IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 8 April 2024, accepted 18 May 2024, date of publication 27 May 2024, date of current version 3 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3405540

RESEARCH ARTICLE

Fuzzy Few-Shot Learning for Predictive Maintenance and Cost-Benefit Analysis in Semiconductor Manufacturing: A Case Study on Photomask Haze

HAO-HSUAN HUANG AND YUN-HSUN HUANG
Department of Resources Engineering, National Cheng Kung University, Tainan 701, Taiwan
Corresponding author: Yun-Hsun Huang (z10808014@email.ncku.edu.tw)

This work was supported in part by the Ministry of Science and Technology of Taiwan under Grant 112-2410-H-006-079-.

ABSTRACT Predictive Maintenance (PdM) plays a key role in production management by extending the lifespan of components and reducing maintenance costs. In the manufacture of semiconductors, the scarcity of data related to haze defects makes it difficult to draw correlations between environmental factors and haze formation. The uncertainty associated with a reliance on indirect evidence (i.e., cumulative time in the environment) has seldom been explored in the literature. Therefore, we developed a PdM framework based on fuzzy few-shot learning to deal with photomask haze in the semiconductor industry. The robustness of the model was evaluated in a three-month pilot study conducted in a wafer foundry. This paper also provides cost-benefit analysis of model implementation. The results demonstrate that the proposed haze photomask detection model outperforms existing models in terms of hit rates and false alarm rates. It also proved effective in lowering labor costs and power consumption, as evidenced by the fact that the number of haze candidates was well below the daily inspection cap, which allowed the decommissioning of one photomask inspection device, reducing the photomask inspection volume by 21%, which is equivalent to an annual labor cost reduction of USD 18,780. These results should help to promote the adoption of predictive maintenance applications, even in situations with small sample sizes.

INDEX TERMS Fuzzy few-shot learning, photomask haze, predictive maintenance, imbalanced data, cost-benefit analysis.

I. INTRODUCTION

A. BACKGROUND

In semiconductor manufacturing, the maintenance of production components is crucial to product quality and production yield [1]. The manual inspection of hardware can have a profound negative impact on production capacity and associated costs (reaching 15% - 70% of total production expenses) [2]. The advent of smart factories has led to the widespread adoption of Predictive Maintenance (PdM), which uses statistical analysis and machine learning based on historical data to determine the optimal timing of maintenance interventions [3].

PdMs use a record of historical usage in conjunction with the current operations to predict the health status and thereby improve component reliability and usability. PdM is generally implemented in three stages: component-related data collection, data processing and analysis, and PdM decision-making [4]. Generally, PdM models employ two approaches to decision-making: Remaining Useful Life (RUL) prediction and component status classification. RUL prediction typically employs Time-To-Failure (TTF) as a quantitative prediction target, whereas component status classification involves the categorization of components as normal versus susceptible [3]. Common quantitative prediction methods include

The associate editor coordinating the review of this manuscript and approving it for publication was Rongbo Zhu.

statistical approaches [5], [6], time series methods [7], machine learning techniques [8], [9], [10], deep learning methods [11], [12], as well as Continuous Semi-Markov Decision Process (CSMDP) combined with Reinforcement Learning (RL) [13]. Recent prediction methods consider long-range dependencies and heavy tail characteristics in the assessment of RUL [14]. The methods for classification tasks include Hierarchical Clustering (HC) [15], Support Vector Machine (SVM) [16], [17], Decision Tree (DT) [18], and Random Forest (RF) [19]. In the field of deep learning, methods such as Convolutional Neural Network (CNN) combined with LSTM have been employed [20]. All these methods have been validated for their effectiveness in building PdM models for RUL prediction or fault classification.

Most PdM models generally require a large amount of failure data for training aimed at estimating the RUL or predicting impending anomalies [21], [22]. Data collection is hindered by a severe imbalance between normal data (abundant) and abnormal data (scarce), which can render conventional machine learning or deep learning methods ineffective [23]. Several researchers have proposed the pre-training of deep learning models or the use of transfer learning to reduce data size requirements [24]; however, engineers have found that the inference process lacks interpretability. This can make it difficult to reverse-engineer the relationships between critical influencing factors and prediction results, thereby hindering the definition of standards for the maintenance of components. Data augmentation is another approach to increasing sample size; however, the representativeness of newly generated virtual samples is often questionable. Note also that the heterogeneity of components due to material variations, vendor differences, and disparities in usage can make it difficult to accurately quantify an RUL for each component [25]. Finally, it is exceedingly difficult to obtain data directly related to the lifespans of components [26]. The results obtained using indirect data introduce a degree of fuzziness and uncertainty. Despite these issues, there has been very little research on the establishment of data-driven PdM models under data constraints.

Machine learning and deep learning methods both require large amounts of data for model training and the avoidance of overfitting, and even statistical methods require at least 30 abnormal samples to establish statistical significance. A lack of data suitable data and uncertainties associated with data collected via indirect methods can lead to suboptimal prediction and classification outcomes with potentially grievous effects on the lifespan of components. Although there have been studies in recent years investigating few-shot learning for data imbalance [27], [28], [29], there is limited research on building few-shot learning models using indirect variables to address the scarcity of key direct variables affecting component abnormalities. Most previous research on photomask haze have concentrated on physical and chemical phenomena. Few studies have adopted a data-driven approach to the exploration of PdM.

In the current study, we sought to fill this gap by developing a PdM framework based on Fuzzy Few-Shot Learning (FFSL) to alleviate the problem of data imbalance by training the PdM model using only abnormal data. We addressed the problems of components heterogeneity and fuzziness in indirect data by implementing attribute classification and fuzzy similarity, respectively. The framework is implemented in two stages: detection and self-monitoring. In the detection stage, Fuzzy Similarity (FS) is used to determine whether the features of a given normal component are similar to those of abnormal components in a form of one-class classification. In the event that the model incorrectly identifies a component as abnormal when it is actually normal, then that component enters a self-monitoring stage in which a Self-inspection Index (SI) is used to determine whether the inspection criteria have been met.

The proposed model was validated in a three-month pilot run involving the detection of photomask haze in a semiconductor wafer foundry. The study had three main objectives. First, we sought to establish a framework and operational process for FFSL-based PdM. Second, we compared model performance metrics with those of common machine learning models, including hit rate, false alarm rate, and prediction loading rate (indicating the number of photomasks that need to be inspected per unit of manpower for use in assessing the impact of the model on inspection workload). Third, we evaluated the cost savings attributable to the implementation of FFSL-based PdM model in the factory.

This study makes the following contributions. First, this is the first study to develop a PdM model that addresses the challenges of limited sample size, data heterogeneity, and fuzziness. We also derived crystallization dynamic features associated with haze formation to be used in the PdM model. Note that this innovation is applicable to the establishment of PdMs for a wide range of consumables or components with limited data pertaining to abnormal operations. Second, the model was implemented as a pilot run in a wafer foundry, during which data was collected on 50,000 photomasks over a period of 3 months. Note that this was the first study on photomasks to cover PdM research, and first PdM study to be verified through big data analysis. Third, we developed a practical metric by which to measure the labor costs saved by the PdM. The proposed prediction loading rate makes it far easier to assess the cost-benefit of model implementation, which has been largely disregarded in the literature.

B. CASE DESCRIPTION OF PHOTOMASK HASE DETECTION

The process of manufacturing semiconductors is complex, typically involving 100∼200 processing steps of high precision [30]. In photolithography, Ultra-Violet (UV) light is used to transfer patterns from photomasks to wafers; however, the on-going miniaturization of electronic components has reduced linewidths (from 350∼500 nm to 100∼250 nm) and exposure wavelengths (from 436 nm to 193 nm), thereby increasing the likelihood of transferring defects in

the photomask layer to the wafer during the photolithography process.

Photomask types can be categorized as Binary Intensity Masks (BIM) or Phase Shift Masks (PSM). Under exposure to a light beam, BIMs tend to generate diffraction patterns and image blurring. This issue can be resolved by creating a PSM, which involves applying a phase-shifting layer atop the BIM with the aim of inverting the phase of the light beam to improve image resolution. PSMs are widely utilized in Deep Ultra-Violet (DUV) and Extreme Ultra-Violet (EUV) lithography; however, PSMs are more susceptible to haze formation than are BIMs [31].

