Bias Suppression Framework for Detrending Mean of Multi-Output Gaussian Process Regression in LED Remaining Storage Life Prognosis

LIM SZE LI HARRY [1], **PHAM LUU TRUNG DUONG**[1], **HYUNSEOK PARK** [2], **AND NAGARAJAN RAGHAVAN** [1], **(Member, IEEE)**
[1]Engineering Product Development (EPD) Pillar, Singapore University of Technology and Design, Singapore 487372
[2]Department of Information Systems, Hanyang University, Seoul 133-791, Republic of Korea

Corresponding author: Lim Sze Li Harry (harry_lim@mymail.sutd.edu.sg)

ABSTRACT Advancements in storage prognosis tend to be limited by the inherent challenge to collect sufficient significant degradation data over an extensive period. Using only sparse fleet data, multi-output Gaussian process regression (MOGPR) is one of the few techniques that offers a practical data driven approach to model non-monotonic degradation profiles with low mean absolute percentage error (MAPE). This accuracy in the *storage* prognosis context is, however, sensitive to the choice of the *detrending mean*. Working with light emitting diodes (LED) sparse lumen degradation data under storage conditions in this study, the MAPE is observed to be highly correlated to the *detrending bias* – the difference between the detrending mean and the test mean. We explore various approaches to suppress this bias and advocate a generic framework for fleet storage prognosis. The approaches include *detrending* using (A) static training data mean, (B) dynamic observed test data mean, (C) static bounded training data set pairs, (D) dynamic weighted mean of unbounded training data set pairs and (E) moving average of weighted mean of unbounded training data set pairs. Our analysis shows that the moving average approach (Method E) of computing weighted mean of unbounded training data set pairs results in the most stable detrending mean to suppress detrending bias and helps achieve an MAPE lower than 1%.

INDEX TERMS Storage prognosis, light emitting diodes (LED), multi-output Gaussian process regression (MOGPR), detrending mean, detrending bias, moving average (MA).

I. INTRODUCTION

Storage has been a conventional approach to buffer the uncertainty in supply chain, be it to fulfill surge in demand or to address surplus in supply [1], [2]. The equipment or commodity in storage does not perform its designated functional roles, hence, contributes the least value in the system lifecycle. As such, it is also intuitive that the designers prioritize design requirements for operational use over storage. With in-transit inventory optimization, manufacturers can minimize inventory holding and storage duration which further reduces the manufacturers' vested interest in storage.

More disruptive "Black Swan" events have recently been observed, be it a pandemic like COVID-19 or environmental disasters like earthquakes, fires, floods, and super typhoons. These have introduced disruptive demands for emergency response equipment on one hand and led to long-term storage of high value assets like aircrafts in aviation industries during the pandemic on the other hand [3], [4]. Due to the lack of degradation models for such storage condition profiles, the affected industries have to suffer the uncertainty in the reliability of equipment or commodity placed under long-term storage or commit costly resources to inspect and

The associate editor coordinating the review of this manuscript and approving it for publication was Zhaojun Steven Li .

IEEE *Access*

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

service these unused equipment or commodities in order to preserve their reliability till activation [5]–[7]. Sabouri *et al.* [8] propose an algorithm to develop the inspection policy that minimizes the infinite horizon discounted expected penalty, replacement, and inspection costs.

In addition to the lack of financial motivation for the manufacturers to develop these degradation models for storage profile conditions, the manufacturers are in general not positioned within the equipment life cycle to collect appropriate storage data as well [9]. The relevant data can only be collected by the user or by an assigned warehousing operator, who often does not possess the technical knowledge, toolkits and infrastructure to collect data, analyze it appropriately and develop useful degradation models for storage life prognosis [9]. Furthermore, meaningful data has to be collected over an extensive period of time, due to the relatively slower rate of degradation while in storage, leading to both lower velocity and volume for "good data" [10], [11].

Physics-based approaches, data-driven approaches and a hybrid of both are conventional prognostic approaches to develop degradation models. An *et al.* [12] provide a review for selecting data-driven or physics-of-failure prognostic approaches and conclude that Gaussian process regression (GPR) is an easy and fast option for implementation; however, it suffers from cubic computational complexity. Liu *et al.* [13] highlight that this complexity of $O(n^3)$, due to inversion and determinant computation of the $n \times n$ kernel matrix, limits the use of GPR for large datasets. As our storage fleet data is normally sparse, this limitation becomes less of a constraint. Neural networks (NN) provide advantages to handle cases with large noise and complex models, however, they demand extensive and diverse training sets. Elforjani and Shanbr [14] demonstrate the advantage of NN with back propagation (BP), over support vector machine (SVM) regression and GPR for bearing remaining useful life (RUL) prediction with acoustic emission signals. However, given the *storage* context with far lower data velocity and volume, adoption of NN is a challenge for us to estimate the remaining storage life (RSL). Bishnoi *et al.* [15], on the other hand, demonstrate the utility of GPR over NN, with a small dataset, to provide rigorous property estimates, without the issue of overfitting, which is a common problem associated with NN.

For the physics-based approaches, particle filter (PF) and Bayesian method (BM) provide more superior performance to data-driven approaches like GPR and NN; however, they require a prior established underlying physical model governing the degradation process along with the constraint of constant loading conditions [11]. In the context of storage prognosis, such degradation models may not be available as the physics of degradation under storage is hard to decipher and is often, unknown. This poses challenges to adopting PF and BM approaches for RSL prediction.

Focusing on readiness of aerospace products, Lu *et al.* [16] introduce the development of storage life assessment and evaluation with accelerated life test, life evaluation process and storage life assessment method. Mense *et al.* [17] provide a data driven approach with binary logistic regression (BLR) model to assess recent system failures during non-operating storage and transportation. However, the reliability predictions are sensitive to the accuracy and completeness of the information provided to perform the prediction. Li *et al.* [18], on the other hand, proposed a method to estimate the storage life distribution parameters by fitting to exponential component and the storage period can subsequently be determine with test results. However, this approach requires the storage degradation profile to follow the exponential distribution.

There are several good case studies of GPR being used for prognostic applications in the recent past. Yang *et al.* [19] use GPR to present a novel approach to estimate the state-of-health (SoH) for lithium-ion batteries, based on four specific parameters from the charging curve, instead of typical charging cycles. Richardson *et al.* [20] used the GPR to forecast battery state of health to suggest the relative advantage of data-driven approach over physics-of-failure approach when modeling complex battery degradation trends. Extending the typical GPR (single output), they explored the multi-output Gaussian process regression (MOGPR) to exploit the correlation of degradation time-series trend between different battery cells under the same loading environment.

Liu *et al.* [21] presented a comprehensive review on the performance of MOGPR under asymmetric and symmetric scenarios. Under the asymmetric or multi-fidelity scenarios, there are differences in the number of training points across outputs and they attempt to improve the predictions of more costly high-fidelity output by transferring information from the less costly low fidelity outputs. Under the symmetric scenarios, the outputs have the same number of training points, and the authors attempt to improve predictions with all related outputs. The symmetric scenario discussed in their work has some synergy and synchrony with our intent to RSL prediction for fleet storage scenario where the prediction of an equipment degradation trend can be made and improved with degradation trend from other equipment, within the fleet under the same environmental conditions.

Duong *et al.* [22], [23] used MOGPR to predict the lumen maintenance life for light emitting diode (LEDs) over a long horizon of their active service life. While similar in concept, the training data sets in that work are large and the degradation of active components are more prominent than that in the storage context. To this end, we have previously explored the basic MOGPR to obtain residual storage life (RSL) prediction and achieved 80-90% estimation accuracy for Li-ion batteries in storage with sparse degradation data points, in the absence of an established degradation model (Ref. [24]). In order to further improve the accuracy of the predictions, we attempt to refine the basic MOGPR approach.

As part of the MOGPR procedure, it involves a step of *detrending* the degradation time series prior to performing the training. Typically, the arithmetic mean of the training data is used to detrend the training data. However, for the test data, the arithmetic mean of the test data will not be available until all the test data has been observed, and as such, the

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

IEEE *Access*

arithmetic mean of the training data itself is often used, as an estimate, to detrend the test data. It is observed that this choice of the detrending mean may introduce *bias* into the MOGPR training, due to inherent differences between the means of the training and test data sets. These differences may arise from manufacturing differences within specification tolerances or differing usage history prior to storage. As the magnitude of the degradation in storage is typically small, these inherent differences, while seemingly small, may induce a relatively significant bias that can result in unacceptable distortions in the degradation time series predictions, compromising on the RSL estimation accuracy quite a lot.

To minimize this bias, there is a need to explore approaches to detrend the test data, possibly with the mean of the incremental test data observed up to current time of test. A better prediction with lower mean absolute percentage error (MAPE) was obtained by trying this approach in our work [25]. However, this approach was unstable when the number of test data points is low. Subsequently, we also explored possibility of having bounded training sets with an equally weighted mean of both training data sets to detrend the test data. While this further lowered the bias and MAPE, resulting in improved RSL [25], it greatly constrains the selection of the training sets to be the upper and lower bound of all collected data (which most often may not be the case in practical circumstances) and enforces stringency on the test data sets which have to strictly stay within the bound of the training data sets in order to achieve a lower MAPE.

This paper investigates into the effects of *bias* in the detrending mean on the MAPE (or RSL prediction) while performing MOGPR and proposes more versatile approaches to reduce MAPE in the degradation profile. Section II provides an overview of MOGPR and the concept of detrending mean in MOGPR. Sections III describes the data, performance evaluation metric and simulation set-up. Section IV examines the various detrending approaches and discusses the key results and observations. Finally, Section V presents a conclusion of our study with specific recommendations for future exploration.

II. DETRENDING IN MULTI-OUPUT GAUSSIAN PROCESS REGRESSION

A. MULTI-OUTPUT GAUSSIAN PROCESS REGRESSION

MOGPR models the degradation time series, $f(t)$ with a Gaussian probability distribution over the function [26]:

$$f(t) \sim N(m(t), \kappa_{MOGP}) \tag{1}$$

where $m(t)$ is the function mean, estimated by the arithmetic mean of the test time series, and κ_{MOGP} is the function covariance, derived as the product of the label covariance and temporal covariance:

$$\kappa_{MOGP} = \kappa_c(\lambda, \lambda', \theta_c) \times \kappa_t(t, t', \theta_t) \tag{2}$$

where $\kappa_t(t, t', \theta_t)$ is the temporal covariance matrix for test time series, $\kappa_c(\lambda, \lambda', \theta_c)$ is the label covariance matrix,

t denotes a specific time within the time series, t' denotes the remaining times within the time series, λ denotes the label input for a specific times series, λ' denotes the label inputs for remaining times series, and θ_c and θ_t are the vector of hyper-parameters for the label and temporal covariance, respectively. In our context, to model the degradation time series, the intuition for the function mean ($m(t)$) can be perceived as the trend. The function covariance (κ_{MOGP}) can be perceived as the relative changes from the trend. In a basic GPR, also termed as single-output GPR (SOGPR), for time series, we describe the relationship from one data point to the other data points in the time series, with the temporal covariance. For simplicity, we have adopted the common square exponential (SE) covariance for $\kappa_t(t, t', \theta_t)$:

$$\kappa_t = \kappa_{SE} = \sigma_f e^{\left(-\frac{(t-t')^2}{2\theta_t^2}\right)} \tag{3}$$

where σ_f denotes the scaling hyper-parameter for f. The hyperparameters are optimized through maximizing the log marginal likelihood expression:

$$\log p(y \mid t, \theta_t) = -\frac{1}{2} Y^T \kappa_t^{-1} Y - \frac{1}{2} \log |\kappa_t| - \frac{n}{2} \log(2\pi) \tag{4}$$

where $Y = [(y(t_1), \ldots, (y(t_n)]^T$ denotes the array of data points from the degradation time series, n denotes the number of training points (observations).

Extending SOGPR to MOGPR with multiple outputs, we leverage on the relationship of the observed time series test data set with one or more similar time series training data sets to enhance our prediction. This relationship is described by the *label* covariance matrix. The term "label" is merely used to index the training and test data sets. $\kappa_c(\lambda, \lambda', \theta_c)$ is a positive semi-definite covariance matrix, obtained by the free form parameterization by Cholesky decomposition to provide the elements of the lower triangular matrix [6]:

$$\kappa_c = LL^T \begin{bmatrix} \theta_{c,1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ \theta_{c,k-m+1} & \cdots & \theta_{c,k} \end{bmatrix} \tag{5}$$

where $k = m \times (m+1)/2$ is the number of correlation hyper-parameters. The diagonal elements represent the correlation of the times series, whereas the non-diagonal elements represent the correlation between multiple time series.

B. DETRENDING MEAN AND BIAS

Different approaches to preprocess the degradation time series can have a significant impact on the prediction performance. Ahmed *et al.* [27] listed some of the common approaches, including but not limited to deseasonalisation, log transformation and detrending. The basic MOGPR can be specified completely by just the mean function and covariance function. It is common and practical, though not necessary, to use a Gaussian Process (GP) with zero mean function. For simplicity, we start with a fixed (deterministic) value as the GP mean function.

