Task failure prediction for wafer-handling robotic arms by using various machine learning algorithms

Ping Wun Huang and Kuan-Jung Chung

Abstract
Industries are increasingly adopting automatic and intelligent manufacturing in production lines, such as those of semiconductor wafers, optoelectronic devices, and light-emitting diodes. For example, automatic robot arms have been used for pick-and-place workpiece applications. However, repairing automatic robot arms is time-consuming and increases the downtime of equipment and the cycle time of manufacturing. In this study, various machine learning (ML) models, such as the general linear model (GLM), random forest, extreme gradient boosting, gradient boosting machine, and stacked ensemble, were used to predict the maximum Cartesian positioning shift (i.e. the maximum eccentric distance) in the next handling time period (e.g. 1 min). A charge-coupled-device-based fault diagnostic system was developed to measure the critical positions of the robotic arm when transferring wafers. A novel data augmentation method was used to determine the correlation parameters in the dataset for the ML models. The prediction error for each algorithm was determined using the root mean square error (RMSE). The results revealed that the GLM exhibited the lowest prediction errors. The RMSEs of the GLM were 0.024, 0.032, and 0.046 mm for 3421 pickups, the last 1000 pickups, and 100 pickups, respectively, for the prediction target. Thus, the GLM is a promising model for predicting the task failure of wafer-handling robotic arms.

Keywords
Machine learning, wafer-handling robotic arm, positioning shift measurement, data augmentation, predictive maintenance

Date received: 16 February 2021; accepted: 2 March 2021

Introduction
An industrial robot is a programmable piece of mechatronic equipment that is generally used in production lines to perform repetitive or dangerous tasks automatically for ensuring high fabrication quality. For instance, numerous robots are used in semiconductor wafer, optoelectronic device, and light-emitting diode fabrication industries to perform various tasks. Typically, triaxial robotic arms are placed at etching, deposition, and photolithography stations to pick and place wafers between cassettes with high accuracy, repeatability, and reliability. However, repairing robotic arms is time-consuming and considerably increases the downtime of equipment and the cycle time of manufacturing. Figure 1 displays a real-world example in which three wafers were dropped by a robotic arm when transferring them. The subsequent maintenance procedures were complex and time-consuming. First, an engineer collected all the damaged pieces from the working area. Then, the failed components of the robot were inspected and repaired. Finally, the handling process was restored, and a trial run was performed to ensure that the robotic arm was repaired. The task failure of handling robots is the main cause of yield loss in semiconductor manufacturing. Therefore, a reliable method for reducing the occurrences of task failures should be developed urgently to improve the availability of handling robots.

In 1993, Ntuen and Park proposed a system-level methodology that involves using a reliability block diagram (RBD) to evaluate the reliability of certain robots. The module was divided into mechanical, sensor, and control modules in series. Subsequently, Korayem and Iravani applied failure mode and effect analysis (FMEA) and quality function deployment to improve the reliability and operational quality of the
3P and 6R robotic arms. Yamada and Takata⁵ and Lueth et al.⁶ investigated the optimal factors to reduce the operational failures of robotic arms when performing specific tasks. Conventional reliability analysis was performed in the aforementioned studies to evaluate the failures of various types of robot arms for certain tasks.⁷ The errors were instrumental in the development of effective solutions to improve the operational reliability. However, these solutions are only effective for certain work conditions. To address this problem, proactive reliability assessment methods or prognostics and health management (PHM or predictive maintenance) have been used in real-time situations for equipment maintenance. In most cases, PHM is used to predict the residual useful life (RUL). This model actively and recursively predicts and updates information periodically regarding the monitored critical parameters. Therefore, the state of health (SOH) of assets can be determined. Furthermore, the model assists engineers in determining the assets’ best time to maintenance (BTTM).