Haze occurring between the photomask surface and the pellicle is the most common form of photomask defect. From a chemical perspective, haze is mainly ammonium sulfate ($(NH_4)_2SO_4$) [32]. The mechanism underlying haze formation involves the adsorption of sulfur ions (S^{2-}) from the air onto the photomask surface, which then oxidizes to form SO_2. The exposure stage of photolithography involves the use of light beams (wavelength = 193) carrying energy sufficient to cause the photochemical decomposition of oxygen molecules into oxygen atoms. The SO_2 then combines with oxygen atoms as SO_3, which forms into sulfuric acid in the presence of H_2O in the air. The presence of any NH_3 in the air during this stage can lead to the formation of ammonium sulfate crystals, resulting in haze formation [33].

The factors that contribute to haze formation can be categorized as photomask attributes, cleaning processes, environmental factors, and exposure energy-induced effects [34]. The risk of ammonium sulfate adsorption (haze formation) is positively correlated with the presence of photomask impurities (S^{2-} and NH_3) and light energy, such that shorter wavelengths (193 nm) are more likely than long wavelengths (248 nm) to induce haze formation [31]. Haze formation can also be attributed to cleaning methods and environmental conditions [35]. Haze does not form when the concentration of sulfur ions is less than 5 ppb, even under an accumulated exposure energy of 30 kJ/cm^2 [36]. Even in the presence of sulfur ions on the photomask surface, haze does not form under low NH_3 concentrations when using exposure energy of 10 kJ/cm^2 [37]. In other words, haze formation requires that both sulfur ions and NH_3 exceed given concentrations. The photomask storage environment is the primary source of sulfur ions and NH_3 on photomasks [38], [39]. Researchers have demonstrated that the lifespan of a photomask can be extended simply by improving the storage environment or reducing the time that photomasks remain in potentially contaminated environments. However, due to the varying environmental exposure paths or footprints of photomasks driven by their production requirements, this time variable introduces uncertainty and fuzziness into the haze [40], [41]. The risk of haze formation is proportional to the number of exposures and cumulative exposure energy [31]. The threshold for cumulative exposure energy (at 193 nm) depends on the density of ammonium sulfate seeds. Haze formation occurs when the seed density exceeds a given threshold, such that the exposure energy is sufficient to induce seed agglomeration [36]. The timely cleaning of photomasks can eliminate haze and allow the continued use of photomasks [37]. However, the large number of photomasks that pass through the production line each day can make it difficult to differentiate haze-contaminated photomasks. Engineers require methods to facilitate the detection of haze and thereby reduce the time required for manual inspections.

In this case, the existing methods use SVM to select the photomasks that need to be inspected. However, this assembly line had only 5 personnel using only 5 photomask inspection devices to inspect 500∼600 photomasks each day, thereby limiting the daily inspection capacity to 75 photomasks. Under these constraints, it was inevitable that many haze-contaminated photomasks would go undetected. In developing the proposed PdM model, our aim was to ensure that photomasks would be sent for cleaning in a timely manner and reduce costs due to unnecessary manual inspection.

The remainder of this paper is organized as follows. Section II presents an overview of the FFSL methodology, including its capabilities and applications. Section III outlines the data sources and processing. Section IV presents our research results. Section V summarizes the findings and provides conclusions and recommendations for future work.

II. PROPOSED METHODOLOGY

Figure 1 illustrates the proposed PdM framework based on FFSL, which is implemented as follows: (1) Identify key factors influencing haze formation via cause-and-effect analysis, including categorical and numerical factors. (2) Categorize photomasks according to attributes (e.g., vendor), which could affect component lifespans. (3) Extract critical features of photomask haze with equations pertaining to crystalline dynamics. (4) Utilize fuzzy similarity in the detection stage to compare the similarity between the tested photomasks and those contaminated with haze. If the similarity threshold is exceeded, then an inspection is initiated. If the inspection returns normal results, then proceed to the self-monitoring stage. (5) In the self-monitoring stage, a self-inspection index determines whether the photomask meets the inspection criteria. The abbreviations used in the proposed PdM model and their corresponding explanations are provided in Table 1. Detailed explanations of each step are provided below.

A. CAUSE AND EFFECT ANALYSIS

Cause and effect analysis was performed using a fishbone diagram to identify the factors with a meaningful effect on the target. Figure 2 presents the factors with an influence on haze formation, based on previous research and expert knowledge. Photomask lifetime can be affected by categorical variables, such as photomask attributes (e.g., vendor manufacturing quality, photomask types, and pellicle types) and the cleaning process (e.g., cleaning method and contamination in

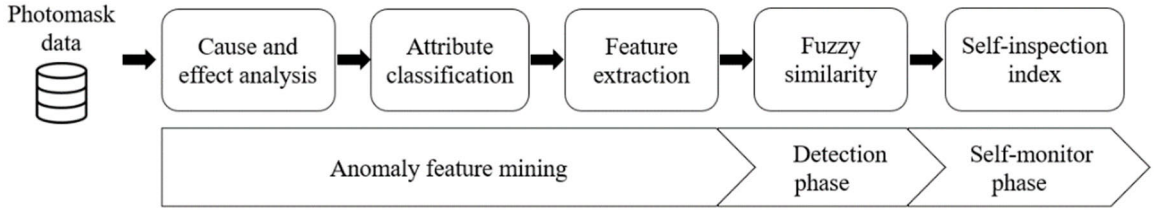

FIGURE 1. Fuzzy few-shot learning framework for PdM dealing with photomask haze.

TABLE 1. Explanation of abbreviation used in the proposed PdM.

Abbreviation	Explanation
Y	The transformation fraction, indicating the relative degree of crystallinity
k	The crystallization rate constant
t	The duration time
a	The Avrami constant
exp	The natural exponential function
π	The circular constant
υ	The growth velocity of spherulites
d	The thickness of the crystal disk
N_0	The nucleation density
Y_i	The transformation fraction of the i-th inspection
F_i	The crystalline dynamics feature of the i-th inspection
E_i	The exposure energy of the i-th inspection
C_i^{ACC}	The cumulative exposure count at the time of the i-th inspection
T_i^{ACC}	The cumulative time in the contaminated environment at the time of the i-th inspection
F_i^e	The crystalline dynamics features in the e-th environment at the i-th inspection
Y_i^*	The set of five transformation fractions at the i-th inspection
$T_i^{ACC,e}$	The accumulated times at the i-th inspection in the e environments
H^{med}	The median of the transformation fraction set of all haze photomasks
H_j^*	The transformation fraction set of the j-th haze photomask
Med	The median function
ΔC_i^{ACC}	The change in cumulative exposure count at the i-th inspection
z	The count window for inspection frequency

the cleaning environment). During the movement of photomasks from storage rooms to production lines, their RUL is gradually shortened by exposure to energy-induced factors (e.g., exposure energy and cumulative exposure count) and environmental factors (e.g., S^{2-} and NH_3 concentrations). Note that the adsorption of these impurities is proportional to the time spent in the contaminated environment. This means that the cumulative time that photomasks remain in contaminated environments is an indirect factor that should be considered in quantifying the retention of S^{2-} and NH_3 on photomasks.

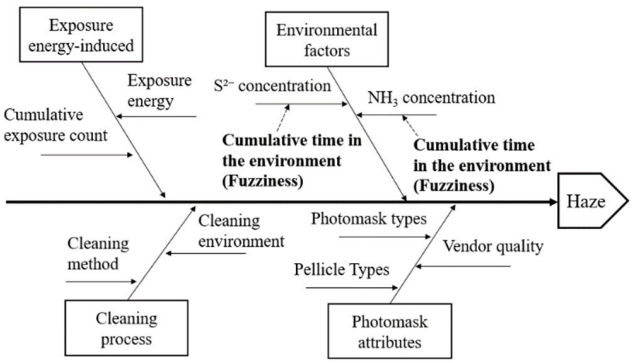

FIGURE 2. Cause-and-effect analysis pertaining to the formation photomask haze.