IEEE *Access*

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

For the training data set, this fixed mean function value can be easily obtained by computing the arithmetic mean of the training data set. As illustrated in Figure 1, by taking the difference between each data point and this fixed mean function value, we are effectively *detrending* the training data set into an adjusted training set with zero mean. We can then apply the common zero mean GP to this adjusted training set. However, in the case of the test data set, the test data points are collected incrementally, and hence, the true arithmetic mean can only be computed after the complete test data set has been collected. It will, however, be meaningless by then for practical application of MOGPR prediction. As such, the arithmetic mean of the training data set is typically used as the best estimate to *detrend* the test data before subjecting the adjusted test data to MOGPR based predictions.
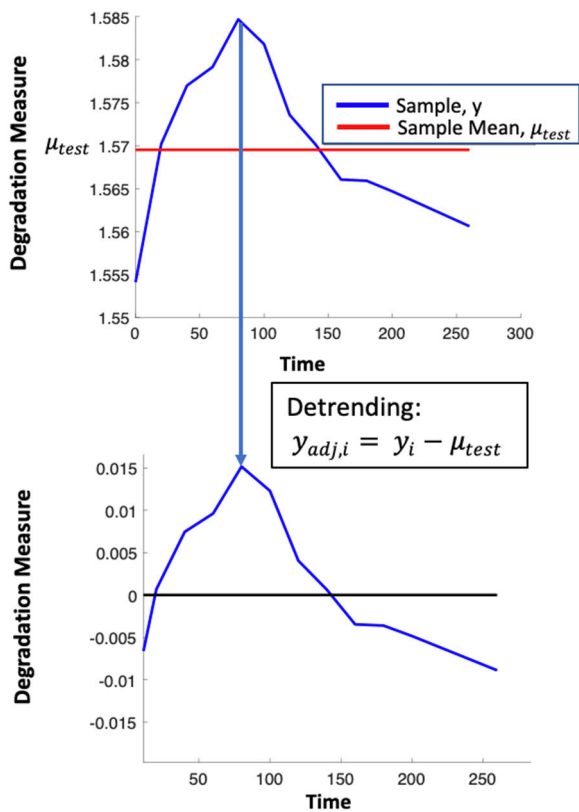
FIGURE 1. Illustration of detrending to shift the sample mean (red line) to zero (black line) by subtracting all the data points with the sample mean.

The inherent differences between the test and training data mean are known to introduce *bias*, as seen in Figure 2. This bias leads to distortions in the MOGPR predictions. Based on our earlier works (Ref. [24] & [25]), we observed a strong negative correlation between the detrending bias and the prediction MAPE. To understand the effects of this bias, various approaches to detrend the test data will be considered in our study here, as summarized in Table 1. In doing so, we hope to develop a generic and practical framework that can be widely used for bias-suppressed prognosis.
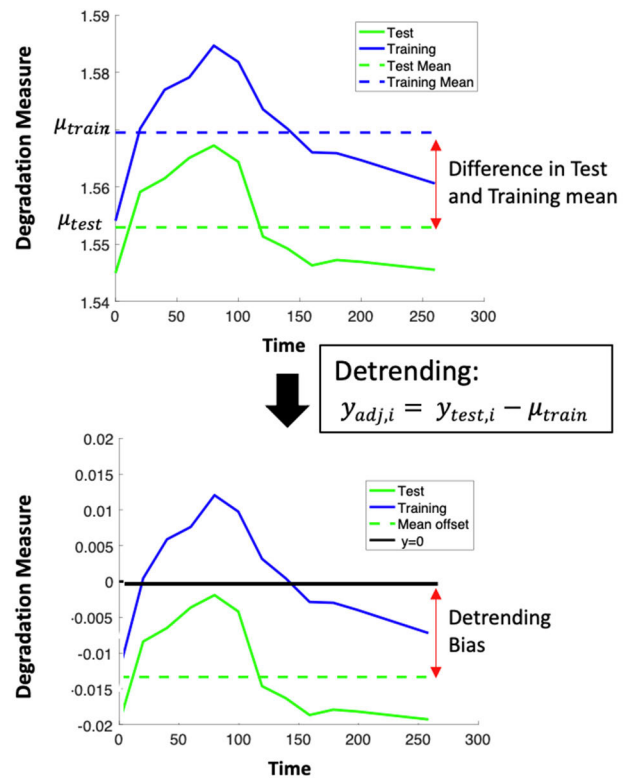

FIGURE 2. Illustration of detrending the test data set with the training mean which introduces a bias that is the difference between the test and training mean.

III. SIMULATION SET-UP

A. CHOICE OF LED TECHNOLOGY FOR THE STUDY
In this study, we have selected LED as the technology of focus to illustrate the utility of MOGPR in storage prognosis and develop our modeling framework, due to the increased proliferation of LED and associated prognostic challenges, as elaborated in the ensuing paragraphs.

Since 2011, LED-based products have steadily penetrated the market compared to conventional incandescent, halogen, florescent and high-intensity discharge (HID) lamps. Market penetration for LED-based products is projected to reach 75-85% by 2030 [28]. With such proliferation, there have been significant research significant on the reliability of LEDs. This is an upheaval undertaking. LEDs typically have long operating lifetime with few observable degradations. It is costly and time consuming for manufacturers to collect failure or degradation data to comprehensively assess and improve the reliability of LED.

B. CURRENT STATUS OF LED PROGNOSIS STUDIES
Furthermore, failures can occur outside laboratory conditions. Trivellin *et al.* [29] provided a comprehensive review of the failures of LEDs in real-world applications, beyond laboratory environment where stress parameters are carefully controlled or isolated. Several real-world LED degradation mechanisms like electrical overstress, assembly issues, mishandling and chemical contamination were reported to be

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

IEEE *Access*

crucial in the upstream manufacturing process to prevent downstream catastrophic service failures. Driel *et al.* [28] reviewed that there have been a whopping 88 failure unique modes for LED documented since 2011 and projected, based on the Goel-Okumoto maturity model, and that another 42 unique failure modes remain to be discovered. The possible interplay of these large number of multiple failure modes presents a big challenge to develop a comprehensive physics-of-failure model for LED.

Nayak *et al.* [30] attempt to predict lumen degradation with traditional statistical models like Normal, Weibull and Lognormal models. They conclude that lognormal provides the best fit with Akaike Information Criterion (AIC). However, such models are deterministic and may not be useful for non-monotonic degradation scenarios, as observed for LEDs under storage.

Recent advances in artificial intelligence and machine learning have provided additional options for reliability analysis. As the LED lamp system comprises of sub-systems like LED driver, LED module, diffuser, reflector and interconnects, the system can be decomposed into sub-systems for analysis. Instead of the traditional deterministic reliability block diagram (RBD) approach for LED system reliability assessment, Ibrahim *et al.* [31] leverage on fault tree analysis (FTA) results and expert knowledge to develop a Bayesian network (BN) to model the complex system reliability to address the uncertainties and interaction between sub-systems. The proposed BN approach provides an accurate lifetime prediction with shorter testing time. While this approach provides better reliability insights for LED manufacturers to assess and improve their product design, it requires the further modelling of sub-systems which is often untenable for the storage agency.

Jing *et al.* [32] implemented long short-term memory (LSTM) and recurrent neural network (RNN) to predict the lifetime of ultra-violet (UV) LEDs against the nonlinear least squares (NLS) regression approach, as recommended by Illuminating Engineering Society of North America (IESNA) TM-21 industry standard. They achieve more robust results with improved accuracy of 29.7% lower lifetime prediction error. Magnien *et al.* [33], on the other hand, used a Bayesian regression approach to quantify the parametric uncertainty in a temperature driven prediction model with the temperature sensitive electrical parameter (TSEP) measurements to improve fatigue lifetime model for LED. Similarly, Ibrahim *et al.* [34] employed the model-based Bayesian method to estimate the remaining useful life with lumen degradation data from accelerated degradation tests. An exponential decay model was developed with parameters obtained through Markov Chain Monte Carlo (MCMC) sampling with Metropolis-Hasting (MH) algorithm. This approach did provided better prediction results against IESNA TM-28. TM-28 is the existing standard employed by the industry to project long term lumen maintenance for LED light source. All of these approaches, however, are still data intensive and focus on the degradation of LED under active usage conditions.

Purwanto *et al.* [35] presented an aging study of LED with counter-intuitive results. LED samples exposed to least level of radiation energy (simulated sunlight) displayed the highest physical degradation in conversion efficiency and transparency of LED, under ambient temperature and short exposure duration of 500 hours. They suggest that there may be other more dominating degradation factors, namely oxygen and humidity exposure, that are responsible for this observation. Degradation may be significant even under storage. This re-iterates the complexity of LED degradation process when different physics-of-failure mechanisms can have counter-intuitive interactions and effects and dominate the degradation under operating and/or storage conditions.

C. LED STORAGE DATA SET

The LED data sets are obtained from the work of Singh and Tan [36] at Chang Gung University (CGU) from the Center for Reliability Science and Technology (CReST®) in their proposed moisture-electrical-temperature (MET) test to evaluate the outdoor reliability of high power LEDs. In order to simulate the morning condition when the streetlamp is "OFF", a subset of 20 blue LED samples is maintained in the "OFF" condition, with a constant temperature of 85°C and 85% relative humidity, as the accelerating environmental factor. The "OFF" condition is akin to LEDs placed under long-term storage, hence, is applicable for our study. In our previous work [25], we have explored some basic flavors of MOGPR to predict remaining storage life for these LEDs and the same data sets have been employed here for this study.
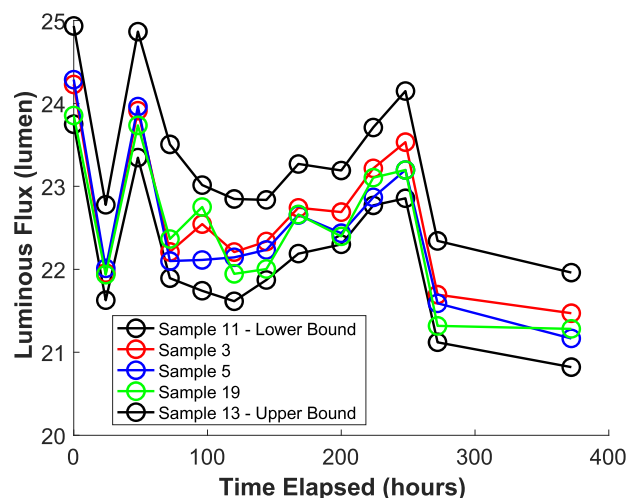

FIGURE 3. Non-monotonic degradation curves due to the partial recovery behavior in light emitting diodes. The samples are intentionally selected to provide a diversity of completely bounded samples and inter-twining samples to simulate unbounded samples.

Each LED sample is a time series that consists of 12 lumen degradation data points, collected every 24 hours, for optical and electrical measurements. 5 LED sample data sets are selected, as illustrated in Figure 3, where 1 LED sample can be considered as the test data set, while the required training sets are selected from the remaining 4 LED samples

IEEE *Access*

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

TABLE 1. Exploration of approaches with different detrending means overview of different detrending means and their characteristics.

Detrending Approaches	*Ideal*	Method A	Method B	Method C	Method D	Method E
Description of Detrending Mean	Mean of test data set	Mean of training data set	Mean of observed test data	Average of bounded training pair means	Weighted mean of training pair means	Moving average of weighted means of training pair means
Nature of Mean	• Static	• Static • Bias for Test	• Dynamic • Fluctuates with test data profile • Tends to true mean as time increases	• Static • Equally weighted • Bounded training pair	• Dynamic • Weighted based on proximity • Any training pair • Fluctuates with test data profile	• Dynamic • Weighted based on moving average of proximity • Any training pair • Less fluctuation with test data profile
Intuition	• True mean • Ideal case	• Simple proxy for test set • Practical	• Use test data to compute mean that incrementally tends to true mean	• Use upper and lower bound data sets as training pair • True test mean will be bounded by upper bound mean and lower bound mean	• Weight based on proximity of test data to training pair • Remove the need to identify upper and lower bound data sets	• Weight based on MA of proximity of test data to training pair • Remove the need to identify upper and lower bound data sets
Training Samples	13	13	11-13	11-13	11-13, 3-13, 5-13	11-13, 3-13, 5-13
Test Samples	3, 5	3, 5	3, 5	3, 5, 19	19	19

to simulated different scenarios. The test data can also be subsequently reused as training sample in other simulations. Samples 11 and 13 serve as the upper and lower bound, respectively. Samples 3 and 5 follow the general degradation profiles of the bounded training samples, and they are distributed around the 50% (mid-point) and 30% (nearer to the lower bound), respectively, of the bounding pair. Sample 19 displays more profile irregularities with reference to the other four samples and is distributed around 30% (nearer to the lower bound), like Sample 5. Sample 19 is also selected as it intertwines with Sample 3 and Sample 5; hence, it can be used to simulate scenarios where the test data transcends the training pairs.