Proactive models are categorized into model-based, data-driven, or hybrid models. The data-driven model mainly relies on the Wiener or Gamma stochastic process to determine the performance degradation routes for devices under test (DUTs).⁸⁻¹³ The lifetime of a product is defined as the time when the specific value of the product first passes the critical first passage time. Machine learning (ML)-based data-driven models have been applied to predict lifetimes and monitor the SOH of critical products.¹⁴⁻¹⁹ These models use neural networks or other ML algorithms to determine the aging behavior of assets and achieve high prediction accuracy. Although proactive techniques effectively update the RUL, the prediction frequency of these techniques for some applications, especially in the case of fast degradation or immediate failure behaviors, is unreliable. For instance, when performing in situ prognostics of electronics, proactive techniques achieved the worst prediction accuracy of short circuits because the transition from the abnormal state to the failure state was fast.²⁰ Furthermore, the RUL is not adequate for predicting discrete and repetitive task failures of handling robots because the success of the next one handling cycle is more concern than the prediction of residual useful life. For a real situation, the specimen drops on the current pick-up cycle but the prediction of RUL displays that the failure will occur on many handling cycles later. Thus, it is necessary to develop effective methodology to predict the probability of success for the next one pick-up cycle of robotic arms.

In this study, a proactive data-driven methodology based on ML was developed to evaluate task reliability, that is, the maximum Cartesian positioning shift in the next time interval, which is the main parameter related to failure. A charge-coupled-device-(CCD)-based fault diagnostic system was used to capture pictures of the pickup motion of robotic arms. The X–Y position of the arm was then determined from the images by using a vision identification software. Considering the requirement of reliable, cost-effective, and precise prediction quality in smart manufacturing, the supervised statistical-based machine learning methodologies regarding with the general linear model (GLM), random forest (RF), extreme gradient boosting (XGB), gradient boosting machine (GBM), and stacked ensemble (SE) were applied to predict the maximum Cartesian positioning shift (i.e. the maximum eccentric distance) in the next 1 min time interval. To improve the prediction accuracy, a novel data augmentation method that combines root cause analysis and statistical parameter calculation was used to determine the failure causation and parameter correlation. Uncertainty analysis was performed by calculating the root mean square error (RMSE) for all algorithms to select the best ML model (i.e. the model with the lowest RMSE).

**Wafer-handling diagnostics and prognostics system**

Figure 2 displays the layout of the T-axis, R-axis, Z-axis, and Table 1 presents the specifications of the robotic arm. This triaxial robot arm is typically used in semiconductor production lines to pick wafers from the cassette and place them on the chuck for manufacturing. On completion of the manufacturing process, the robotic arm removes the wafers from the chuck and places them back on the cassette. Typically, the T (θ)-axis denotes the rotation of the arm, the R (radial)-axis denotes the reach and retraction ability of the arm, and the Z (vertical)-axis denotes the vertical movement. Figure 3 displays the robot fault diagnostics system that consists of a wafer-handling robotic arm, chucks, cassettes with wafers, a CCD camera module, and a...
laptop that is used as the control unit for the system.\textsuperscript{21} A visual identification software program was used to record and translate pixels of the photograph to X–Y position data. In the diagnostic process, first, all the process parameters were decided and the initialization process was performed by home-positioning the robot. Second, when the robot arm began its first task to deliver the test wafer to the chuck, the CCD camera captured the motion. The software program recorded the geometry of the wafer and computed the central point of the wafer \((X_0,Y_0)\), which was the original point for vision identification, shown in Figure 4. The robot performed its designated tasks and the vision subsystem measured all the in-process wafers. The central points of the first \((X_1,Y_1)\) to Nth \((X_n,Y_n)\) wafers were recorded on a computer which was recognized to be the original/raw data.