B. ATTRIBUTE CLASSIFICATION

The useful lifespan of a photomask depends on cleaning processes and photomask attributes. Categorical factors include pellicle types, photomask types, vendor quality, cleaning method, and cleaning environment. Considering that cleaning is performed by the same photomask vendor, we consolidated cleaning method and cleaning environment under the category 'vendor'. This categorization refers to variations in photomask lifespan attributable to the manufacturing quality, cleaning method, and cleaning environment associated with a given vendor.

C. FEATURE EXTRACTION OF PHOTOMASKS HAZE

This study employed the principles of crystalline dynamics proposed by Avrami in 1939 [42], [43], [44] to calculate haze features. As shown in Figure 3, the Avrami equation is derived through a Poisson distribution [45], [46], which suggests a positive sigmoidal relationship between the extent of crystal growth and time.

The relationship between crystal growth (Y) and time (t) can be written as follows:

$$Y = 1 - exp(-kt^a) \tag{1}$$

When the process of crystallization on the photomask is in the initial nucleation stage, the transformation fraction is close to 0% (indicating minimal nucleation). With time, the process advances into the nucleation growth stage, during which transformation fraction increases rapidly (indicating rapid crystal growth). When nucleation undergoes a gradual slowdown, the transformation fraction reaches or exceeds

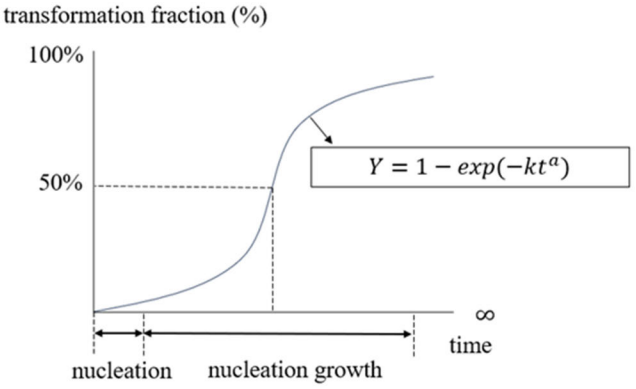

FIGURE 3. Stages of crystal growth as a function of time.

50% (indicating significant nucleation) [46]. Note that the occurrence of haze nucleation cannot be defined by the transformation fraction for other ratios, thus precluding the use of sensitivity analysis.

The Avrami constant is associated with nucleation type and the nucleation mechanism associated with a given material [47]. Nucleation type can be categorized as homogeneous (crystallization based on the properties of the solution) and heterogeneous (crystallization that requires the assistance of small external particles as seeds). Note that nucleation can also be one-dimensional (rod), two-dimensional (disk), or three-dimensional (sphere). Haze formation is associated with heterogeneous nucleation involving plate-like crystalline structures based on an Avrami constant of 2 [47]. The rate constant (k) for the crystallization of plate-like crystals in the process of heterogeneous nucleation can be expressed as follows. Note that to overcome the difficulty of obtaining data pertaining to the rate of microscopic crystallization, we simulated this data based on a photomask haze database ($k = 0.03$).

$$k = \pi \upsilon^2 dN_0 \tag{2}$$

Single exposure energy, cumulative exposure count, and cumulative time in an impure environment play key roles in the formation of haze on photomasks; therefore, we incorporated these factors into Eq. (1) to be reformulated as Eq. (3). The transformation fraction (Y_i) of haze is positively correlated with crystalline dynamics features (F_i), which can be calculated via multiplication using single exposure energy, the cumulative exposure number, and the cumulative time of environmental impurities using Eq. (4). These crystalline dynamics features indicate that haze formation can occur when a photomask is exposed to impurities for a set duration and continuously exposed to a set level of exposure energy.

$$Y_i = 1 - exp(-0.03F_i^2) \tag{3}$$

$$F_i = E_i \times C_i^{ACC} \times T_i^{ACC} \tag{4}$$

As shown in Eq. (5), there are five different levels of cumulative time of environmental impurities (T_i^{ACC}). These impurity environments (CDA, NCDA, SINV, INV, and FAB) are detailed in Table 2. By substituting these variables into Eq. (4), it is possible to derive the corresponding crystalline dynamics features (F_i^e) can be derived for each environment, as shown in Eq. (6). This means that Eq. (3) can be used to obtain five transformation fractions, as shown in Eq. (7). This means that the proportion of each factor changes with the Avrami equation. When photomask haze is detected, this set of transformation fractions can then be used to perform reverse inference indicating the transformation fraction and factors that are responsible. This could be highly beneficial for engineers seeking to interpret available data in tracing root causes.

TABLE 2. Impurity environments and risk of haze formation.

Storage environment	Explanation	Haze risk
CDA	Typical storage area, moderate cleanliness maintained using high-pressure dry air increases risk of minor impurities on the photomask surface.	Low
NCDA	Typical storage area lacking high-pressure dry air – inferior to CDA.	Medium
SINV	Semi-storage area described by engineering staff as a potential location for photomasks with defects or a history of exposure to impurities.	High
INV	Retreat area where photomasks are returned to the warehouse or transferred to the warehouse from other fab facilities. Poor environmental conditions.	High
FAB	Other areas outside through which photomasks are transported are difficult to control in terms of impurities.	Medium

$$T_i^{ACC} = \{T_i^{ACC,CDA}, T_i^{ACC,NCDA}, T_i^{ACC,SINV}, T_i^{ACC,INV}, T_i^{ACC,FAB}\} \tag{5}$$

$$F_i^e = E_i \times C_i^{ACC} \times T_i^{ACC,e}, \quad e = CDA, NCDA, SINV, INV, FAB \tag{6}$$

$$Y_i^* = \{Y_i^{CDA}, Y_i^{NCDA}, Y_i^{SINV}, Y_i^{INV}, Y_i^{FAB}\} \tag{7}$$

D. FUZZY SIMILARITY

Consider two photomasks that spend the same cumulative time in an impure environment but are transported along different paths. The two samples are likely to differ in terms of impurity adsorption, thereby introducing inherent fuzziness and uncertainty into the correlation between time spent in an impure environment and haze formation. Under these conditions, the transformation fractions calculated using the crystalline dynamics features will also tend to exhibit fuzziness.

In the current study, we employed Fuzzy Similarity (FS) [48], [49], [50] to derive a more complete assessment of haze-affected photomasks based on the five transformation fractions. The formula presented in Eq. (8) can be used to calculate the fuzzy similarity between the transformation fraction set of the tested photomask and the transformation

fraction set of a known haze-affected photomask. Here, the values range from 0 to 1 with a higher value indicating a higher risk of haze formation. Fuzzy similarity of > 0.5 indicates that the photomask is probably in the middle to late stages of nucleation and should be inspected due to an elevated risk of developing haze.

$$FS = 1 - \frac{\sum_{i=1}^{n} \left| H^{med} - Y_i^* \right|}{\sum_{i=1}^{n} \left(H^{med} + Y_i^* \right)}, \quad FS \in [0, 1] \tag{8}$$

It is important to consider that extreme outliers can be found in the transformation fraction sets of all haze photomasks. For example, it is plausible that haze could occur on a photomask even in situations where all five transformation fractions are very small. To mitigate the influence of outliers on haze features, we employ the median instead of the mean, as follows:

$$H^{med} = Med\left(H_j^*\right), \quad j = 1, 2, \ldots, m \tag{9}$$

E. SELF-INSPECTION INDEX

In some cases, photomasks identified as hazy based on fuzzy similarity are not actually hazy. Nonetheless, these photomasks should be continuously monitored using the Self-inspection Index (SI), because they face an elevated risk of developing haze as the exposure count increases. The self-inspection index is calculated as the change in the cumulative exposure count divided by the count window, which indicates the risk based on the cumulative time spent in the contaminated environment. Based on engineering experience, when the total cumulative time in an impure environment ($\sum_i^n T_i^{ACC}$) exceeds 2,000 hours, it is considered high risk. Thus, a tighter count window is usually set (600 counts). If the total cumulative time is less than 2,000 hours, it is considered medium risk, and a looser count window is set (1,000 counts). As shown in Eq. (10), a self-inspection index exceeding 1 indicates that the photomask should be inspected.

$$SI = \frac{\Delta C_i^{ACC}}{z}, \quad \left\{ \begin{array}{ll} z = 1{,}000, & if \ \sum_i^n T_i^{ACC} > 2{,}000 \\ z = 600, & if \ \sum_i^n T_i^{ACC} < 2{,}000 \end{array} \right\} \tag{10}$$

Another scenario is when photomasks are determined as normal based on fuzzy similarity but have haze. In this case, the occurrence is extremely rare. Detecting these cases with PdM involves significant labor costs. Therefore, it is more feasible to detect haze-contaminated photomasks retrospectively through abnormal semi-finished products.