D. TEST REGIME

The Multi-task Gaussian Process (MTGP) Toolbox [37] for MATLAB® is used to implement the MOGPR simulation to generate the projected degradation profile for the test data set. We have performed the MOGPR simulation over the various approaches of detrending on the test data, as listed in Table 1, referred to here on as Methods A to E.

E. ERROR METRICS

To visualize the typical remaining useful life (RUL) or RSL prediction, the predicted degradation profile is plotted together with test data point with a user defined condition failure threshold. The time when the predicted degradation

profile cuts the failure threshold defines the predicted RUL/RSL, as illustrated in Figure 4. This can be validated against the true RUL/RSL determined with actual failure data. The prediction error is defined as:

$$e_j = \tau - \hat{t}_j \qquad (6)$$

where τ denotes the true RSL based on the observed time series data, $\hat{t}_j$ denotes the predicted RSL computed based on the remaining time from the time that j^{th} test data is collected to the time where median generated by the MOGP regression intersects the defined failure threshold, j refers to the number of test data used for MOGP regression.

While this standard threshold failure-based approach works for monotonic degradation profile, it is not useful for non-monotonic degradation profile as there may be multiple points of intersection of the degradation trace with the threshold. For non-monotonic degradation profiles, we have adopted the mean absolute percentage error (MAPE) as our measure of prediction performance:

$$MAPE = \frac{1}{n} \sum_{1}^{n} \left(\left| \frac{\hat{y}_i - y_i}{y_i} \right| \right) \times 100\% \qquad (7)$$

where n is the number of data points used, y_i is the luminous flux observed at data point i, and $\hat{y}_i$ is the predicted luminous flux at data point i, obtained from the median of the MOGPR prediction. MAPE effectively measures the performance of the approach to retrace or replicate the degradation trend. For this paper, only the predicted points, i.e., the remaining

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

IEEE *Access*

points that are yet to be observed at the time of simulation, are subjected to the MAPE computation. This will prevent the dilution of the errors with the increasing number of observed data points.
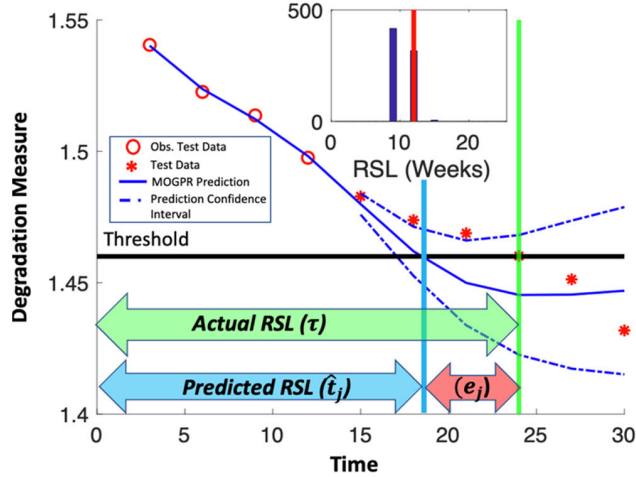

FIGURE 4. Illustration of a typical degradation trace for a monotonic degradation scenario. The predicted RSL is defined by the time (with respect to current time) when the degradation curve intersects the failure threshold. Error (e_j) is computed based on the difference between predicted (t_j) and actual RSL (τ).

IV. METHODS AND RESULTS

A. METHOD A: DETRENDING USING STATIC TRAINING MEAN

The ideal approach for MOGPR is to detrend the test data set with its true mean (as illustrated in Figure 1). However, the test data is observed incrementally, and the true mean can only be computed after all test data are observed. The true mean, hence, cannot be determined or used, in advance, at each iteration step. In lieu of the true test mean, the mean of the training data set is often used as an estimate for the true test mean (as illustrated in Figure 2). This premise is on the assumption that the training set for MOGPR should be largely similar in characteristics to the test set. In this case, both the training and test data sets are detrended by the same training mean, prior to training and prediction, where:

$$\hat{\mu}_{test} = \mu_{train} \tag{8}$$

where $\hat{\mu}_{test}$ = estimator of test mean to be used as detrending mean for test sample, μ_{train} = mean of training sample.

As mentioned in Section II.B, this method introduces a detrending bias when $\mu_{test} \neq \mu_{train}$. The training mean is used to detrend Test Sample 3 and Test Sample 5, with detrending bias of 0.67 and 0.82, to yield an MAPE of 2.9% and 3.8%, respectively. To establish the ideal MAPE baseline reference, the true means of the respectively test data are used to detrend the test data (which is the ideal scenario we refer to in Table 1), prior to subjecting the test data to the MOGPR with single training data set (Sample 13). The detrending bias in this case is zero. For Test Samples 3 and 5, the ideal MAPE

turns out to be 0.5% and 0.4%, respectively. From Figure 5, the MAPE obtained by detrending with training mean shows a significantly higher MAPE than that obtained with true mean.
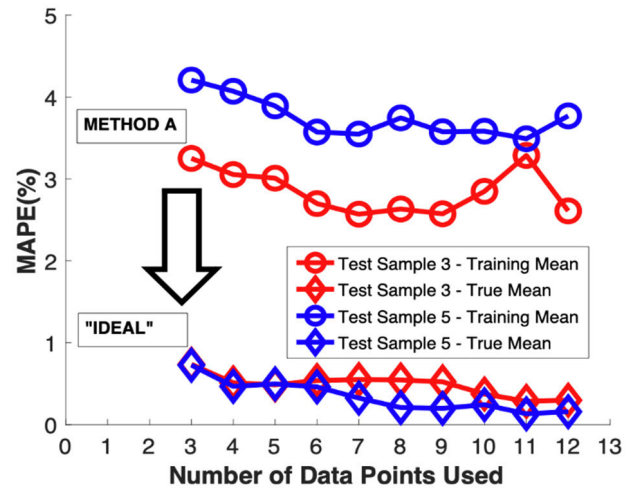

FIGURE 5. Plot of MAPE for test sample 3 & 5 detrended with true test mean and training mean. The plots from training mean reveal significantly higher MAPE than that from the true test mean.
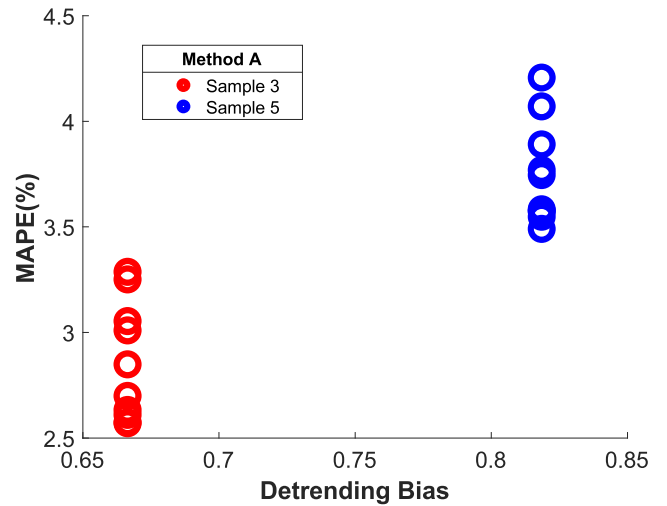

FIGURE 6. Scatter plot to show the relationship between MAPE and detrending bias when sample 3 & sample 5 are detrended with training mean (of sample 13) at multiple time instants. Sample 5 shows a larger detrending bias and MAPE, as compared to sample 3.

Figure 6 shows the plot of the MAPE against the detrending bias. Since the training mean is fixed, the detrending bias is consistent throughout each iteration for Sample 3 and Sample 5, hence, the plot appears as 2 vertical lines of plots. It should be noted that Sample 5 has a larger bias than Sample 3, and therefore, the larger MAPE is also observed for Sample 5.

B. METHOD B: DETRENDING USING DYNAMIC OBSERVED TEST MEAN

Observing the loss of prediction performance due to the detrending bias between the training mean and the true test

Method B: Detrending by dynamic test mean computed with k observed data points
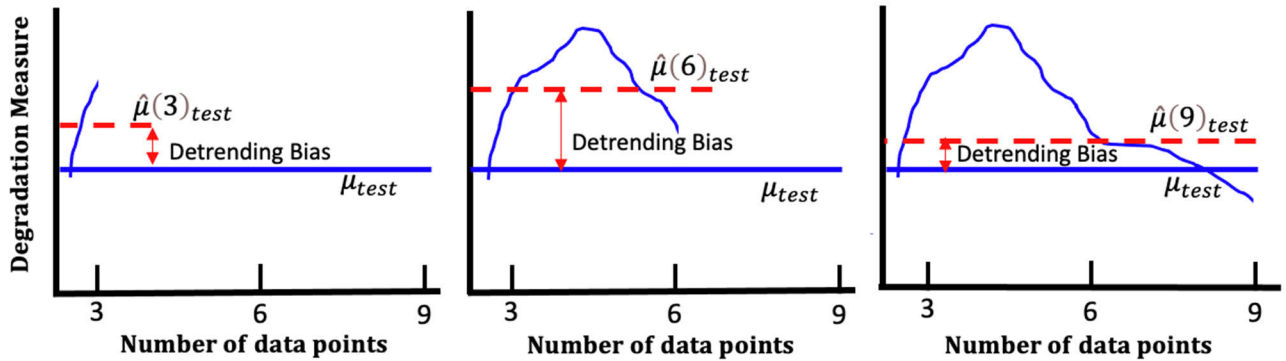
$$y_{adj,k} = y_{test,k} - \hat{\mu}(k)_{test}$$

FIGURE 7. Illustration of method B with dynamic observed test mean, computed with all test data observed thus far. As the number of observed points (k) approaches to the total number of observed test data, this dynamic observed test mean approaches the true test mean.

mean, we previously proposed an incremental observed test mean to be used (Ref. [24]), instead of the training function mean, to detrend the test data, as illustrated in Figure 7. This observed test mean is derived by averaging only the observed test data points, available at the respective measurement time intervals.

$$\hat{\mu}(k)_{test} = \left(\sum_{i=1}^{k} y_i\right)/k \qquad (9)$$

where $\hat{\mu}(k)_{test}$ = estimator of true test mean to be used as detrending mean for test sample, y_i = luminous flux at observation data point i, k = number of observation data points available, thus far in the LED unit under test. Based on Equation (9), it should be noted that the detrending mean is dynamic with every new test data observed.

In our simulation, detrending Samples 3 and 5 with their observed test means as compared to detrending with the training Sample 13, the MAPE is reduced by more than 40%. As the number of sampling points increases, this observed test mean will tend towards the true mean of the test data. It should naturally serve as a better estimate candidate for the true mean of the test data. While this approach offers lower MAPE, hence, better prediction performance, than the training mean approach, the region of instability (orange dashed box), exhibited as high MAPE, is observed when number of sampling points are small, as observed in Figure 8.

This instability reduces rapidly with the increase in number of data points. It usually offers a more stable MAPE only after 5 data points when there are high fluctuations in the initial actual degradation profile. This shortcoming can be explained as when the number of observed test data points are small, there are insufficient data to represent the true test data set. Consequentially, these observed test means computed (when data points are small) cannot achieve a good estimator of the true mean, leading to high detrending bias. This leads to erratic MAPE performance during the initial phase of
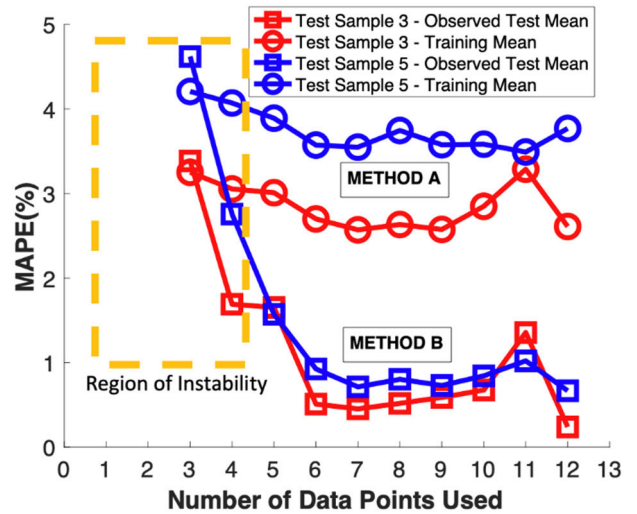

FIGURE 8. Plot of MAPE for test sample 3 & 5 with observed test mean and training mean (of sample 13) as detrending mean. The observed test mean approach reduces the MAPE significantly from the training mean approach with the increase in number of data points used.

prediction when the number of observed test data points is small. Figures 9 and 10 show the MAPE against detrending bias plot. Similarly, it is observed that the larger the detrending bias, the higher the MAPE. Here, a more obvious positive correlation between the MAPE and detrending mean can be clearly observed.

