu, P., Zhang, Z., Guo, K., Wang, D., ... Wu, C. (2019). Lifetime prediction of a multi-chip high-power LED light source based on artificial neural networks. *Results in Physics*, 12, 361–367. <https://doi.org/10.1016/j.rinp.2018.11.001>
- Moallemi, A., Burrello, A., Brunelli, D., & Benini, L. (2021). Model-based vs. Data-driven Approaches for Anomaly Detection in Structural Health Monitoring: A Case Study. *Conference Record - IEEE Instrumentation and Measurement Technology Conference, 2021-May*. <https://doi.org/10.1109/I2MTC50364.2021.9459999>
- Park, S., & Ko, J. H. (2021). Robust Inspection of Micro-LED Chip Defects Using Unsupervised Anomaly Detection. *International Conference on ICT Convergence, 2021-October*. <https://doi.org/10.1109/ICTC52510.2021.9620801>.
- Pode, R. (2020). Organic light emitting diode devices: An energy efficient solid state lighting for applications. *Renewable & Sustainable Energy Reviews*, 133, 110043. <https://doi.org/10.1016/j.rser.2020.110043>
- Qian, C., Xu, B., Chang, L., Sun, B., Feng, Q., Yang, D., ... Wang, Z. (2021). Convolutional neural network based capacity estimation using random segments of the charging curves for lithium-ion batteries. *Energy*, 227. <https://doi.org/10.1016/j.energy.2021.120333>
- Schubert, E., & Kim, J. (2005). Solid-State Light Sources Getting Smart. *Science*, 308, 1274–1278. <https://doi.org/10.1126/SCIENCE.1108712>
- Si, X.-S., Wang, W., Hu, C., & Zhou, D. (2011). Remaining useful life estimation - A review on the statistical data driven approaches. *European Journal of Operational Research*, 213, 1–14. <https://doi.org/10.1016/j.ejor.2010.11.018>

- Sun, B., Jiang, X., Yung, K., Fan, J., & Pecht, M. (2017). A Review of Prognostic Techniques for High-Power White LEDs. *IEEE Transactions on Power Electronics*, 32, 6338–6362. <https://doi.org/10.1109/TPEL.2016.2618422>
- Tsui, K. L., Chen, N., Zhou, Q., Hai, Y., & Wang, W. (2015). Prognostics and health management: A review on data driven approaches. *Mathematical Problems in Engineering*, 2015. <https://doi.org/10.1155/2015/793161>
- Wang, Z. (2021). Current status and prospects of reliability systems engineering in China. *Frontiers of Engineering Management*, 8(4), 492–502. <https://doi.org/10.1007/s42524-021-0172-2>
- Wen, M., Jing, Z., Ibrahim, M. S., Fan, J., & Zhang, G. (2021). A hybrid degradation modeling of light-emitting diode using permutation entropy and data-driven methods. 2021 22nd International Conference on Electronic Packaging Technology (ICEPT).
- Wiley, J., & Leong Gan, C. (2020). Prognostics and Health Management of Electronics: Fundamentals, Machine Learning, and the Internet of Things. *Life Cycle Reliability and Safety Engineering 2020* 9:2, 9(2), 225-226. <https://doi.org/10.1007/S41872-020-00119-Y>.
- Wong, K. L., Bosello, M., Tse, R., Falcomer, C., Rossi, C., & Pau, G. (2021). Li-Ion batteries state-of-charge estimation using deep LSTM at various battery specifications and discharge cycles. GoodIT 2021 - Proceedings of the 2021 Conference on Information Technology for Social Good.
- Zhao, Y., Zheng, Z., & Wen, H. (2010). Bayesian statistical inference in machine learning anomaly detection. *Proceedings - 2010 International Conference on Communications and Intelligence Information Security, ICCIIS 2010*, 113-116. <https://doi.org/10.1109/ICCIIS.2010.48>.