F. PERFORMANCE INDICATORS AND PROCESS VALIDATION

In this paper, the hit rate, false alarm rate, and prediction loading rate are calculated using a confusion matrix. Hit rate (also known as recall) measures the proportion of photomasks correctly predicted by the model as hazy divided by the total number of haze-contaminated photomasks (see Eq. (11) and Table 3), wherein a higher index indicates more accurate haze detection. The false alarm rate indicates the proportion of normal photomasks falsely predicted by the model as haze, as shown in Eq. (12). Prediction loading rate measures the proportion of photomasks predicted by the model as haze divided by the daily inspection ceiling constraints imposed by the available human resources, wherein a higher index indicates a heavier workload, as shown in Eq. (13). In this case, the daily inspection limit (M) was set at 75 photomasks for a workforce of 5 individuals.

TABLE 3. Impurity environments and risk of haze formation.

		Predicted	
		Bad (Haze)	Good (Normal)
Actual	Haze	True positive (A)	False negative (B)
	Normal	False positive (C)	True negative (D)

$$Hitrate = Recall = A/(A + B) \tag{11}$$

$$False\ alarm\ rate = 1 - Precision = C/(A + C) \tag{12}$$

$$Prediction\ loading\ rate = (A + C)\ /M \tag{13}$$

As shown in Figure 4, the validation process involved a training phase followed by a trial run. In the training phase, we first categorized the vendors of hazy photomasks and then computed the feature set for the haze transformation fraction. The feature set serves as a parameter set for the calculation of fuzzy similarity (FS) in the trial run phase. FS is computed by calculating the transformation fraction of the photomask being inspected with the feature set of the haze transformation fraction. If FS is less than 0.5, then the photomask is categorized as good, indicating that the model considers the photomask normal. If FS is greater than 0.5, then the photomask is categorized as bad, indicating that the model considers the photomask at risk of developing haze, such that an intervention will be required by the engineering department. A photomask confirmed as haze-affected is sent for cleaning and the data in the haze features database is updated. In cases where a photomask suspected of being bad is found to be normal, the photomask is shifted to the self-monitoring stage, as shown in Figure 5. Note that the prediction loading rate is used only in the trial run phase.

III. DATA SOURCE AND PREPARING

Photomask data were collected from a foundry in Taiwan producing 12-inch wafers [51]. The data included monitoring variables, such as indirect environmental factors, exposure energy-induced factors, photomask attributes, and cleaning processes, (see Table 4). The indirect environmental factors included cumulative storage time in the five environments. Exposure energy-induced factors included exposure energy and cumulative exposure count. Photomask attributes included photomask types, pellicle types, and vendors. Due to a lack of data related to cleaning, we used the details

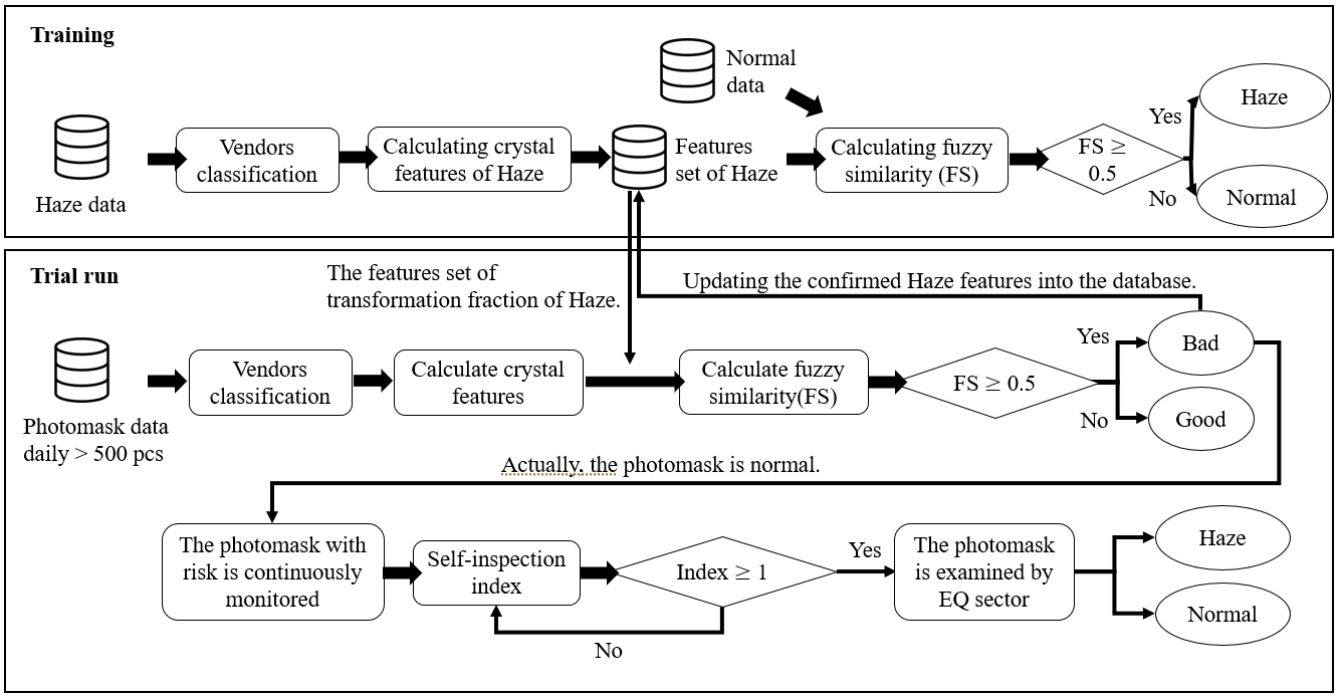

FIGURE 4. Decision flowchart of PdM for photomask haze.