C. METHOD C: DETRENDING USING STATIC BOUNDED TRAINING PAIRS

To alleviate the initial shortcoming presented by detrending with the dynamic observed test mean, our group further proposed to use the equally weighted average of the training means of the upper and lower bound training data series,
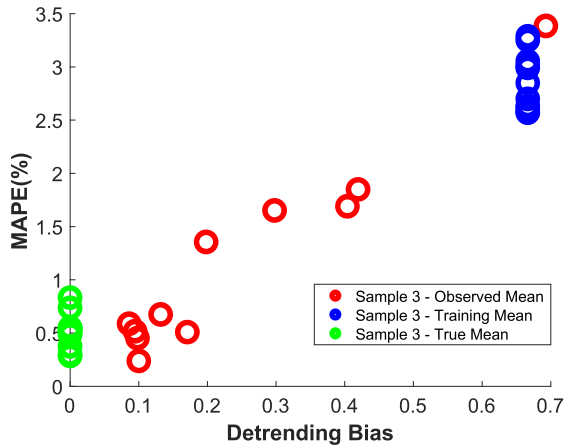
L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

IEEE*Access*

FIGURE 9. Scatter plot MAPE against the bias between true test mean, training mean and observed test mean as the detrending means for each simulation run for test sample 3.
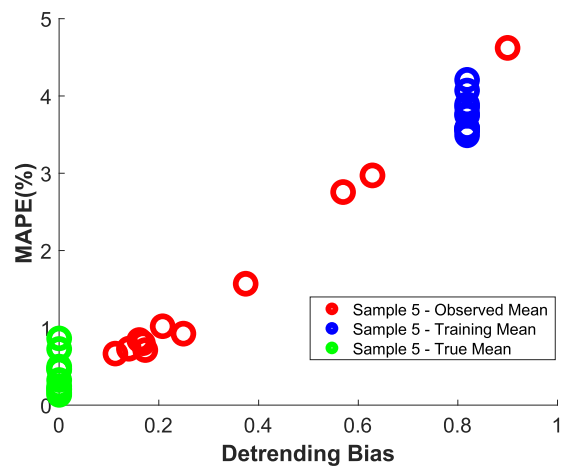

FIGURE 10. Scatter plot MAPE against the bias between true test mean, training mean and observed test mean as the detrending means for each simulation run for test sample 5.

to detrend the test data series, as illustrated in Figure 11 (Ref [25]).

The detrending mean is estimated by the mid-point of the upper and lower bound:

$$\hat{\mu}_{test} = w_{train-UB}\mu_{train-UB} + w_{train-LB}\mu_{train-LB} \quad (10)$$

$$w_{train-UB} = w_{train-LB} = 0.5 \quad (11)$$

where $\hat{\mu}_{test}$ = estimator of test mean to be used as detrending mean for test sample, $w_{train-UB} = w_{train-LB}$ = weight of mean of upper bound and lower bound training samples, respectively, $\mu_{train-UB}$ = mean of upper bound training sample, $\mu_{train-LB}$ = mean of lower bound training sample.

As both training data sets are completely available, they do not suffer from the inadequacy to represent the test data set, as in the previous case of the incremental test mean approach. By estimating the test mean with the equally weighted average of the upper and lower bound training data series, we help reduce the unilateral bias, which exists when
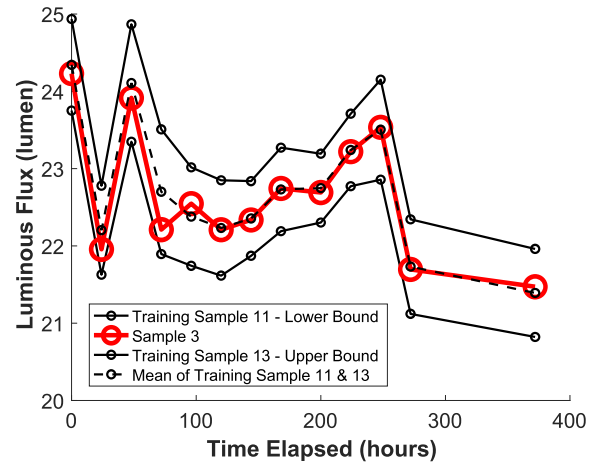

FIGURE 11. Degradation profile of test sample 3 bounded completely by training sample 11 (lower bound) and training sample 13 (upper bound).

performing detrending with either one of the training means alone. By restricting the training sets to be within the upper and lower bound, we restrict the detrending bias to be no greater than half of the difference between the mean of upper and lower bound training sets, as illustrated in Figure 12. This also implies that the smaller the difference between the means of the bounded training pair, the smaller the possible range of detrending bias. In this specific case, Sample 3 is situated close to the middle (50%) of the training pair, as such, it can be observed that their equally weighted mean (22.74 lumen) becomes a very good estimator for the true test mean (22.67 lumen). It provides better stability over the entire range of data points (entire lifecycle span of the prognosis) as compared to Method B, as shown in Figure 13. The detrending bias is greatly suppressed to 0.07 in this specific case, as seen in Figure 14. The MAPE-detrending bias plot reveals that the performance of Method C to be comparable to that derived by the true test mean, with average MAPE of 0.9%.

Similar procedure is repeated for Sample 5 and Sample 19. The degradation profiles of Sample 5 and Sample 19, relative to the bounded training pair (Sample 11 & 13) are shown in Figure 15 and Figure 16, respectively. As elaborated previously, Sample 3 lies around the mid-point between the bounded training pair, leading to low detrending bias, whereas Sample 5 and Sample 19 reside around the 30% (nearer to the lower bound) between the bounded training pair. The 30% offset, while within the constraint of half of the difference between the mean of upper and lower bound training sets, leads to the higher detrending bias (in blue and green) than that for Sample 3 (in red), as seen in Figure 17. As such, the resultant average MAPE from Sample 5 & 19 are also observed to be larger, at 1.4% and 2%, respectively, as observed in Figure 18.

It should be noted that while the equally weighted bounded training pair approach yields a more stable MAPE, it does not guarantee a better performance than the observed test mean, as the detrending bias is constrained to half the difference

IEEE *Access*

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis
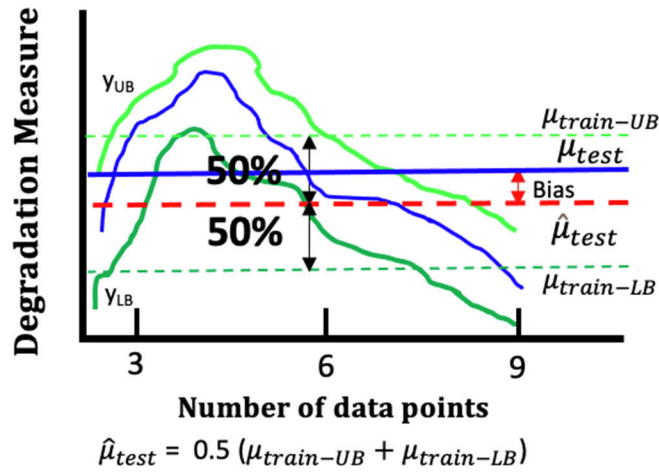

FIGURE 12. Illustration of method C with the plot of means of upper & lower bound training pair, their equally weighted mean, and the true mean of test sample. The bounded training pair approach constrains the detrending bias to a maximum of half the difference between the means of training pair, i.e., the true test mean will fall either between (1) the detrending mean and mean of upper bound training sample, or (2) the detrending mean and mean of lower bound training sample.
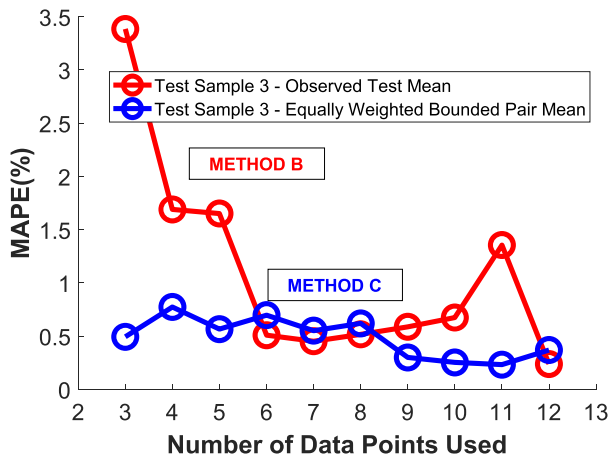
FIGURE 13. Plot of MAPE for test sample 3 with observed test mean (method B) and equally weighted bound pair mean (method C) as detrending mean. The equally weighted bound pair mean approach provides better stability over the entire range of data points (entire lifecycle span of the prognosis).

FIGURE 14. Scatter plot of MAPE against the bias between true test mean, training mean, observed test mean and equally weighted bounded pair mean as the detrending means for each simulation run for test sample 3. The plot includes analysis carried out at multiple instants of time spanning the whole lifecycle.

between the means of the bounded training pair. When the difference between the means of the bounding pair is large, it provides opportunity for larger detrending bias; and hence, higher MAPE. The observed test mean approach, on the other hand, tends towards the MAPE performance of the true mean approach as the number of data points increases. Furthermore, while equally weighted bounded training pair approach eliminates the challenges presented by the earlier approaches, it assumes that the upper and lower training data
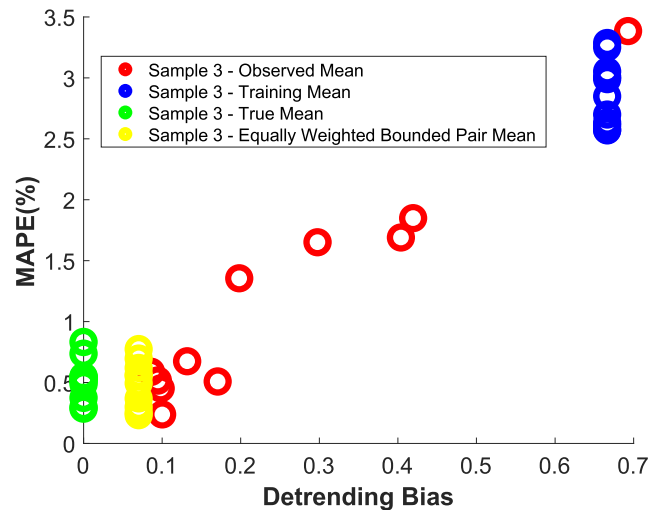
sets can be determined upfront. While in most practical cases, this may be possible, however, it is challenging to obtain the optimal bounding pair for each test sample and we cannot exclude the possibility that the test data series may become a new upper or lower bound, or it may intermittently transcend the upper and lower bound training data, hence, invalidating the basis to lead to the reduction of the bias from the true test mean.
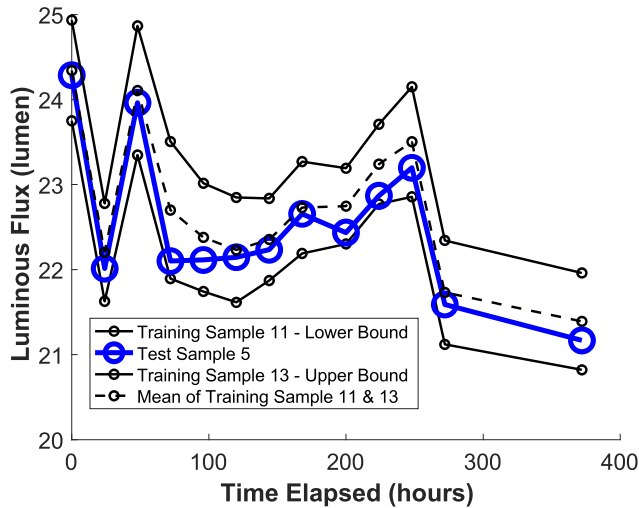
L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

IEEE *Access*

FIGURE 15. Degradation profile of test sample 5 bounded completely by training sample 11 and training sample 13.

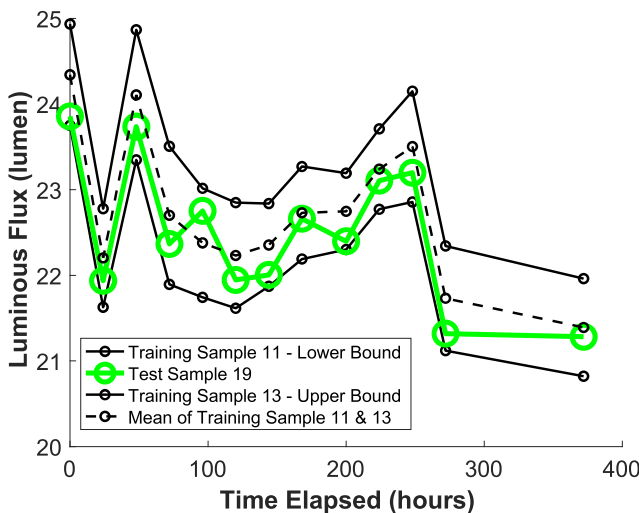**FIGURE 16.** Degradation profile of test sample 19 bounded completely by training sample 11 and training sample 13.