The raw data have to be preprocessing before they were used as input dataset for machine learning. First step of doing data preprocessing was to calculate the XY-shift, and the second step was to perform the data augmentation shown in the later section 3.4. In the prognostic model, the threshold value is defined by the critical eccentric distance that precedes failure. Thus, the maximum acceptable limit of the position deviation during the wafer-handling tasks is determined. The eccentric distance is calculated using the following expressions: \(X_n - X_0\) and \(Y_n - Y_0\). Typically, the maximum eccentric distance of the XY-shift can be calculated using the following expression: 
\[
\sqrt{(X_n - X_0)^2 + (Y_n - Y_0)^2}
\]

The threshold value is \(\pm 1\) mm for each axis. In reality, it is difficult to promptly predict the maximum eccentric distance for next one pick-up from present pick-up data because the calculation time is too short. Instead, a short time interval (i.e. 1 min) including several pick-ups were used in this study. Thus, the handling failure is referred that

Table 1. PRI robot arm specifications.

| Repeatability (\(\mu\)m) | Wafer sizes (mm) | User interface |
|--------------------------|------------------|---------------|
| \(\pm 25\)               | 75–300           | EQT32         |


![Figure 2. Handling robot and the T-axis, R-axis, and Z-axis.](image)

Figure 2. Handling robot and the T-axis, R-axis, and Z-axis.

![Figure 3. Fault diagnostic system of the robotic arm\textsuperscript{21}: (a) framework and (b) system in operation.](image)

Figure 3. Fault diagnostic system of the robotic arm\textsuperscript{21}: (a) framework and (b) system in operation.

![Figure 4. Pick-up position shift.](image)

Figure 4. Pick-up position shift.
the predictive value of the maximum positioning shift (i.e. the maximum eccentric distance) in the next one time interval (i.e. 1 min) is higher than the threshold value.

**ML techniques**

Many novel ML algorithms have been developed for various applications. For instance, deep neural-based networks provide state-of-art frameworks for text-independent speaker verification, object detection, and media recommendation. However, conventional algorithms corresponding to bagging, boosting, decision trees, and linear-based regression methodologies provide comprehensive formulations and case studies to solve problems in industries and businesses. Furthermore, an ensemble of multiple statistic-based learning algorithms is used to obtain superior predictive performance to that of any of the constituent learning algorithms. In this study, measurement data were collected from the wafer-handling diagnostics and prognostics system and used as the input of five ML models, namely the GLM, SE, RF, GBM, and XGBoost. The following subsections briefly introduce these learning algorithms.

**SE**

SE is a supervised ensemble ML algorithm in which the optimal combination of a collection of prediction algorithms is obtained through a stacking process. SE enables regression, binary classification, and multiclass classification by using a select algorithm pool including GLM, GBM, RF, deep learning, and XGBoost. Wolpert introduced the SE method in the early 1990s and developed an approach to combine several base predictors into a high-level model for increasing prediction accuracy. Subsequently, the SE method was developed several times and is now called super-learner regression or stacked regression. A k-fold cross-validation process is generally applied to determine an optimal weighted combination of predictions from multiple learning algorithms. Optimality is defined using a specified objective function, such as the minimum RMSE for continuous data or the maximum area under the receiver operating characteristic curve for discrete data. Super-learner regression is a powerful tool for estimating complex real-world relationships while avoiding unsubstantiated parametric assumptions and overfitting to improve predictive performance. The general super-learner algorithm for prediction is described in accordance with the research of van der Laan et al. The learning dataset is defined as follows:

\[ X_i = (Y_i, C_i), \quad i = 1, \ldots, n \]  

where \( Y \) is the outcome of interest and \( C \) is a \( p \)-dimensional set of covariates. The objective is to evaluate the function \( \psi_0(C) = E(Y|C) \). This function can be expressed as the minimizer of the expected loss as follows:

\[ \psi_0(C) = \arg \min_{\psi} E[L(X, \psi(C))] \]  

where the loss function is typically the square of the error loss \( (L = (Y - \psi(C))^2) \). From a pool that comprises a collection of algorithms, one or multiple constituent algorithms can be selected for a given problem. In this study, GLM, GBM, RF, SE, and XGBoost were used in the general sense as any mapping from the data into a predictor. Data review were conducted for detecting patterns and outliers by visually navigating data before ML models were created. A k-fold cross-validation was applied to our time-series case that involved using a small subset of data for training, predicting the later data points, and then evaluating the accuracy of the predicted data points. The predicted data points were subsequently included as part of the next training dataset, and the subsequent datapoints were forecasted. The time-axis layout of the five-fold cross-validation was applied in this study.