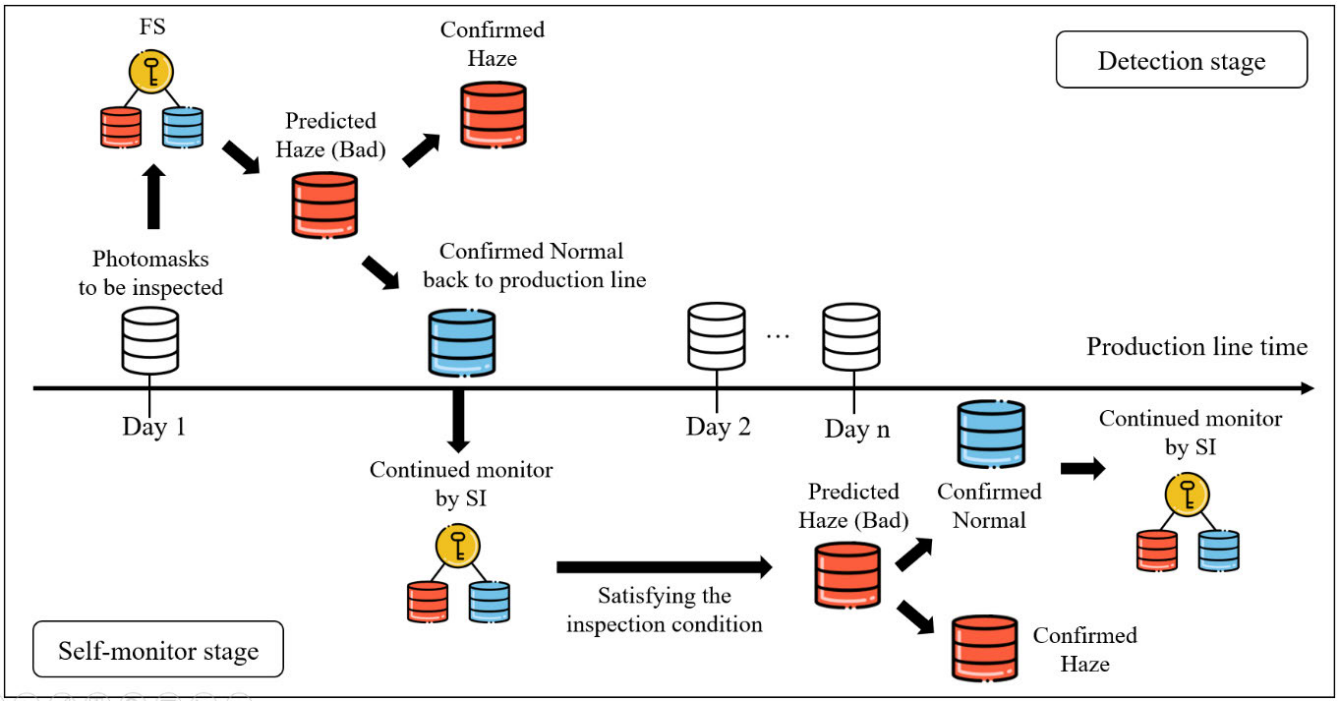

FIGURE 5. Two-stage flowchart of PdM during trial run daily.

pertaining to the cleaning methods and clean environment provided by the respective vendors.

As shown in Table 5, the sample comprised a total of 54,146 photomask samples (both the training and trial run phases), which included only 52 instances of haze (a severe imbalance in the data). A total of 731 photomasks were included in the training phase (28 instances of haze). A total of 53,415 photomasks were included in the trial run phase

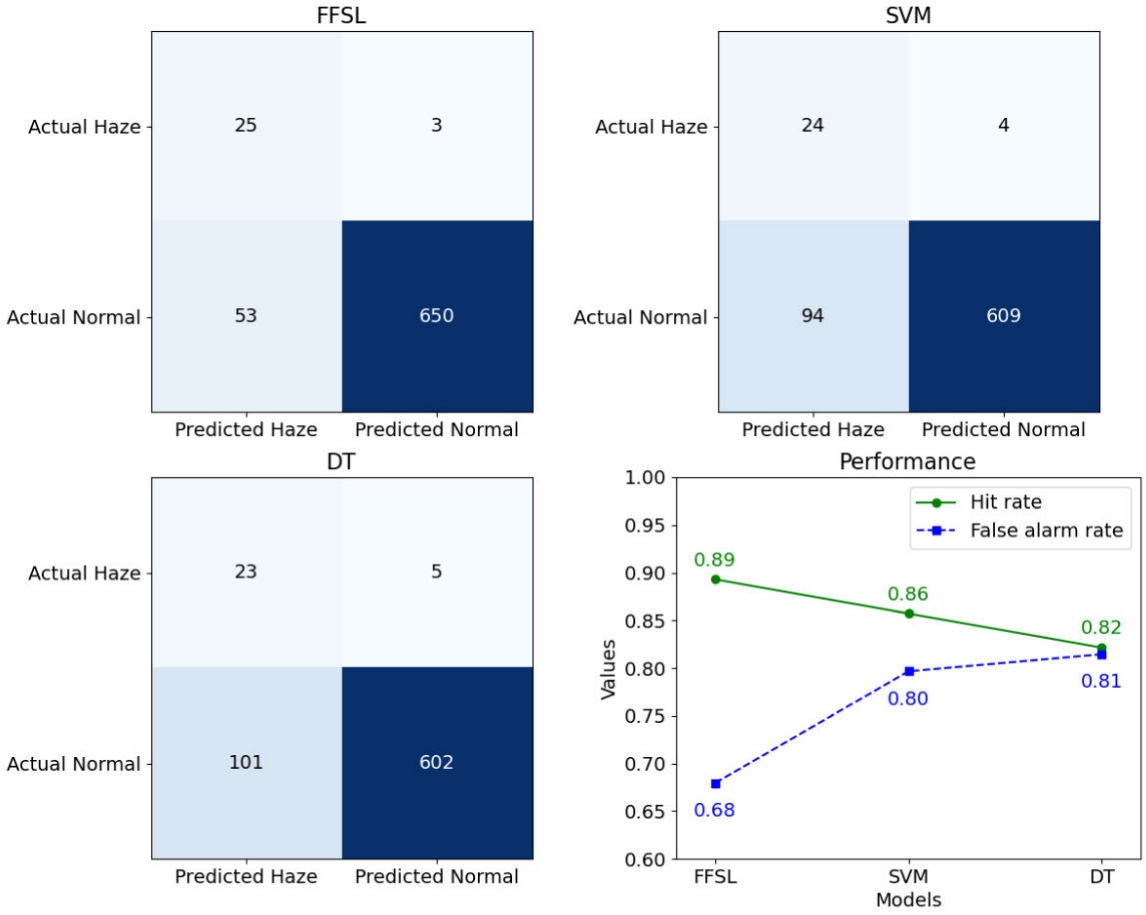

FIGURE 6. Confusion matrix and performance in the training.

TABLE 4. Summary of raw data variables.

Category	Variables	Type	Unit	Contents
Indirect environmental factors	CDA NCDA SINV INV FAB	Cumulative time	hour	Details see Table 1.
Exposure energy-induced	Exposure energy	Value	J/cm²	-
	Cumulative exposure count		count	-
Photomask attribute	Photomask types	Category	-	PSM
	Pellicle types		nm	193nm
	Vendors		-	five vendors
Cleaning process	Cleaning method	-	-	All are related to the vendor of the photomask
	Clean environment	-	-	

TABLE 5. Summary of sample sizes used in training and the trial run.

Phase	Period	Haze	Normal	Total
Training	-	28	703	731
Trial run	May	9	17,947	17,956
	June	6	16,321	16,327
	July	9	19,123	19,132
	Total	24	53,391	53,415
Sum		52	54,094	54,146

(25 instances of haze). In the model comparison, SVM and DT both used standardized raw data as independent variables for haze prediction. Additionally, during training, both SVM and DT adjust for imbalanced data by using the inverse of the sample proportions as weights to correct for unbalanced classes, thereby preventing the model from merely correctly distinguishing normal photomasks.

IV. RESULTS

The results obtained using training data are shown in Figure 6. The hit rates were as follows: FFSL (0.89), SVM (0.86), and DT (0.82), indicating that the proposed model exhibits higher accuracy in discriminating haze photomasks. The false alarm rates were as follows: FFSL (0.68), SVM (0.80), and DT (0.81), indicating that the proposed model has a lower rate of misclassifying normal photomasks as haze photomasks,