FIGURE 17. Scatter plot for MAPE against the detrending bias for sample 3, sample 5, & sample 19, detrended by the equally weighted training bound pair mean.

FIGURE 18. Comparison of MAPE for test sample 3, test sample 5 and test sample 19, detrended by the equally weighted training bound pair mean.

To analyze the performance of equally weighted means of the bounded pair approach when the test sample transcends the bounded pair, Sample 19 is used as the test sample with Sample 3 and Sample 5 replacing Sample 11, as the lower bound, as seen in Figure 19 and Figure 20, respectively. It should be noted that Sample 19 intersects Sample 3 and Sample 5 as the "lower bound", at several points. Figure 21 shows the relative MAPE performances with the previous static equally weighted bounded training pair approach when the test sample 19 is completely bounded within the upper and lower bound of training samples 11-13 pair, and when the test samples transcend the bounded pairs of training samples 3-13 pair and training samples 5-13 pair. It can be observed that the MAPE has increased, compared to that achieved with training sample 11-13 pair, as test sample 19 is no longer completely bounded by the training pair, in both cases of
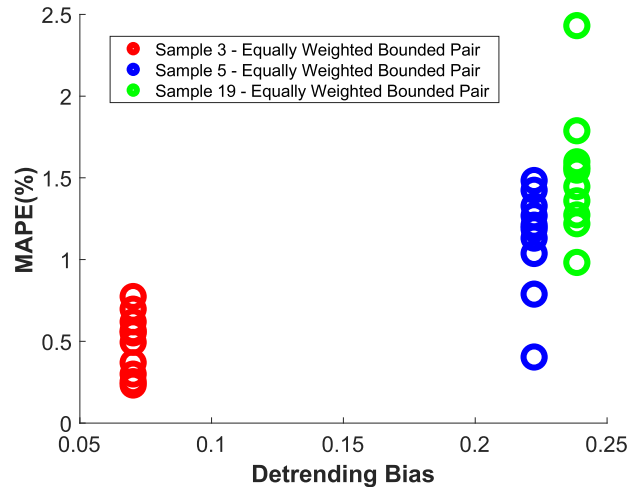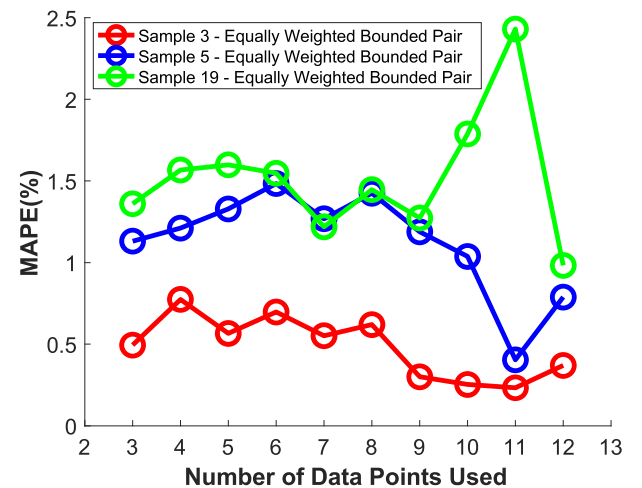
training sample 3-13 pair and training sample 5-13 pair. This violates the "completely bounded test sample" constraint for Method C. Nevertheless, it should be noted that even this "degraded" Method C where the test sample transcends the bounded pair offers better prediction when compared to the most standard training mean approach (Method A).

Between the two runs with the "degraded" Method C where the test sample transcends the bounded pair, training sample 5-13 pair provides better prediction with lower MAPE, compared to training sample 3-13 pair. This is coherent with our postulation of the detrending bias-MAPE correlation, where bias arising from the training sample 5-13 pair (magenta) and training sample 3-13 (cyan) pair are 0.42 and 0.50, respectively, as illustrated in Figure 22. This analysis illustrates that the performance of the equally weighted means of the bounded pair approach degrades when the test data series transcend the bounded training pair.
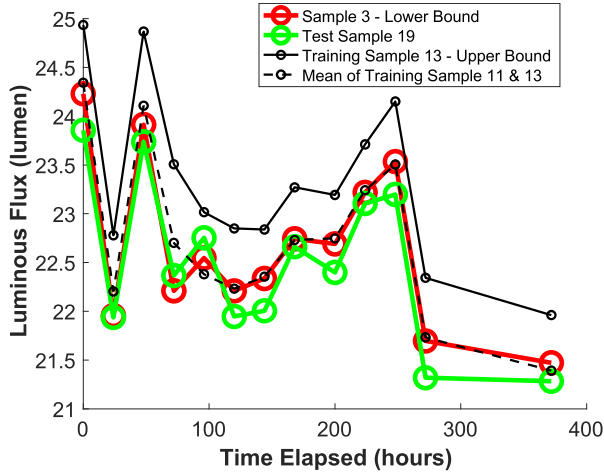
IEEE *Access*

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

FIGURE 19. Degradation profile of test sample 19 transcends the bounds of training sample 3 and training sample 13.

FIGURE 20. Degradation profile of test sample 19 transcends the bounds of training sample 5 and training sample 13.
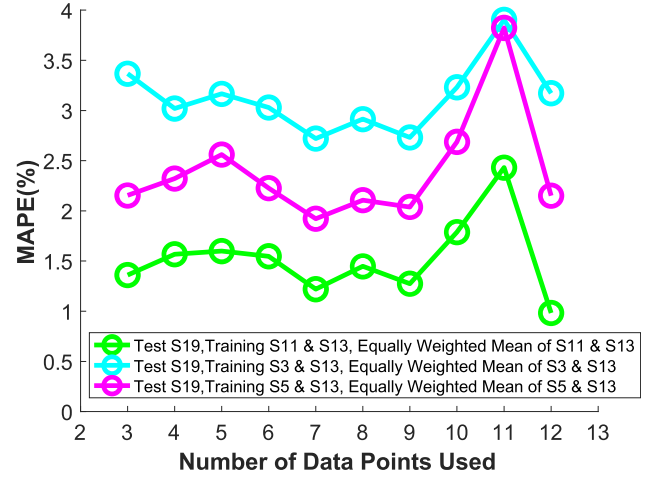

FIGURE 21. Comparison of MAPE for sample 19 detrended with unbounded equally weighted sample 3–13 pair and sample 5–13 pair.
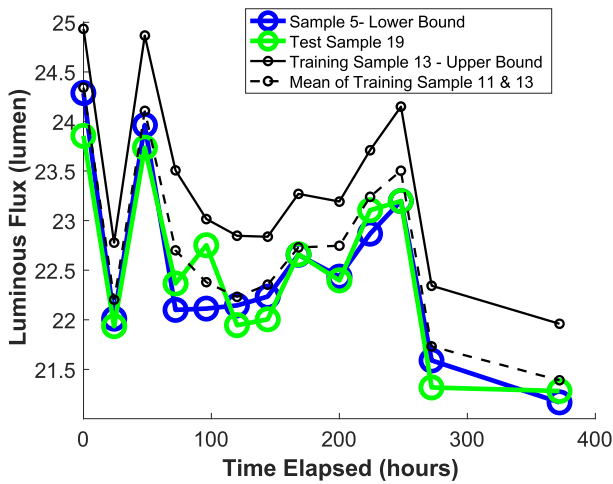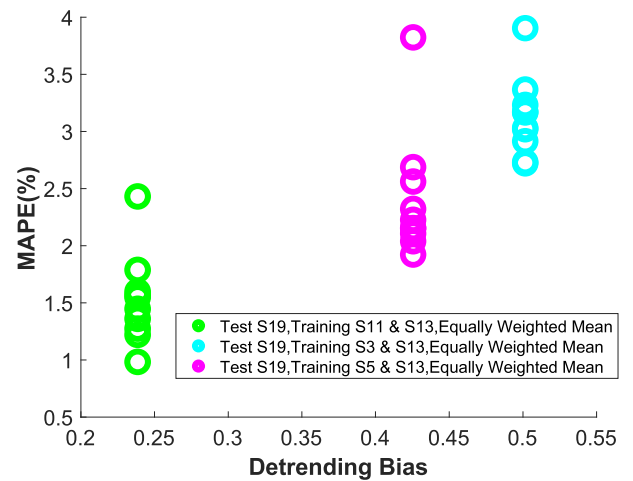

FIGURE 22. Scatter plot of MAPE against the detrending bias for test sample 19, with equally weighted mean with various bounded and unbounded training pairs.

D. METHOD D: DETRENDING WITH DYNAMIC WEIGHTED MEAN OF UNBOUNDED TRAINING PAIR

To relax the constraint to be able to determine upper and lower bound data sets as the training pair, prior to the complete observation of all test data points, a detrending mean obtained by dynamically computing the weighted average of any two training data set pairs is proposed here. In this case, the weight of the mean of each training sets is assigned based on the proximity to the test data. This allows the test data to permanently or intermittently transcend the bounds of either of the training data sets, while maintaining a lower bias from the true test mean.

$$\hat{\mu}(k+1)_{test} = w_{train1}\mu_{train1} + w_{train2}\mu_{train2} \quad (12)$$

$$w_{train2} = (y_{test,k} - y_{train1,k})/(y_{train2,k} - y_{train1,k}) \quad (13)$$

$$w_{train1} = 1 - w_{train2} \quad (14)$$

where $\hat{\mu}(k+1)_{test}$ = estimator of test mean to be used as detrending mean for the test sample for $k+1$ runs,

w_{train1} = weight of mean of training sample 1, μ_{train1} = mean of training sample 1, w_{train2} = weight of mean of training sample 2, μ_{train2} = mean of training sample 2, $y_{test,k}$ = luminous flux of test sample at observation data point k, $y_{train1,k}$ = luminous flux of training sample 1 at observation data point k, $y_{train2,k}$ = luminous flux of training sample 2 at observation data point k.

As the test sample is no longer completely bounded by the training sample, it is intuitive to reduce the bias of the detrending mean by estimating a mean by allocating dynamic weightage to the means of training samples, based on the proximity to observed test sample, as seen in Eqns. (12)-(14). Method D is illustrated in Figure 23. To contrast the effects of Method C and Method D, Method D is repeated for sample 3-13 and 5-13 training pair sets, using test sample 19 for our analysis.

From Figure 24, it can be observed that the dynamic weighted approach provides a lower MAPE in general, than

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

IEEE *Access*

Method D: Detrending by dynamic weighted mean of unbounded training pair

$$y_{adj,k} = y_{test,k} - \hat{\mu}(k)_{test}$$

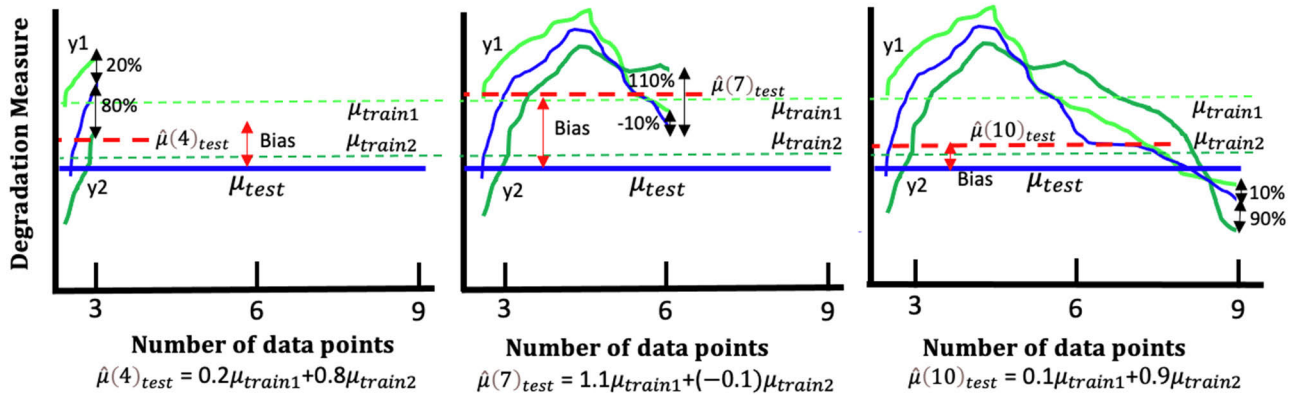$\hat{\mu}(4)_{test} = 0.2\mu_{train1}+0.8\mu_{train2}$