**Tree-based ensemble machines**

RF is a popular supervised data-mining technique that comprises decision trees and uses both bootstrap aggregation (bagging) and random variable selection to construct trees. This method was introduced by Ho. He used a method to concordantly grow trees in random subsets of features. When growing trees to form a forest, the training data were projected into a randomly selected subspace involving each node or tree. Then, the randomized nodes were optimized to make a proper decision. Breiman extended this method of combining node optimization and bagging techniques to construct a forest with uncorrelated trees. This process is similar to a classification and regression tree procedure. The out-of-bag error was used as an estimate of the generalization error, and variable importance was measured through the permutation process. For regression, the prediction produced by each tree in the forest was summarized and processed using an averaging method. Subsequently, Breiman introduced gradient boosting techniques to develop various tree models, such as the gradient-boosted regression tree and GBM. GBM combines multiple weak trees and creates a strong learner to optimize predictive performance. XGBoost was introduced by Chen in 2014. Subsequently, XGBoost became a popular and powerful ML tool because it requires limited effort for data preprocessing and has few hyperparameters that require tuning. XGBoost exhibits superior data learning and antinoise ability to solve complex classification and regression problems. Furthermore, XGBoost is easier to implement than artificial-neural-network-based algorithms and can help decision makers clearly understand the working process and outcome generation in the predictive model. When training the
model, the training error of the previous tree is trained as the learning goal of the current tree, and the results of multiple trees are overlapped to obtain the final result. To obtain superior model capability, the optimal splitting node must be determined to minimize the objective function $O^{(i)}$, which is defined as follows:

$$
O^{(i)} = \sum_{i=1}^{n} l(y_i, f^{(i-1)}_i + f_k(x_i)) + \Omega(f_k) \quad (3)
$$

where $n$ is the number of training samples; $l(.)$ is the loss function between the observed value and the predicted one; $f_k(.)$ is the result of the prediction of the regression tree at the $k$th iteration; and $\Omega(f_k)$ is the regularization function, which can be expressed as follows:

$$
\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (4)
$$

where $\gamma$ and $\lambda$ are both model hyperparameters that users can define, $T$ is the number of leaf nodes, and $w$ is the leaf weight of the tree. The objective function is optimized and simplified using the following second-order Taylor expansion:

$$
O^{(k)} = \sum_{i=1}^{n} \left[ g_k(x_i) + \frac{1}{2} h_k^2(x_i) \right] + \Omega(f_k) \quad (5)
$$

where $g_k = \frac{\partial l(y_i, f^{(i-1)} + f_k(x_i))}{\partial f_k}$ and $h_k = \frac{\partial^2 l(y_i, f^{(i-1)} + f_k(x_i))}{\partial f_k^2}$ are first- and second-order gradient statistics. Finally, the optimal weight $w_m^*$ and the corresponding optimal objective function $O^*(q)$ are obtained as follows:

$$
w_m^* = -\frac{\sum_{i=1}^{n} g_k(x_i)}{\sum_{i=1}^{n} h_k + \lambda} \quad (6)
$$

$$
O^*(q) = -\frac{1}{2} \sum_{m=1}^{T} \frac{(\sum_{i=1}^{n} g_k(x_i))^2}{\sum_{i=1}^{n} h_k + \lambda} + \lambda T \quad (7)
$$

**GLM**

GLMs were introduced by Nelder and Wedderburn. GLMs are a combination of linear and nonlinear functions associated with a distribution of the exponential family involving normal, inverse normal, binomial, Poisson, gamma, and logistic functions. The GLM consists of the following three components:

- **Random component**: $n$ values of the response variable $(y_1, ..., y_m)$
- **Systematic component**: a linear composition for the regression model of $\eta = \beta x^T$, where $x^T = (1, x_1, x_2, ..., x_m), i = 1, ..., n$, represents the explanatory or independent variables
- **Link function**: a monotone and differentiable function $g$, which connects the random and systematic components and relates the response variable mean $(\mu)$ with the linear structure of

$$
\eta = \beta x^T + f \quad (8)
$$

Each distribution has a unique link function. For instance, the widely used distribution for the estimation of air pollution health impacts is the Poisson function denoted by $Y \sim \text{Poisson}(\mu)$. Its link function is $\eta = \log(\mu)$; thus, $\mu = e^\eta$. By using the matrix notation,

$$
\log(\mu) = \beta x^T + f
$$

where $Y = (y_1, ..., y_m)$, $f$ is the function $f$ evaluated at time $t = 1, ..., m$, and $x^T$ is the $m \times n + 1$ design matrix containing a column of 1s and the other explanatory times-series variables $(x_1, ..., x_n)$.

**Data review and augmentation**

Figure 5 displays the variations in the maximum eccentric distance with one-minute time intervals. The data were collected from the XY-shift measurements. The duration of the time-series data was 17,921 min, of which 14,500 min (80%) were used for training and 3421 min (20%) were used for testing (prediction). The failure threshold was set as 1 mm. Failure occurred in the 17,921st min, with the maximum eccentric distance being 1.132 mm. The profile of time-series data regarding the maximum eccentric distances during 17,921 min pick-ups exhibited a randomly vibrated-like curve that is great challenge for data-driven ML modeling. Additional labels had to be generated to identify the latent features, which resulted in the fast and complex variation of the maximum eccentric distance profile. In this study, a novel data augmentation method was developed by combining root cause analysis (RCA) with statistical parameter calculation to determine the causation and correlation of data for ML-based modeling. The RCA was performed through FMEA to find the error source of positioning shift. The results revealed that teeth wear of motor drive belts is a major failure mechanism to conduct the failure mode of positioning shift. Thus, the advanced parameters were created according to the measured data to enhance the
wear-induced degradation behavior of the belt. Table 2 presents the parameters of input dataset for ML modeling. The time interval of 1 min and the maximum value of XY-shift among 1-min pick-ups denoted by XY_Max are basic parameters from measurements. Different lookback sizes and statistic-based data augmentations are advanced parameters described in the following text.

First, various lookback sizes were used for the time-series dataset to simulate the accumulated wear behavior of the belt in the physical domain. The parameters of different “lookback” sizes at each time interval of the dataset are going to create various accumulated quantities of wear for machine to learn the critical failure behavior of the belt. This behavior is defined by the number of recent data points to be used when predicting each future value in the time-series domain. The lookback size \( n \) for the present data denotes that the maximum eccentric distance (XY_Max) for recent past \( n \)-minute intervals, including the present time interval, is named XY\(_{n}\)Max, \( n = 1, ..., i \). For instance, XY\(_{10}\)Max denotes the “maximum value” of XY\(_{n}\)Maxs at the present minute interval where it is selected by looking back the values of XY\(_{n}\)Maxs at recent past 10 1-min intervals, including the present time interval. In this study, lookback sizes of 10, 30, 50, and 100 were used.

Second part of advanced parameters was the statistic-based data augmentation detailing the average, and standard deviation for XY_Max considering look size \( n \). For instance, “XY\(_{10}\)Ave” shown in Table 2 denotes the “average value” of XY\(_{n}\)Max at the present minute interval where it is selected by looking back the values of XY\(_{n}\)Maxs at recent past 10 1-min intervals, including the present time interval. Figure 6 displays the comparison of XY\(_{n}\)Max and XY\(_{10}\)Max profiles in the maximum eccentric distance with 1-min time intervals. It can be seen that XY\(_{n}\)Max and XY\(_{10}\)Max represent the almost same variety of maximum eccentric distances during the whole time intervals, but the values of XY\(_{10}\)Max are greater or equal to ones of XY\(_{n}\)Max to reveal higher probability of dropping wafer. From the maintenance point of view, it is better consideration to use XY\(_{10}\)Max as a prognostic parameter. In this study, XY\(_{10}\)Max was selected to be the prediction function (target) to determine the maximum eccentricity in the next 1 min time interval.