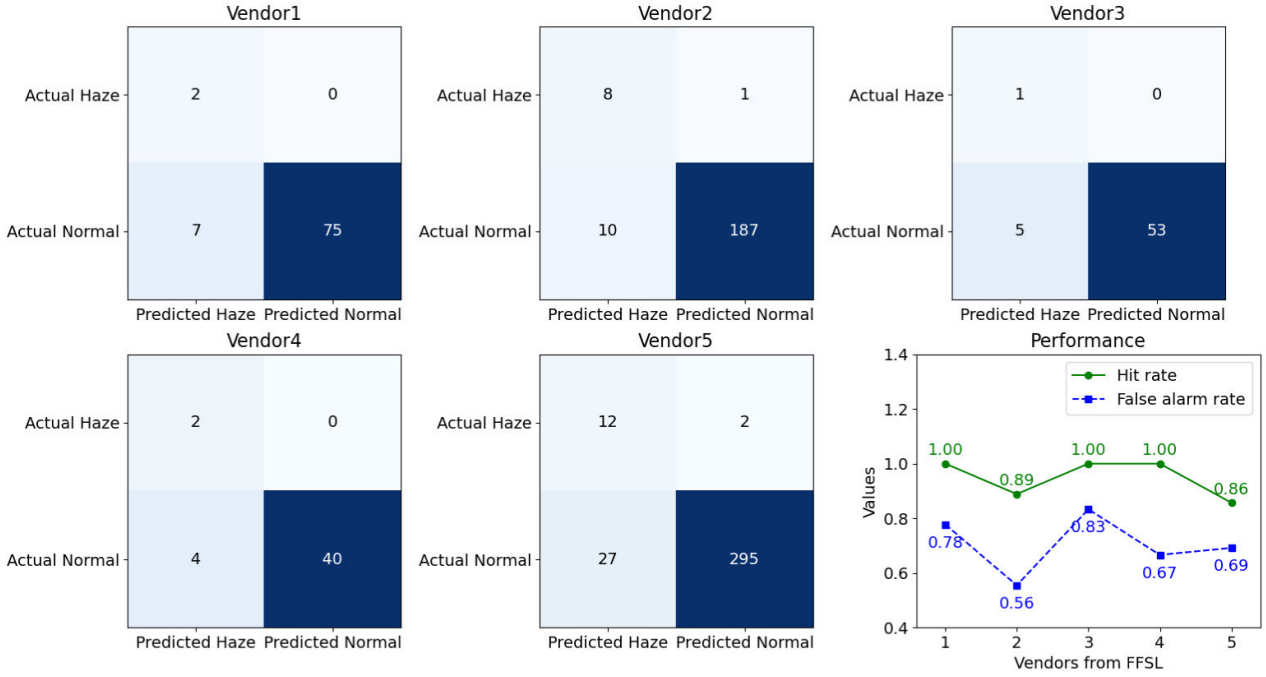

FIGURE 7. Confusion matrix and performance of vendors from FFSL.

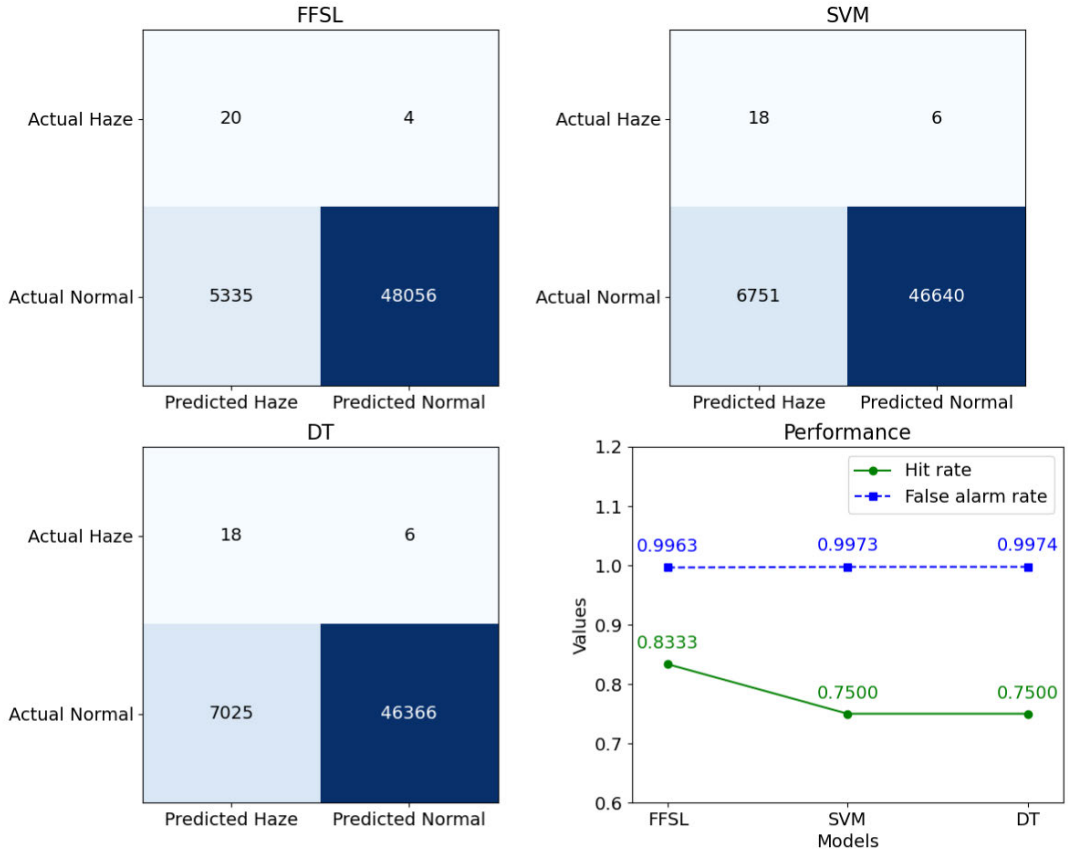

FIGURE 8. Confusion matrix and performance in the trial run.

thus demonstrating higher accuracy in discriminating normal photomasks. All of the scores were above 0.5; however, the use of non-linear transformations (e.g., photomask attribute classification and feature extraction) in a multivariate feature

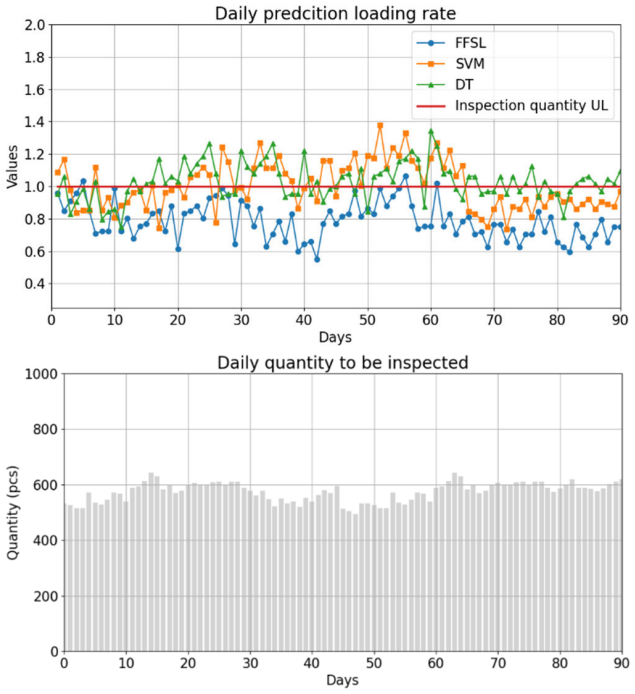

FIGURE 9. Daily prediction loading rate and quantity of photomasks requiring inspection.

space enhanced the ability of the proposed model to discriminate between normal and haze photomasks, thereby reducing the number of false alarms (n = 53). The proposed model significantly outperformed SVM (n = 94) and DT (n = 101).

As shown in Figure 7, further analysis of the differences among vendors revealed a hit rate between 0.86 and 1.0. Note that Vendor5 presented the highest number of hazy photomasks (haze = 12) and the lowest hit rate (0.86), making it the most representative. The false alarm rates ranged from 0.56 to 0.83. Vendor5 was the only vendor with a statistically significant large sample (predicted haze = 39 > 30), so its corresponding false alarm rate (0.69) is the most representative.

The trial run results in Figure 8 revealed that the hit rate of FFSL (0.83) was higher than that of SVM (0.75) and DT (0.75), indicating stable discriminative capability for hazy photomasks. Between the training and trial run phases, the hit rate decreased by only 0.06. All of the false alarm rates were close to 1: FFSL (0.9963), SVM (0.9973) and DT (0.9974). Due to a notable imbalanced data, the number of normal photomasks misclassified as haze far exceeded the number of photomasks correctly identified as haze (i.e., FFSL: 5,335 > 20), thereby rendering this indicator ineffective. Thus, we employed the prediction loading rate to measure the number of normal photomasks classified by the model as haze. As shown in Figure 9, the average daily prediction loading rate for FFSL was 0.79, which means that the FFSL model needs to inspect only 60 photomasks, based on inspection upper limit (UL) of 75 photomasks daily. This represents a 21% reduction in the costs for inspection. Note that in this batch of 60 inspections, the hit rate was 86%. By contrast, the average prediction loading rate for SVM and DT was higher than 1, which exceeded the daily inspection capacity. Taken together, FFSL clearly outperformed the other methods in terms of hit rate and prediction loading rate in the trial run (see Table 6).