$\hat{\mu}(7)_{test} = 1.1\mu_{train1}+(-0.1)\mu_{train2}$

$\hat{\mu}(10)_{test} = 0.1\mu_{train1}+0.9\mu_{train2}$

FIGURE 23. Illustration of method D with dynamic weighted mean of unbounded training pair at different iteration of observed test data points. The weight of the mean of each training set is assigned based on the proximity to the previous test data.
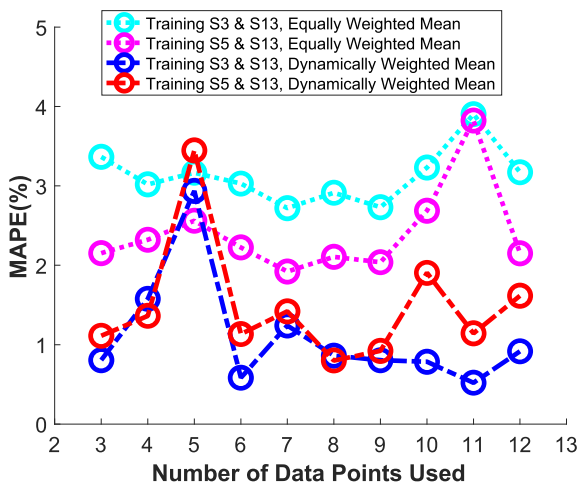

FIGURE 24. Comparison of MAPE for unbounded test sample 19, detrended by the equally weighted and dynamically weighted mean, with training sample 3–13 pair and sample 5–13 pair.

the equally weighted training pair mean approach when the test data transcends the training pair. The MAPE performance by the dynamic weighted mean approach does not have the restriction for the test sample to remain completely bounded by the training sample pair. We attribute the lower MAPE for the prediction to the lower detrending bias by the dynamic weighted mean approach, as seen in Figure 25. Similarly, the detrending biases are distributed laterally due to the dynamic nature of the detrending mean, as compared to the equally weighted mean approach where the static detrending biases are lined vertically. The dynamic weighted mean approach also suffers from the similar challenge of initial instability as the observed test mean approach. This is all the more apparent when there are high fluctuations of the degradation profile in

the initial stage. The dynamic weighted approach will also result in highly fluctuating detrending means, and hence, detrending bias, in tandem with the degradation profile.
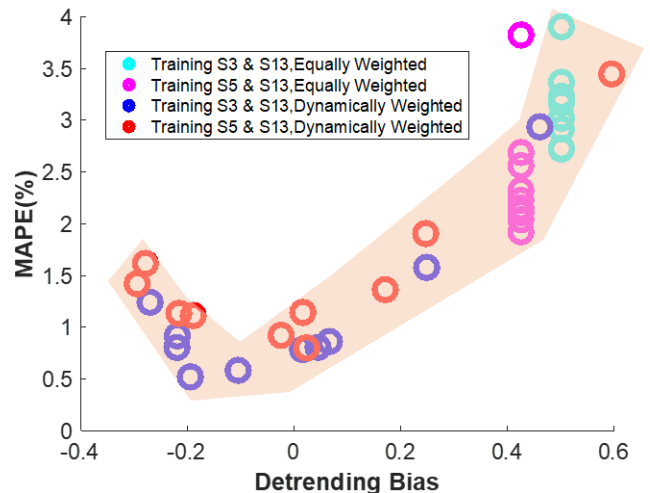

FIGURE 25. Scatter plot MAPE against the detrending bias for unbounded test sample 19, detrended by the equally weighted and dynamic weighted mean, with training sample 3–13 pair and 5–13 pair.

E. METHOD E: DETRENDING WITH DYNAMIC WEIGHTED MEAN OF UNBOUNDED TRAINING PAIRS USING MOVING AVERAGES

While the dynamic weighted training pair mean approach removes the hard constraint for the test data sets to remain within the bounds of the training data sets, the approach produces highly fluctuating detrending means as the proposed detrending mean is a function of the individual observed test data that fluctuates around the static true test mean, at the time of observation. This is especially so when the degradation
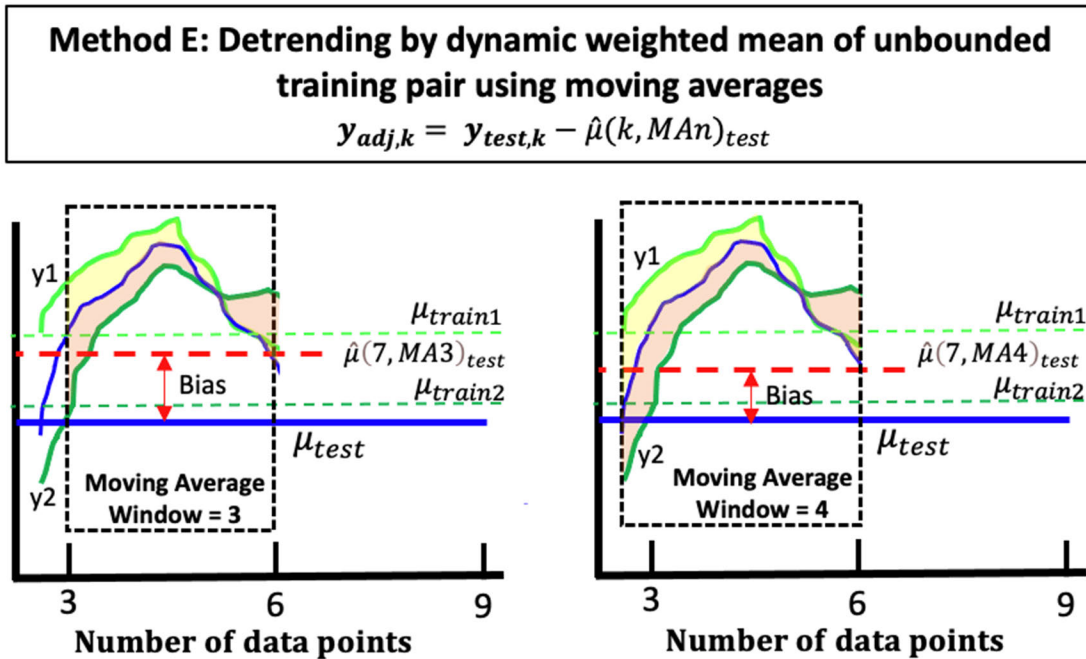
FIGURE 26. Moving average of the dynamic weighted means is introduced to smoothen the fluctuation of detrending mean, hence, the detrending bias. In general, the increase in moving average window reduces the detrending bias.

profile is undulating or non-monotonic. To smoothen this fluctuation, the previous dynamic weighted mean approach is further refined with the moving average (MA) preprocessing methods, with different sliding windows, considered in Ref. [14]. Method E flattens the fluctuations of the detrending means, and generally reduces the bias from the true test mean, as illustrated in Figure 26. The generalized form of the dynamically weighted training pair mean, with moving average of sliding window, *n* is defined as:

$$\hat{\mu}(k+1, MAn)_{test} = w_{train1}\mu_{train1} + w_{train2}\mu_{train2} \quad (15)$$

$$w_{train2} = \frac{1}{n}\sum_{i=k+1-n}^{k}\left[\frac{(y_{test,i} - y_{train1,i})}{(y_{train2,i} - y_{train1,i})}\right] \quad (16)$$

$$w_{train1} = 1 - w_{train2} \quad (17)$$

where $\hat{\mu}(k+1, MAn)_{test}$ = estimator of test mean to be used as detrending mean for test sample for the $(k+1)^{th}$ run, w_{train1} = weight of mean of training sample 1, μ_{train1} = mean of training sample 1, w_{train2} = weight of mean of training sample 2, μ_{train2} = mean of training sample 2, $y_{test,i}$ = luminous flux of test sample at observation data point i, $y_{train1,i}$ = luminous flux of training sample 1 at observation data point i, $y_{train2,i}$ = luminous flux of training sample 2 at observation data point, i.

The dynamic weighted means of any training sample pair with moving averages are extended to the same three test sets, i.e., test sample 19 with training sample 11-13, 3-13 and 5-13 pair sets, with varying moving average windows of window size, $n = 2$ to 5. It should also be noted that MA(1) is effectively the dynamic weighted training pair mean as discussed in the Section IV (Method

D). Figures 27-29 illustrate the relative reduction of the MAPE from the equally weighted bounded pair approach (denoted by "zero" size of moving average), as discussed in Section IV (Method C), to the dynamic weighted mean of training pair by moving averages with varying window sizes.

Figure 27 illustrates the case where the test sample 19 is completely bounded by the training pair. The new approach shows only slight improvement from the equally weighted bounded pair approach ("zero" case) as the previous approach has already yielded relatively good performance, given that it complies with our completely bounded constraint for the approach.

Figure 28 and Figure 29 illustrate the cases where the test sample 19 is not bounded by the training pair. Significant improvement in the MAPE, compared to the equally weighted bounded pair approach, can be observed. By using the dynamic weighted mean of any training sample pair with moving averages, the MAPE can be reduced to about 1%, regardless of whether the test sample is bounded, or not, by the training sample pair across various test samples. This approach provides flexibility to select training samples without compromising the performance of the prediction of the degradation profile.

The corresponding detrending bias are also plotted in Figure 30-32. The boxplots in Figures 30-32 show that this moving average approach suppresses the bias and the fluctuation or dispersion of the bias. This contributes to a lower MAPE during each iteration of the MOGPR. Larger windows of moving averages (*n*) will smoothen the fluctuations encountered in the previous approach, thereby providing a

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis
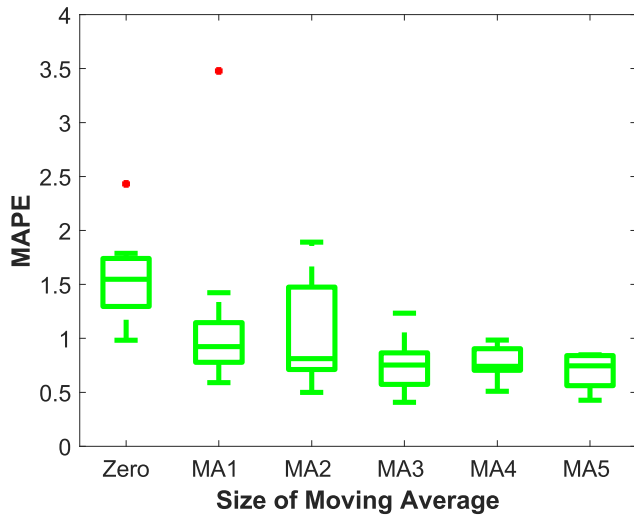
IEEE *Access*

FIGURE 27. Boxplot for MAPE for test sample 19, completely bounded by training samples 11 and 13, across different sizes of moving average to dynamically weight the means of the training samples.
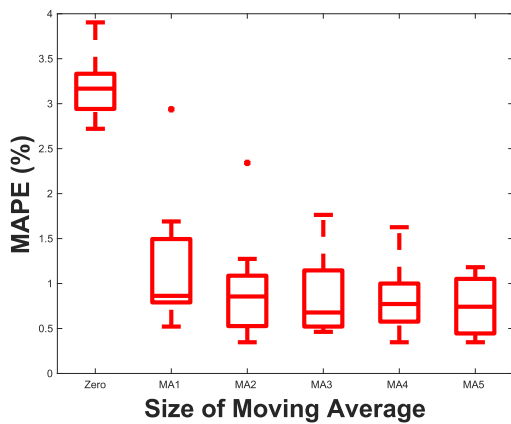

FIGURE 28. Boxplot for MAPE for test sample 19, with training samples 3 and 13, across different sizes of moving average to dynamically weight the means of the training samples.

FIGURE 29. Boxplot for MAPE for test sample 19, with training samples 5 and 13, across different sizes of moving average to dynamically weight the means of the training samples.

FIGURE 30. Boxplot for detrending bias for test sample 19, completely bounded by training samples 11 and 13, across different sizes of moving average to dynamically weight the means of the training samples.

FIGURE 31. Boxplot for detrending bias for test sample 19, with training samples 3 and 13, across different sizes of moving average to dynamically weight the means of the training samples.

more consistent estimation of the detrending mean. This will lead to a lower detrending bias. However, the larger moving average windows will also be a limiting factor in that the very first prediction can only be made after $n + 1$ test data points have been observed. This implies that prediction will be deferred until sufficient data points have been observed and fed as input to derive the moving averages, MA(n). This may be a challenge in the situation when there are sparse or little data points collected, as in our case for RSL prediction.