Results and discussion

Table 3 displays the significant test conditions of tree-based ensemble machines regarding with XGB, GBM, and RF. Compared to the GBM and RF, XGB used the largest number and the smallest max. depth of trees. XGB and GBM also represent the higher sample rates than that of RF. Table 4 shows the significant test conditions of GLM. The Gaussian was used as the family and link distribution, and regulation was applied to solve problems with overfitting. Besides, the test conditions of SE were quite simple that the GLM was selected as a meta learner, and other four algorithms were ensembled to obtain the optimal combination. Table 5 presents the training evaluation by five-fold
cross validations for regression. It can be seen that GLM represents the best results of validation regarding with the average 0.9999 and standard deviation 0.0000. However, these five algorithms demonstrate the premium results for data training. They have already to predict the target (XY10_Max) for the test dataset.

Figure 7 illustrates the prediction profiles of the five ML algorithms for XY10_Max. These ML models satisfactorily predicted the target because all profiles of prediction were similar to that of the real data. The prediction accuracy of the ML models was evaluated using the RMSE (Figure 8), which is a standard measure for determining the error of an ML model in predicting quantitative data. In this study, 3421 pickups (the entire test data), the last 1000 pickups, and the last 100 pickups were used to evaluate each algorithm. The GLM achieved the best prediction with the lowest RMSEs (average RMSE was 0.034 mm) for the

Table 3. The significant test conditions of XGB, GBM, and RF.

| Algorithms | Test conditions | Tree-based ensemble machines | XGB | GBM | RF |
|------------|----------------|-------------------------------|-----|-----|----|
| Histogram type | N/A | N/A | Uniform Adaptive |
| Booster | XGB Tree | N/A | N/A |
| Number of trees | 611 | 293 | 50 |
| Max. depth | 5 | 7 | 20 |
| Min. rows | 3 | 10 | 1 |
| Min split improvement | 0 | 0.0001 | 0.0001 |
| Model type | Regression | Regression | Regression |
| Numerical imputation | Model based | Model based | Model based |
| Sample rate | 0.8 | 0.8 | 0.632 |
| Learning rate | 0.05 | 0.1 | N/A |

Table 4. The significant test conditions of GLM.

| Algorithms | Test conditions | GLM |
|------------|----------------|-----|
| Model type | Regression | Regression |
| Family | Gaussian | Gaussian |
| Regularization | Alpha | Lambda |
| 2.2085 × 10⁻⁵ |
| Numerical imputation | Mean | Mean |
| Precision of the vector | 0.0001 | 0.0001 |
| Standardize | True | True |

Table 5. The training evaluation by 5-fold cross validations.

| Evaluation | Algorithms | 5-fold cross validation scores (regression R²) | Average | Std |
|------------|------------|--------------------------------------------|---------|-----|
| GLM | 0.9999 | 0.0000 |
| SE | 0.9992 | 0.0001 |
| XGB | 0.9973 | 0.0002 |
| GBM | 0.9972 | 0.0003 |
| RF | 0.9960 | 0.0005 |

Figure 7. Prediction profiles of the five ML algorithms for XY10_Max.
aforementioned three conditions. The GBM provided the next best prediction with an average RMSE of 0.049 mm. The combination of the linear and nonlinear functions with the exponential family in GLM provided a flexible and fast response to the randomly vibrating profile, which resulted in high prediction accuracy. Furthermore, if the number of weak trees is increased to optimize predictive performance in the GBM, results similar to those of the GLM may be obtained. The remaining algorithms also exhibited high RMSEs because all the predictions are close to the prediction function XY10_Max. It can be seen in Figure 7 that they are likely to follow the “tempo” of vibration-like profile.