TABLE 6. Summary of model performance in the trial run.

Model	Hit rate	False alarm rate	Prediction loading rate
FFSL	0.8333	0.9963	0.7882
SVM	0.7500	0.9973	1.0024
DT	0.7500	0.9974	1.0332

Note: The existing methods utilize SVM.

Cost-benefit analysis revealed a daily labor cost savings of 21%. This is equivalent to reducing the workforce by one, which corresponds to labor cost savings of roughly USD 18,780 annually, based on an average monthly salary of USD 1,565 for full-time workers in Taiwan in 2022 [52]. Decreasing the photomask inspection volume would also allow the deactivation of one photomask inspection device, which would further reduce costs related to equipment operation and electricity consumption.

V. CONCLUSION

The deployment of PdM is often hobbled by data heterogeneity, a reliance on indirect data, and/or a paucity of abnormal data, resulting in suboptimal performance in distinguishing anomalies or estimating the remaining useful life. Research on the implementation of PdMs under these constraints has been limited, as has research on the use of these methods for the prediction of photomask haze. The PdM framework presented in the current study employs cause-and-effect analysis to elucidate relationships, attribute classification and feature extraction for the processing of categorical and numerical variables, and fuzzy similarity and self-monitoring indicators to differentiate haze from normal photomasks. In a three-month validation run in a wafer foundry, the proposed FFSL framework was evaluated in terms of performance, robustness, and cost-benefit performance.

The proposed model outperformed SVM and DT in terms of accuracy in correctly identifying haze photomasks, while significantly reducing the number of false alarms. The ability to discriminate between haze and normal photomasks can be attributed to the nonlinear transformation of critical features via attribute classification and feature extraction. The proposed model also reduces the number of manual inspections to below daily inspection caps. Note that when using SVM or DT, the number of required inspections exceeded the capacity of the existing manpower. Cost-benefit analysis revealed that this model could reduce the inspection overhead by 21%, which is equivalent to a USD 18,780 reduction in labor costs annually. It would also allow the decommissioning of one photomask inspection device.

The proposed PdM framework and verification process could be used internally, i.e., for process management within the factory as well as the continual collection of abnormal data to mitigate the data imbalance. The framework could also be used externally; i.e., to assess the performance of vendors in terms of product quality based on metrics, such as the average useful life. As the abnormal data accumulates, when the abnormal data for each vendor reaches a statistically significant sample, it can be used to evaluate differences in vendor manufacturing quality, broadening the application of PdM and aligning it with practical factory management. In terms of model limitations, if there are too many categories of components, it may result in insufficient samples for each category, which could lead to inadequate representation of abnormal features and ultimately degrade the performance of haze photomask detection.

REFERENCES

[1] N. Vafaei, R. A. Ribeiro, and L. M. Camarinha-Matos, "Fuzzy early warning systems for condition based maintenance," *Comput. Ind. Eng.*, vol. 128, pp. 736–746, Feb. 2019.

[2] C. N. Madu, "Competing through maintenance strategies," *Int. J. Quality Rel. Manage.*, vol. 17, no. 9, pp. 937–949, Dec. 2000.

[3] T. P. Carvalho, F. A. A. M. N. Soares, R. Vita, R. D. P. Francisco, J. P. Basto, and S. G. S. Alcalá, "A systematic literature review of machine learning methods applied to predictive maintenance," *Comput. Ind. Eng.*, vol. 137, Nov. 2019, Art. no. 106024.

[4] E. Florian, F. Sgarbossa, and I. Zennaro, "Machine learning-based predictive maintenance: A cost-oriented model for implementation," *Int. J. Prod. Econ.*, vol. 236, Jun. 2021, Art. no. 108114.

[5] X. Han, Z. Wang, M. Xie, Y. He, Y. Li, and W. Wang, "Remaining useful life prediction and predictive maintenance strategies for multi-state manufacturing systems considering functional dependence," *Rel. Eng. Syst. Saf.*, vol. 210, Jun. 2021, Art. no. 107560.

[6] S. Van der Auweraer and R. Boute, "Forecasting spare part demand using service maintenance information," *Int. J. Prod. Econ.*, vol. 213, pp. 138–149, Jul. 2019.

[7] H. Mirshekali, A. Q. Santos, and H. R. Shaker, "A survey of time-series prediction for digitally enabled maintenance of electrical grids," *Energies*, vol. 16, no. 17, p. 6332, Aug. 2023.

[8] M. Fernandes, A. Canito, V. Bolón-Canedo, L. Conceição, I. Praça, and G. Marreiros, "Data analysis and feature selection for predictive maintenance: A case-study in the metallurgic industry," *Int. J. Inf. Manage.*, vol. 46, pp. 252–262, Jun. 2019.

[9] B. Taşcı, A. Omar, and S. Ayvaz, "Remaining useful lifetime prediction for predictive maintenance in manufacturing," *Comput. Ind. Eng.*, vol. 184, Oct. 2023, Art. no. 109566.

[10] D. Natanael and H. Sutanto, "Machine learning application using cost-effective components for predictive maintenance in industry: A tube filling machine case study," *J. Manuf. Mater. Process.*, vol. 6, no. 5, p. 108, Sep. 2022.

[11] T. Zonta, C. A. da Costa, F. A. Zeiser, G. de Oliveira Ramos, R. Kunst, and R. da Rosa Righi, "A predictive maintenance model for optimizing production schedule using deep neural networks," *J. Manuf. Syst.*, vol. 62, pp. 450–462, Jan. 2022.

[12] Y. Lv, X. Guo, Q. Zhou, L. Qian, and J. Liu, "Predictive maintenance decision-making for variable faults with non-equivalent costs of fault severities," *Adv. Eng. Informat.*, vol. 56, Apr. 2023, Art. no. 102011.

[13] A. Adsule, M. Kulkarni, and A. Tewari, "Reinforcement learning for optimal policy learning in condition-based maintenance," *IET Collaborative Intell. Manuf.*, vol. 2, no. 4, pp. 182–188, Dec. 2020.

[14] W. Song, H. Liu, and E. Zio, "Long-range dependence and heavy tail characteristics for remaining useful life prediction in rolling bearing degradation," *Appl. Math. Model.*, vol. 102, pp. 268–284, Feb. 2022.

[15] M. Zhu and X. Zhou, "Hierarchical-clustering-based joint optimization of spare part provision and maintenance scheduling for serial–parallel multi-station manufacturing systems," *Int. J. Prod. Econ.*, vol. 264, Oct. 2023, Art. no. 108971.

[16] T. Vanderschueren, R. Boute, T. Verdonck, B. Baesens, and W. Verbeke, "Optimizing the preventive maintenance frequency with causal machine learning," *Int. J. Prod. Econ.*, vol. 258, Apr. 2023, Art. no. 108798.

[17] W. J. Lee, H. Wu, H. Yun, H. Kim, M. B. G. Jun, and J. W. Sutherland, "Predictive maintenance of machine tool systems using artificial intelligence techniques applied to machine condition data," *Proc. CIRP*, vol. 80, pp. 506–511, Jan. 2019.

[18] Y. E. Mourabit, Y. El, H. Zougagh, and Y. Wadiai, "Predictive system of semiconductor failures based on machine learning approach," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 12, pp. 199–203, 2020.

[19] S. Ayvaz and K. Alpay, "Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time," *Expert Syst. Appl.*, vol. 173, Jul. 2021, Art. no. 114598.

[20] A. Nasser and H. Al-Khazraji, "A hybrid of convolutional neural network and long short-term memory network approach to predictive maintenance," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 1, p. 721, Feb. 2022.

[21] T. Zonta, C. A. da Costa, R. da Rosa Righi, M. J. de Lima, E. S. da Trindade, and G. P. Li, "Predictive maintenance in the industry 4.0: A systematic literature review," *Comput. Ind. Eng.*, vol. 150, Dec. 2020, Art. no. 106889.