It will be necessary to probe deeper into the MA window size to trade-off the expected prediction performance and sufficient remaining life to make meaningful predictions for prognosis of equipment in storage. From our experience, MA(3) seems to provide an intuitive window to reap the effects of the moving average approach, while leaving sufficient remaining life to make meaningful predictions. The studied range, from MA(1) to MA(5), provides coverage for
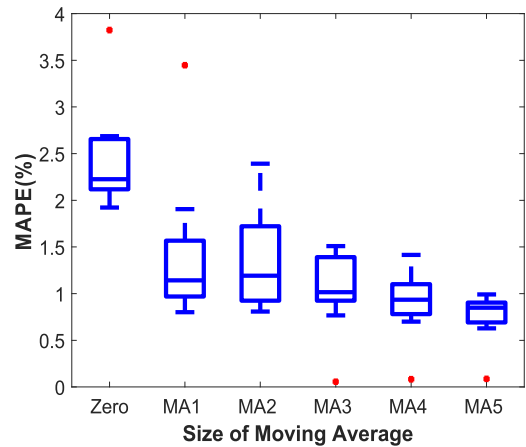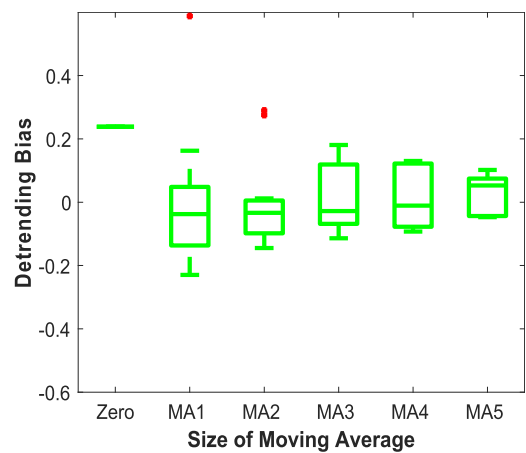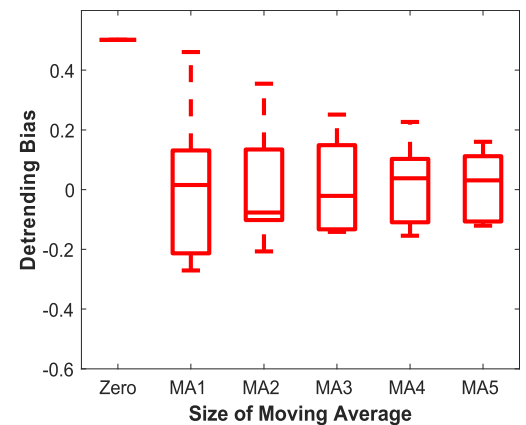
the analysis over our intuitive window of $n = 3$. It is observed that the degree of suppression of bias and improvement to MAPE is not strictly increasing with the increase of moving

IEEE Access

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

TABLE 2. Approaches with different detrending means—summary of performance and characteristics.

| Detrending Approaches	*Ideal*	Method A	Method B	Method C	Method D	Method E
Description of Detrending Mean	Actual Mean of test data set	Mean of training data set	Mean of observed test data	Average of bounded training pair means	Weighted mean of training pair means	Moving average sliding window of weighted means of training pair means
Average Detrending Bias	0 (Static)	0.742 (Static)	0.290 (Dynamic)	0.177 (Static)	0.173 (Dynamic)	0.105 (Dynamic)
Average MAPE Performance	0.45%	3.31%	1.38%	1.45%	1.25%	**0.87%**
Pros	• Theoretical limit	• Simple • Practical	• Tends to ideal approach	• Simple • Stable throughout	• Flexibility to choose upper and lower bound data sets • Remove the need to identify upper and lower bound data set	• Flexibility to choose upper and lower bound data sets
Cons	• Can only be computed at end of collection	• Inherent bias leads to larger MAPE	• Initial Instability	• Need to identify upper and lower bound data set • Need to ensure test data is completely bounded	• Initial Instability	• Deferment of prediction until *n+1* test data have been observed. • Choice of MA windows
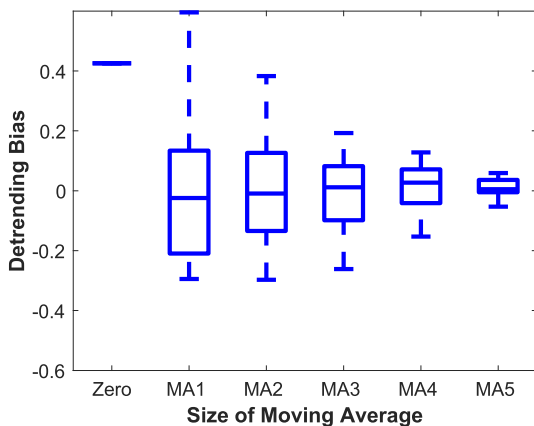
FIGURE 32. Boxplot for detrending bias for test sample 19, with training samples 5 and 13, across different sizes of moving average to dynamically weight the means of the training samples.

average windows, as seen in Figure 30 and Figure 31, where the detrending bias from MA(2) is observed to be lower than that from MA(3). As such, the choice of the ideal moving average window (which very much depends on the time series pattern and evolution of the data) remains inconclusive and will require further investigation.

The effects of bias suppression and improvement to MAPE could also be visualized with a scatterplot. Figure 33 and Figure 34 show the scatterplots of MAPE against detrending bias for test sample 19 with training samples 3-13 pair and 5-13 pair, respectively. A "*V*" shape clusters of data points in the scatterplot suggests the strong correlation between the detrending bias and MAPE.

F. SUMMARY OF METHODS

The approaches we present in the sequence of Methods A to E above gradually suppress the detrending bias and achieve a better MAPE performance. Method A uses the training mean as a simple and practical estimate for the detrending mean. However, due to the inherent bias between the training and test data, it introduces detrending bias which results in significant MAPE compared to the ideal case of detrending with true test mean. Method B uses the observed test data to incrementally compute the detrending mean, which tends to the true test mean in the long run. While it eventually tends to the performance of the ideal case, it may suffer from initial instability. Method C uses an equally weighted mean of a pair of bounding training samples to overcome the initial instability of Method B, however, it imposes the restriction for the test sample to be completely bounded
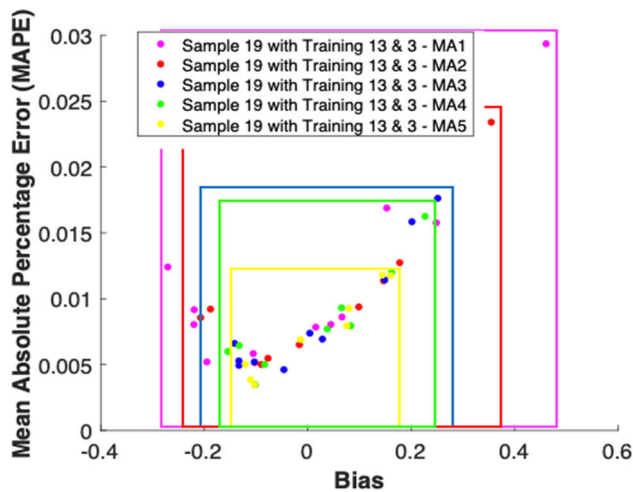
L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

IEEE *Access*

FIGURE 33. Scatterplot illustrates the effects of detrending bias on MAPE for test sample 19, with training samples 3 and 13, over different windows of moving average with dynamic weighted means. The colored average window increases from MA(1) (magenta) to MA(5) (yellow).
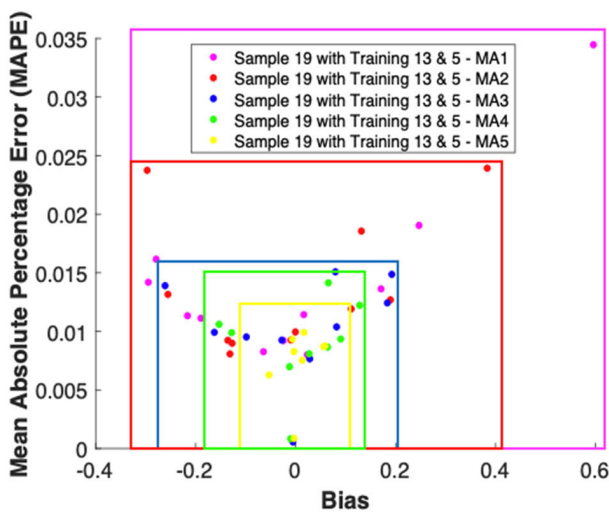

FIGURE 34. Scatterplot illustrates the effects of detrending bias on MAPE for test sample 19, with training samples 5 and 13, over different windows of moving average with dynamic weighted means. The colored boxes illustrate the reduction of the spread of MAPE and bias as moving average window increases from MA(1) (magenta) to MA(5) (yellow).

by the training samples, to yield good detrending bias and MAPE. To relax the "completely bounded" constraint and to cater for flexibility in training samples selection, while maintaining comparable MAPE performance to the equally weighted mean of training pair approach, a dynamic weighted mean of unbounded training pair is proposed. This approach, however, suffers from similar initial instabilities as in the observed test mean approach. Eventually, we propose a moving average pre-processing of the detrending means from the weighted mean of training pair, to smoothen the initial instability, especially when the test data is non-monotonic and fluctuates across the degradation profile. The summary of

the characteristics and performance of the various detrending approaches, with their respective pros and cons, is presented in Table 2. Given the sparse data used for the simulations, the computation load is not demanding. The computation time ranges from around 10 sec for the single training sample approaches, to approximately 20 sec for the training sample pair approaches. The regression is performed on a MacBook Air (Retina, 2018) with 1.6 GHz Dual-Core Intel i5 and 16GB 2133 MHz LPDDR3 RAM.

V. CONCLUSION AND FUTURE WORK

In the absence of extensive datasets and established physics-of-failure models, MOGPR is one of the few data-driven techniques that offers a practical approach to model non-monotonic degradation profiles for prognosis in fleet storage scenario. We observe that the introduction of detrending bias reduces the prediction MAPE performance. As such, we were motivated to develop a framework to suppress this detrending bias further to achieve lower prediction MAPE. The bounded training pair approach provides a good MAPE but it requires the test sample to remain completed bounded by the training sample pair. In practical applications where the test sample data points are incrementally collected, it will not be possible to ensure, in advance, that the entire test sample time series remains completely bounded by the selected training sample pair. In this study, we extend the approach with dynamically evolving weights derived from the proximity of the test data to the training data pair, to address scenarios where the test sample transcends the training sample pair. This relaxes the bounding constraint while achieving similar MAPE performance, thus, yielding improved flexibility for the training samples selection. The technique, however, suffers from initial instability. As such, we further refined the approach with the moving average sliding windows to smoothen the fluctuation of the dynamic detrending mean and bias. This eventually resulted in a drastic reduction of MAPE to about 1%, comparable to that obtained with the ideal true test mean, regardless of whether the test sample is bounded, or not, by the training sample pair. However, the choice of the ideal moving average window remains inconclusive and will require further investigation.

Note that past research works that dealt with GPR mostly focused on active use loading condition data sets, wherein the extent of degradation is far higher than the detrending bias. In such scenarios, the use of the standard training mean approach (Method A) already yields very good outcomes. However, in the storage prognosis scenario, as the extent of degradation becomes comparable to the detrending bias, we are required to more carefully analyze the data using the Methods B-E discussed in this work to achieve similar accuracy standards of MAPE. As such, it is worth mentioning that the approach presented in this work is mostly suited only for storage prognosis scenarios and such intricacies are not relevant to prognosis under active use conditions.

In summary, this paper has laid the foundation to shift the fleet storage industry from traditional periodic inspection

approach to condition-based maintenance by offering simple and practical MOGPR modeling framework. It dispels the misconception on the need for Big Data to leverage artificial intelligence and machine learning for fleet storage prognosis with the use of only sparse degradation data patterns. Based on the correlation observed between the detrending bias and MAPE, we have developed the framework to suppress detrending bias and achieve much better degradation prediction with an improvement in MAPE accuracy for MOGPR by 70% (from 3.3% to 0.87%) through detrending with dynamic weighted mean of unbounded training data sets using the moving average sliding window approach.

Future works may include investigations into the behaviors of the various windows of moving averages on MAPE for different degradation profiles, vis-à-vis different sampling rate. The optimal moving average window may depend on the extent of non-monotonicity in the degradation profile and the full width half maximum (FWHM) of any humps in the time series trend, as well as the density of data points. Alternatively, an operational perspective can be studied wherein the adoption of an incremental window size n that increases with the number of data points is feasible, given that the expectation on accuracy in the initial storage life is not as critical, and a progressive reduction in the MAPE in tandem with the increase in storage life may be acceptable. Finally, beyond detrending bias, there may be other factors, for instance, the (1) number of observations, (2) degree of correlation between the multiple outputs, (3) degree of noise, (4) number of iterative steps for hyper-parameter optimization, introduced during MOGPR that may affect MAPE. A systematic analysis to understand these potential error sources may support further suppression of the MAPE.

REFERENCES

[1] J. R. Bradley, "An evaluation of capacity and inventory buffers as mitigation for catastrophic supply chain disruptions," in *Global Supply Chain Security*, A. Thomas and S. Vaduva, Eds. New York, NY, USA: Springer, 2015, doi: 10.1007/978-1-4939-2178-2_7.

[2] J. R. Bradley, "An improved method for managing catastrophic supply chain disruptions," *Bus. Horizons*, vol. 57, no. 4, pp. 483–495, Jul. 2014, doi: 10.1016/j.bushor.2014.03.003.

[3] *Air Passenger Market Analysis*. Accessed: Oct. 4, 2021. [Online]. Available: https://www.iata.org/en/iata-repository/publications/economic-reports/air-passenger-monthly-analysis—august-2021/

[4] W. Wang and J. Xiong, "Cold-standby redundancy allocation problem with degrading components," *Int. J. Gen. Syst.*, vol. 44, pp. 876–888, Nov. 2015, doi: 10.1080/03081079.2015.1028541.