Figures 9 and 10 present the prediction profiles of the five ML algorithms for the last 1000 and 100 pickup motions, respectively. Unlike other profiles, the prediction profile of the GLM is nearly located on top of the real data profile, which indicates that most of the XY10_Max values were marginally overpredicted (i.e. the predicted value is greater than the real value). Therefore, the GLM is a conservative ML model for the prediction of maximum eccentric distances. As mentioned previously, it is better consideration to maintenance because the expectation of good prediction should be more conservative to stop the machine prior to failure. In this study, task failure occurred in the 17,921st pickup whose value of maximum eccentric distance (1.132 mm) was greater than the one of failure threshold (1 mm), and only the GLM provided a successful prediction in the previous time period (predictive value was 1.063 mm). Before 17,921st pickup, only GLM provided a failure prediction at 17,880st pickup, which the predictive value was 1.023 mm and real value is 0.972 mm. Although it was a little early for predicting failure, it provided a significant warning information to notice that the failure would occur in the near future. In summary, the GLM is the best ML model for task failure prediction.

Conclusions
Proactive reliability assessment by using ML is a promising method for evaluating the SOH of assets in operation and thus determine the BTTM of the assets. In this study, the GLM, RF, XGB, GBM, and SE were used to predict the task failure in terms of the maximum Cartesian positioning shift (i.e. the maximum eccentric distance) in the next handling time period (e.g. 1 min). The comparative prediction results revealed that the GLM exhibited the lowest prediction error (RMSE) and was the only model to successfully predict failure. Therefore, the GLM is a promising model for predicting task failures for wafer-handling robotic arms. In the future, numerous ML-based devices can be used on production lines to predict the maximum eccentric distances for wafer-handling robotic arms.

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this
References

1. Lai MC. Multiobjective path planning for robot manipulators with joint clearance. Taiwan: National Cheng Kung University, 2014.

2. Huang PW and Chung KJ. The prediction of positioning shift for a robot arm using machine learning techniques. In: 14th international microsystems, packaging, assembly, circuits technology conference, Taipei, Taiwan, 23–25 October 2019, pp.58–61. New York: IEEE.

3. Ntuen CA and Park EH. A formal method to characterize robot reliability. In: Annual reliability and maintainability symposium, Atlanta, GA, USA, 26–28 January 1993, pp.395–398. New York: IEEE.

4. Korayem MH and Iravani R. Improvement of 3P and 6R mechanical robots reliability and quality applying FMEA and QFD approaches. Robot Comput Integr Manuf 2008; 24(3): 472–487.

5. Lueth TC, Nasal UM and Rembold U. Reliability and integrated capabilities of locomotion and manipulation for autonomous robot assembly. Rob Autom Syst 1995; 14(2): 185–198.

6. Yamada A and Takata S. Reliability improvement of industrial robots by optimizing operation plans based on deterioration evaluation. CIRP Ann Manuf Technol 2002; 51(1): 319–322.

7. Dhillon BS. Robot reliability: an investigative study. In: The 24th ISSAT international conference on reliability and quality in design, Toronto, Canada, 2–4 August 2018. NF, ISSAT.

8. Doksum KA and Ho’yland A. Models for variable-stress accelerated life testing experiments based on wiener processes and the inverse gaussian distribution. Technometrics 1992; 34(1): 74–82.

9. Singpurwalla ND. Survival in dynamic environment. Stat Sci 1995; 10(1): 86–103.

10. Doksum KA and Normand SL. Gaussian models for degradation processes-part I: methods for the analysis of biomarker data. Lifetime Data Anal 1995; 1(2): 131–144.

11. Whitmore GA and Schenkelberg F. Modeling accelerated degradation data using wiener diffusion with a time scale transformation. Lifetime Data Anal 1997; 3(1): 27–45.

12. Lin YC and Chung KJ. Lifetime prognosis of lithium-ion batteries through novel accelerated degradation measurements and a combined gamma process and Monte Carlo method. Appl Sci 2019; 9(3): 559.