[22] F. Sgarbossa, I. Zennaro, E. Florian, and M. Calzavara, "Age replacement policy in the case of no data: The effect of Weibull parameter estimation," *Int. J. Prod. Res.*, vol. 58, no. 19, pp. 5851–5869, Oct. 2020.

[23] J. Liu, K. Xu, B. Cai, and Z. Guo, "Fault prediction of on-board train control equipment using a CGAN-enhanced XGBoost method with unbalanced samples," *Machines*, vol. 11, no. 1, p. 114, Jan. 2023.

[24] Y. Tian, Y. Wang, X. Peng, and W. Zhang, "A fault diagnosis method for few-shot industrial processes based on semantic segmentation and hybrid domain transfer learning," *Appl. Intell.*, vol. 53, no. 23, pp. 28268–28290, Dec. 2023.

[25] S. Cavalieri and M. G. Salafia, "A model for predictive maintenance based on asset administration shell," *Sensors*, vol. 20, no. 21, p. 6028, Oct. 2020.

[26] W. Zhang, D. Yang, and H. Wang, "Data-driven methods for predictive maintenance of industrial equipment: A survey," *IEEE Syst. J.*, vol. 13, no. 3, pp. 2213–2227, Sep. 2019.

[27] A. Wang, Y. Zhang, H. Wu, K. Jiang, and M. Wang, "Few-shot learning based balanced distribution adaptation for heterogeneous defect prediction," *IEEE Access*, vol. 8, pp. 32989–33001, 2020.

[28] X. Liu, H. Guo, and Y. Liu, "One-shot fault diagnosis of wind turbines based on meta-analogical momentum contrast learning," *Energies*, vol. 15, no. 9, p. 3133, Apr. 2022.

[29] W. Zhang, P. Zhang, X. He, and D. Zhang, "Convolutional neural network based two-layer transfer learning for bearing fault diagnosis," *IEEE Access*, vol. 10, pp. 109779–109794, 2022.

[30] S.-C. Hsu and C.-F. Chien, "Hybrid data mining approach for pattern extraction from wafer bin map to improve yield in semiconductor manufacturing," *Int. J. Prod. Econ.*, vol. 107, no. 1, pp. 88–103, May 2007.

[31] K. Bhattacharyya, W. Volk, B. Grenon, D. Brown, and J. Ayala, "Investigation of reticle defect formation at DUV lithography," in *Proc. IEEE Adv. Semiconductor Manuf. Conf. WorkshopI/SEMI*, Jul. 2002, pp. 478–487.

[32] B. J. Grenon, C. R. Peters, K. Bhattacharyya, and W. W. Volk, "Formation and detection of subpellicle defects by exposure to DUV system illumination," *Proc. SPIE*, vol. 3873, pp. 162–176, Dec. 1999.

[33] B. Wu and A. Kumar, "Theoretical study of mask haze formation," in *Proc. 24th Eur. Mask Lithography Conf.*, Jan. 2008, pp. 1–11.

[34] S. J. Han, B. H. Kim, J. H. Park, Y. H. Kim, S. W. Choi, and W. S. Han, "The study on causes and control methods of haze contamination," *Adv. Microlithography Technol.*, vol. 5645, pp. 109–113, Jun. 2005.

[35] H. B. Kang, J. M. Kim, Y. D. Kim, H. J. Cho., and S. S. Choi, "A study for the control of chemical residuals on photomasks by using a thermal treatment for 65-nm node," *Proc. SPIE*, vol. 5853, pp. 501–506, May 2005.

[36] S. Shimada, N. Kanda, N. Takahashi, H. Nakajima, H. Tanaka, H. Ishii, and N. Hayashi, "A new model of haze generation and storage-life-time estimation for mask," *Photomask Technol.*, vol. 6349, pp. 422–428, Oct. 2006.

[37] H. Ishii, A. Tobita, Y. Shoji, H. Tanaka, A. Naito, and H. Miyashita, "Root cause analysis for crystal growth at ArF excimer laser lithography," *Proc. SPIE*, vol. 5446, pp. 218–224, Aug. 2004.

[38] J. Gordon, L. Frisa, C. Chovino, D. Chan, J. Keagy, and C. Weins, "Study of time dependent 193 nm reticle haze," *Proc. SPIE*, vol. 6730, pp. 469–480, Apr. 2007.

[39] J. Gordon, B. Murray, L. E. Frisa, E. Nelson, C. Weins, M. Green, and M. Lamantia, "Use of excimer laser test system for studying haze growth," *Proc. SPIE*, vol. 5992, pp. 1033–1043, Oct. 2005.

[40] J. M. Kim, J. C. Lee, D. S. Kang, D. H. Lee, C. Shin, M. H. Choi, and S. S. Choi, "Threshold residual ion concentration on photomask surface to prevent haze defects," *Proc. SPIE*, vol. 6607, pp. 556–563, May 2007.
[41] E. Johnstone, L. Dieu, C. Chovino, J. Reyes, D. Hong, P. Krishnan, D. Coburn, and C. Capella, "193-nm haze contamination: A close relationship between mask and its environment," *Proc. SPIE*, vol. 5256, pp. 440–448, Dec. 2003.
[42] M. Avrami, "Granulation, phase change, and microstructure kinetics of phase change. III," *J. Chem. Phys.*, vol. 9, no. 2, pp. 177–184, Feb. 1941.
[43] M. Avrami, "Kinetics of phase change. II transformation-time relations for random distribution of nuclei," *J. Chem. Phys.*, vol. 8, no. 2, pp. 212–224, Feb. 1940.
[44] M. Avrami, "Kinetics of phase change. I general theory," *J. Chem. Phys.*, vol. 7, no. 12, pp. 1103–1112, Dec. 1939.
[45] Y. P. Khanna and T. J. Taylor, "Comments and recommendations on the use of the avrami equation for physico-chemical kinetics," *Polym. Eng. Sci.*, vol. 28, no. 16, pp. 1042–1045, Aug. 1988.
[46] A. T. Lorenzo, M. L. Arnal, J. Albuerne, and A. J. Müller, "DSC isothermal polymer crystallization kinetics measurements and the use of the avrami equation to fit the data: Guidelines to avoid common problems," *Polym. Test.*, vol. 26, no. 2, pp. 222–231, Apr. 2007.
[47] V. Hinrichs, G. Kalinka, and G. Hinrichsen, "An avrami-based model for the description of the secondary crystallization of polymers," *J. Macromolecular Sci., B*, vol. 35, nos. 3–4, pp. 295–302, Jun. 1996.
[48] S.-M. Chen, M.-S. Yeh, and P.-Y. Hsiao, "A comparison of similarity measures of fuzzy values," *Fuzzy Sets Syst.*, vol. 72, no. 1, pp. 79–89, May 1995.
[49] C. P. Pappis and N. I. Karacapilidis, "A comparative assessment of measures of similarity of fuzzy values," *Fuzzy Sets Syst.*, vol. 56, no. 2, pp. 171–174, Jun. 1993.
[50] C. P. Pappis, "Value approximation of fuzzy systems variables," *Fuzzy Sets Syst.*, vol. 39, no. 1, pp. 111–115, Jan. 1991.
[51] *Daily Monitoring Data of Photomasks (from May 2022 to August 2022)*, Photomask Equip. Dept. PED, Wafer Foundry Factory, Taiwan, 2022.
[52] *Statistical Results of Industrial and Service Sector Salaries for December 2022 and Full Year 2022*, Directorate-General Budget, Accounting, Statist., Executive Yuan, Taiwan, 2022.

HAO-HSUAN HUANG received the M.S. degree from the Department of Resources Engineering, National Cheng Kung University, Tainan, Taiwan, in 2015. He is currently pursuing the Ph.D. degree in resources engineering with National Cheng Kung University. His research interests include fault diagnosis and machine learning.

YUN-HSUN HUANG received the Ph.D. degree from the Department of Resources Engineering, National Cheng Kung University, Tainan, Taiwan, in 2007. He is currently an Assistant Professor with the Department of Resources Engineering, National Cheng Kung University. His current research interests include energy modeling and machine learning.