[5] E. C. Martinez, "Storage reliability with periodic test," in *Proc. Annu. Rel. Maintainability Symp.*, Jan. 1984, pp. 181–184.

[6] T. Kitagawa, T. Yuge, and S. Yanagi, "Periodic and non-periodic inspection policies for a one-shot system with minimal repair," *J. Jpn. Ind. Manage. Assoc.*, vol. 66, no. 4, pp. 387–395, 2016, doi: 10.11221/jima.66.387.

[7] Z. Liu and X. Liu, "Storage reliability assessment for the stored equipment under periodical inspection," *Adv. Mech. Eng.*, vol. 10, no. 6, Jun. 2018, Art. no. 168781401878231, doi: 10.1177/1687814018782312.

[8] A. Sabouri, S. M. Shechter, and W. T. Huh, "Inspecting a vital component needed upon emergency," *Prod. Oper. Manage.*, vol. 24, no. 12, pp. 1839–1851, Dec. 2015, doi: 10.1111/poms.12386.

[9] G. Niu, *Data-Driven Technology for Engineering Systems Health Management: Design Approach, Feature Construction, Fault Diagnosis, Prognosis, Fusion and Decisions*. Singapore: Springer, 2016, pp. 43–45. [Online]. Available: https://link.springer.com/book/10.1007/978-981-10-2032-2#about

[10] L. J. Wise, R. D. Schrimpf, H. G. Parks, and K. F. Galloway, "A generalized model for the lifetime of microelectronic components, applied to storage conditions," *Microelectron. Rel.*, vol. 41, no. 2, 317-322, 2001, doi: 10.1016/S0026-2714(00)00220-1.

[11] B. Zheng, F. Cai, and J. Xu, "Evaluation of system storage reliability," *Kybernetes*, vol. 33, no. 2, pp. 438–445, Feb. 2004, doi: 10.1108/03684920410514463.

[12] D. An, N. H. Kim, and J. H. Choi, "Practical options for selecting data-driven or physics-based prognostics algorithms with reviews," *Rel. Eng., Syst. Saf.*, vol. 133, pp. 223–236, Jan. 2015.

[13] H. Liu, Y.-S. Ong, X. Shen, and J. Cai, "When Gaussian process meets big data: A review of scalable GPs," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 11, pp. 4405–4423, Nov. 2020, doi: 10.1109/TNNLS.2019.2957109.

[14] M. Elforjani and S. Shanbr, "Prognosis of bearing acoustic emission signals using supervised machine learning," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5864–5871, Jul. 2018.

[15] S. Bishnoi, S. Singh, R. Ravinder, M. Bauchy, N. N. Gosvami, H. Kodamana, and N. M. A. Krishnan, "Predicting Young's modulus of oxide glasses with sparse datasets using machine learning," *J. Non-Crystalline Solids*, vol. 524, Nov. 2019, Art. no. 119643.

[16] Z. Lv, Z.-M. Zhao, and W.-B. Shi, "Research on storage life assessment method of aerospace products," in *Proc. Int. Conf. Qual., Rel., Risk, Maintenance, Saf. Eng. (QR2MSE)*, Aug. 2019, pp. 448–454, doi: 10.1109/QR2MSE46217.2019.9021153.

[17] L. J. Gullo, A. T. Mense, J. L. Thomas, and P. E. Shedlock, "Models and methods for determining storage reliability," in *Proc. Annu. Rel. Maintainability Symp. (RAMS)*, Jan. 2013, pp. 1–6, doi: 10.1109/RAMS.2013.6517704.

[18] J. Li, X. Zhu, and H. Li, "A method for exponential component storage effect evaluation based on Bayesian inference," in *Proc. 3rd Int. Conf. Mech., Control Comput. Eng. (ICMCCE)*, Sep. 2018, pp. 553–558, doi: 10.1109/ICMCCE.2018.00122.

[19] D. Yang, X. Zhang, R. Pan, Y. Wang, and Z. Chen, "A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve," *J. Power Sources*, vol. 384, pp. 387–395, Apr. 2018.

[20] R. R. Richardson, M. A. Osborne, and D. A. Howey, "Gaussian process regression for forecasting battery state of health," *J. Power Sources*, vol. 357, pp. 209–219, Jul. 2017.

[21] H. Liu, J. Cai, and Y.-S. Ong, "Remarks on multi-output Gaussian process regression," *Knowl.-Based Syst.*, vol. 144, pp. 102–121, Mar. 2018, doi: 10.1016/j.knosys.2017.12.034.

[22] P. L. T. Duong and H. Park, "Application of multi-output Gaussian process regression for remaining useful life prediction of light emitting diodes," *Microelectron. Rel.*, vol. 88, pp. 80–84, Sep. 2018.

[23] P. L. T. Duong and N. Raghavan, "Prognostic health management for LED with missing data: Multi-task Gaussian process regression approach," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Chongqing)*, Oct. 2018, pp. 1182–1187.

[24] L. Sze Li Harry, P. L. T. Duong, and N. Raghavan, "Exploration of multi-output Gaussian process regression for residual storage life prediction in lithium ion battery," in *Proc. Prognostics Health Manage. Conf.*, May 2020, pp. 263–269.

[25] S. L. H. Lim, P. L. T. Duong, H. Park, P. Singh, C. M. Tan, and N. Raghavan, "Assessing multi-output Gaussian process regression for modeling of non-monotonic degradation trends of light emitting diodes in storage," *Microelectron. Rel.*, vol. 114, Nov. 2020, Art. no. 113794, doi: 10.1016/j.microrel.2020.113794.

[26] C. E. Ramsmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.

L. S. Li Harry *et al.*: Bias Suppression Framework for Detrending Mean of MOGPR in LED Remaining Storage Life Prognosis

IEEE *Access*

[27] N. K. Ahmed, A. F. Atiya, and N. E. Gayar, "An empirical comparison of machine learning models for time series forecasting," *Econ. Rev.*, vol. 29, nos. 5–6, pp. 594–621, 2010, doi: 10.1080/07474938.2010.481556.

[28] W. D. V. Driel, B. Jacobs, P. Watte, and X. Zhao, "Reliability of LED-based systems," in *Proc. 22nd Int. Conf. Thermal, Mech. Multi-Physics Simulation Exp. Microelectron. Microsyst. (EuroSimE)*, 2021, pp. 1–4, doi: 10.1109/EuroSimE52062.2021.9410861.

[29] N. Trivellin, M. Meneghini, M. Buffolo, G. Meneghesso, and E. Zanoni, "Failures of LEDs in real-world applications: A review," *IEEE Trans. Device Mater. Rel.*, vol. 18, no. 3, pp. 391–396, Sep. 2018, doi: 10.1109/TDMR.2018.2852000.

[30] P. Nayak, A. N. Padmasali, and S. G. Kini, "Life estimation of high power LED using distribution based reliability analysis," in *Proc. 2nd IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT)*, Jun. 2017, pp. 1626–1631, doi: 10.1109/RTEICT.2017.8256874.

[31] M. S. Ibrahim, J. Fan, W. K. C. Yung, Z. Jing, X. Fan, W. van Driel, and G. Zhang, "System level reliability assessment for high power light-emitting diode lamp based on a Bayesian network method," *Measurement*, vol. 176, May 2021, Art. no. 109191, doi: 10.1016/j.measurement.2021.109191.

[32] Z. Jing, J. Liu, M. S. Ibrahim, J. Fan, X. Fan, and G. Zhang, "Lifetime prediction of ultraviolet light-emitting diodes using a long short-term memory recurrent neural network," *IEEE Electron Device Lett.*, vol. 41, no. 12, pp. 1817–1820, Dec. 2020, doi: 10.1109/LED.2020.3034567.

[33] J. Magnien, M. Dvorzak, U. Kleb, M. Mácke, and E. Kraker, "Probabilistic approach for temperature driven fatigue lifetime data analysis to improve prognostics and health management of LED packages," in *Proc. 26th Int. Workshop Thermal Investigations ICs Syst. (THERMINIC)*, 2020, pp. 173–179, doi: 10.1109/THERMINIC49743.2020.9420536.

[34] M. S. Ibrahim, Z. Jing, W. K. C. Yung, and J. Fan, "Bayesian based lifetime prediction for high-power white LEDs," *Expert Syst. Appl.*, vol. 185, Dec. 2021, Art. no. 115627, doi: 10.1016/j.eswa.2021.115627.

[35] E. Purwanto, P. Dupuis, L. Canale, N. I. Sinisuka, and G. Zissis, "Aging study of remote luminophore at ambient temperature," in *Proc. IEEE Int. Conf. Environ. Electr. Eng. Ind. Commercial Power Syst. Eur.*, Jun. 2019, pp. 1–5, doi: 10.1109/EEEIC.2019.8783248.

[36] P. Singh and T. C. Tan, "Degradation physics of high power LEDs in outdoor environment and role of phosphor in degradation process," *Sci. Rep.*, vol. 6, Apr. 2016, Art. no. 24052.

[37] R. Durichen, M. A. F. Pimentel, L. Clifton, A. Schweikard, and D. A. Clifton, *MTGP—A Multi-Task Gaussian Process Toolbox, Computational Health Informatics Laboratory*. Oxford, U.K.: Univ. Oxford, 2015. Accessed: Jul. 7, 2021. [Online]. Available: https://www.rob.uni-luebeck.de/index.php?id=410&L=0%0A

PHAM LUU TRUNG DUONG received the B.S. and M.S. degrees in electrical engineering from Bauman Moscow State Technical University, in 2006 and 2008, respectively, and the Ph.D. degree in process control from Yeungnam University, South Korea, in 2013. He is currently a Senior Research Scientist at the Engineering Product Development (EPD) Pillar, Singapore University of Technology and Design (SUTD). His research interests include PID control, model predictive control, stochastic systems, prognostics and health management, optimization, reduced order modeling, and uncertainty quantification.

HYUNSEOK PARK received the Ph.D. degree from the Pohang University of Science and Technology, in 2014. After his Ph.D. degree, he spent a couple of years at the Institute of Data, Systems and Society (IDSS), Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. He is an Assistant Professor at the Department of Information Systems, College of Engineering, Hanyang University, Seoul, South Korea. He is the Director of the Future Intelligence Laboratory and works on patent analytics, technology and innovation management, and applications of machine learning and analytics to multi-dimensional industrial problems.

LIM SZE LI HARRY received the B.Eng. degree in mechanical engineering, the M.Sc. degree in industrial and systems engineering, the M.Sc. degree in defense technology and systems, and the M.Tech. degree in knowledge engineering from the National University of Singapore (NUS), in 1998, 2002, 2007, and 2008, respectively. He had also received the M.S. degree in computer science from the U.S.-Naval Postgraduate School (NPS), in 2006, and the M.B.A. degree from Nanyang Technological University (NTU), in 2013. He is currently pursuing the Eng.D. degree with the Engineering Product Development (EPD) Pillar, Singapore University of Technology and Design (SUTD). He is a Principal Engineer at Singapore Armed Forces. His current research interests include modeling and simulation, artificial intelligence and machine learning, prognostics and health management, reliability, and maintenance sciences. He is a Senior Member (SMIES) and a Chartered Engineer (C.Eng.) at the Institute of Engineers, Singapore (IES), and a member and a Certified Reliability Engineer (CRE) at American Society for Quality and Singapore Quality Institute. He holds a Certificate in AI Ethics and Governance (CAEG) (Professional Level) with the Singapore Computer Society.

NAGARAJAN RAGHAVAN (Member, IEEE) received the Ph.D. degree in microelectronics from the Division of Microelectronics, Nanyang Technological University (NTU), Singapore, in 2012. He is an Assistant Professor at the Engineering Product Development (EPD) Pillar, Singapore University of Technology and Design (SUTD). Prior to this, he was a Postdoctoral Fellow at the Massachusetts Institute of Technology (MIT), Boston, and IMEC, Belgium, in joint association with Katholieke Universiteit Leuven (KUL). To date, he has authored/coauthored more than 220 international peer-reviewed publications and five invited book chapters. His work focuses on reliability assessment, characterization, and lifetime prediction of nanoelectronic devices as well as material design for reliability, uncertainty quantification, and prognostics and health management of electronic systems. He was an Invited Member of the IEEE GOLD Committee, from 2012 to 2014. He was the recipient of the IEEE Electron Device Society (EDS) Early Career Award, in 2016, the Asia-Pacific recipient of the IEEE EDS Ph.D. Student Fellowship, in 2011, and the recipient of the IEEE Reliability Society Graduate Scholarship Award, in 2008. He served as the General Chair for IEEE IPFA 2021, Singapore, and has consistently served on the review committee for various IEEE journals and conferences, including IRPS, IIRW, IPFA, and ESREF. He is currently serving as an Associate Editor for IEEE Access.

• • •