13. Huang JH, Cheng HM, Li YR, et al. Prognostic-based reliability assessment of a wafer handling robot arm using wiener stochastic process model and Monte Carlo simulation. In: 24th ISSAT international conference on reliability and quality in design, Toronto, Canada, 2–4 August 2018. New Jersey: ISSAT.

14. Lamoureux B, Massé JR and Mecheb N. An approach to the health monitoring of a pumping unit in an aircraft engine fuel system. In: first European conference of prognostics and health management society, Dresden, Germany, 3–5 July 2012. PHM Society.

15. Herzog MA, Marwala T and Heyns PS. Machine and component residual life estimation through the application of neural networks. Reliab Eng Syst Saf 2009; 94(3): 479–489.

16. Wang WQ, Golnaraghi MF and Ismail F. Prognosis of machine health condition using neuro-fuzzy systems. Mech Syst Signal Process 2004; 18(4): 813–831.

17. Li Z, Wu D, Hu C, et al. An ensemble learning-based prognostic approach with degradation-dependent weights for remaining useful life prediction. Reliab Eng Syst Saf 2019; 184(1): 110–122.

18. Heinermann J and Kramer O. Machine learning ensembles for wind power prediction. Renew Energy 2016; 89(1): 671–679.

19. Zhao R, Yan R, Chen Z, et al. Deep learning and its applications to machine health monitoring. Mech Syst Signal Process 2019; 115(1): 213–237.

20. Vichare NM and Pecht MG. Prognostics and health management of electronics. IEEE Trans compon packaging technol 2006; 29(1): 222–229.

21. Chung KJ, Lin YC and Wu BH. Proactive-based reliability assessment of handling robot arm. In: Proceedings of 2017 world congress on micro and nano manufacturing, Kaohsiung, Taiwan, 27–30 March 2017.

22. Xiao Y, Wu J, Lin Z, et al. A deep learning-based multi-model ensemble method for cancer prediction. Comput Methods Programs Biomed 2018; 153(1): 1–9.

23. Zhang D, Jiao L, Bai X, et al. A robust semi-supervised SVM via ensemble learning. Appl Soft Comput 2018; 65(C): 632–643.

24. Wolpert D. Stacked generalization. Neural Netw 1992; 5(2): 241–259.

25. Laan MJ and Dudoit S. Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon-net estimator: finite sample oracle inequalities and example. Technical Report 2003; 30. Division of Biostatistics, University of California, Berkeley, 2003.

26. Laan MJ, Dudoit S and Vaart AW. The cross-validated adaptive epsilon-net estimator. Stat Decis Int Math Jr Stoch Methods Model 2006; 24(3): 373–395.

27. Laan MJ, Polley EC and Hubbard AE. Super learner. Stat Appl Genet Mol Biol 2007; 6(1): 25.

28. Breiman L. Stacked regressions. Mach Learn 1996; 24: 49–64.

29. Naimi AI and Balzer LB. Stacked generalization: an introduction to super learning. Eur J Epidemiol 2018; 33(5): 459–464.

30. Ho TK. Random decision forests. In: Proceedings of 3rd international conference on document analysis and recognition, Montreal, Canada, 14–16 August 1995, pp.278–282. New York: IEEE.

31. Breiman L. Random forests. Mach Learn 2001; 45 (1): 5–32.

32. Huang L, Li Y, Chen S, et al. Building safety monitoring based on extreme gradient boosting in distributed optical fiber sensing. Opt Fiber Technol 2020; 55: 1–6.
33. Nelder JA and Wedderburn RW. Generalized linear models. *J Royal Stat Soc* 1972; 135(3): 370–384.
34. McCullagh PM and Nelder JA. *Generalized linear models.* 2nd ed. London, UK: Chapman and Hall, 1989.
35. Belotti J, Castanho D, Araujo L, et al. Air pollution epidemiology: a simplified generalized linear model approach optimized by bio-inspired metaheuristics. *Environ Res* 2020; 191: 1–8.