Forecasting Rare Faults of Critical Components in LED Epitaxy Plants Using a Hybrid Grey Forecasting and Harmony Search Approach

Chun-Cheng Lin, *Member, IEEE*, Der-Jiunn Deng*, *Member, IEEE*, Jia-Rong Kang, Sen-Chia Chang, and Chuang-Hua Chueh

***Abstract*—In the LED manufacturing industry, the most expensive and crucial facilities are manufacturing machines. Condition-based maintenance (CBM) for crucial components of a manufacturing machine aims to forecast in advance the precise time when some aging component will be broken and replace it in time, to avoid performing abnormally to manufacture defect products. This study focuses on the CBM for a crucial component called particle filter of a pneumatic conveyor machine in the LED epitaxy plant. Conventional forecasting methods were based on theory of statistics, which requests a large number of data samples and assumes some probability distribution. With advance of machine technology, however, the data samples of broken particle filters to be collected are very few, such that those conventional methods cannot be applied. As a result, this study proposes a novel hybrid grey forecasting and harmony search approach, in which grey forecasting was shown to perform well for small data samples. In the proposed method, operating conditions of particle filters are monitored and collected by industrial sensors; then, those data is preprocessed by data filtering and clustering; finally, a hybrid grey forecasting and harmony search approach is used to fit the curve of the aging condition of a particle filter. Numerical analysis of a real example in an LED epitaxy plant shows that the proposed method performs better than conventional methods.**

***Index Terms*—Condition-based maintenance, industrial sensor, LED manufacturing, rare event detection, data mining, grey forecasting**

Manuscript received September 3, 2015; revised November 6, 2015; accepted for publication December 16, 2015.

The work was supported partially by Industrial Technology Research Institute, Taiwan, under the project of Big Data Technologies and Applications.

D.-J. Deng is the corresponding author of this paper.

C.-C. Lin and J.-R. Kang are with Department of Industrial Engineering and Management, National Chiao Tung University, Hsinchu 300, Taiwan. E-mails: cclin321@nctu.edu.tw, klkang8987@gmail.com.

D.-J. Deng is with Department of Computer Science and Information Engineering, National Changhua University of Education, Changhua 500, Taiwan. E-mail: djdeng@cc.ncue.edu.tw.

S.-C. Chang and C.-H. Chueh are with Computational Intelligence Technology Center, Industrial Technology Research Institute, Hsinchu 310, Taiwan. E-mails: chang@itri.org.tw, chchueh@itri.org.tw.

I. Introduction

Light emitting diode (LED) is a semiconductor electronic component that can emit visible and invisible light. As LED enjoys merits of small size, short response time, little power consumption, and long lifetime, it is applied in various products, e.g., computers, backlight of smart phone screens, and lighting equipment. The LED manufacturing industry mainly consists of three processes. The LED *epitaxy* plant is in charge of manufacturing the polished wafers and the epi wafers, where the latter has better performance than the former. Then, the LED *chip* plant bases on requirements of various LED components to produce chips by processes of electrodes and crystalline grains on the manufactured wafers. Finally, the LED *packaging* plant packages the manufactured chips, and then tests the packaged products.

In the LED manufacturing industry, manufacturing machines are the most expensive and crucial facilities. Hence, it has been challenging to keep the optimal operating condition of manufacturing machines to provide sufficient production capacity and manufacture high-quality products. Most previous related works focused on planning production schedules with various objectives or constraints, e.g., minimization of the total makespan of jobs [1], automatic design of scheduling policies [2], and small-scale multifunction robotic cell scheduling problems [3].

As customers have been more concerned about product quality, LED plants pay more attention to the influence of condition-based maintenance (CBM) on product quality. The CBM is a strategy which replaces in advance the crucial machine component that is going to degenerate or malfunction, so as to raise reliability of manufacturing operations and reduce manufacturing costs. In general, a manufacturing machine includes many complex components that are in charge of controlling physical quantities (e.g., pressure and temperature) or chemical amount (e.g., amount of raw material) of manufacturing processes. However, each machine component must age with time gradually, and finally is out of control (i.e., the set point differs from the real point), so that the manufacturing machine performs abnormally to manufacture defect products. To replace an aging component under proper maintenance cost and time so as to ensure their normal operation, a practical way is to attach an industrial sensor to the component to monitor its operating conditions, e.g., vibration frequency of a motor, air pressure, and temperature. Then, from those monitored conditions, the future health condition and remaining useful life (RUL) [4] of the component are analyzed to provide managers to make their maintenance plans.

In CBM, it has been challenging to correctly identify the time when some component starts to perform abnormally, so as to replace it in advance. Most related works were based on data mining, and the main idea behind those methods is to find patterns of an aging trend or model of the component from the historical data and then to analyze those patterns to forecast the component replacement time. And, most of those methods require a large number of effective data samples to be collected, and usually assume that those samples must satisfy some probability distribution, e.g., Weibull distribution [5]. However, with advance of technology, machine components in modern LED plants perform very accurately and are of a high quality, and hence, data samples of broken components are very few (less than 10 within 2 or 3 years, in general), so that conventional forecasting methods cannot be applied.

This study focuses on the CBM for a crucial component called particle filter of the pneumatic conveyor machine in the LED epitaxy plant. The data samples of broken particle filters to be collected currently are very few. Grey forecasting [6] has been shown to perform well for problems with small samples. Hence, to cope the CBM with such small samples, this study proposes a hybrid grey forecasting and harmony search approach. First, the operating conditions of machine components are monitored and collected by industrial sensors; then, the collected data is preprocessed by data filtering and clustering analysis, in which data filtering aims to decrease interference of data noise and enhance the influence of key data on forecasting; and clustering analysis aims to find the hidden association and difference among the data. Then, a hybrid grey forecasting and geometric selective harmony search approach is used to find the grey parameters and generate the optimal curve to fit the aging behavior of particle filters.

II. Preliminaries

A. The CBM of particle filters of a pneumatic conveyor machine in the LED epitaxy plant

The LED epitaxy plant includes three processes: crystal growth, polishing, and epitaxy (Fig. 1). In the crystal growth process, raw material is melt down at a high temperature, and then becomes ingots after a sequence of crystal growth processes, including neck growth, crown growth, body growth, and tail growth. Then, in the polishing process, those ingots are processed by cropping, slicing, rough grinding, polishing, lapping, and cleaning, to become polished wafers. Finally, in the epitaxy process, the polished wafers are processed by chemical vapor deposition methods to become epi wafers of various sizes and thickness.

This study is concerned with the CBM of particle filters of a pneumatic conveyor machine in the LED epitaxy process, as shown in Fig. 2. In manufacturing epi wafers, a large amount of chemical air is required. The chemical air is stored in a chamber on the left side in Fig. 2, and is drew out by a pump on the right side in Fig. 2. Since the chemical air contains uncountable particles that are harmful to the pump, the air drew out from the chamber to the pump is required to pass through a particle filter (see Fig. 2), to avoid those particles from entering the pump to increase the risk of damage to the pump.

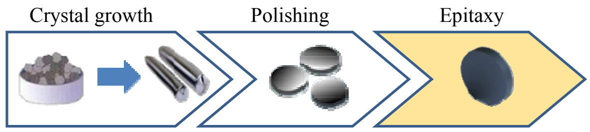


Fig. 1. The LED epitaxy process.

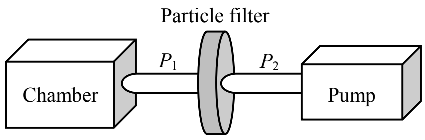


Fig. 2. Illustration of a pneumatic conveyor machine.

The health condition of a pneumatic conveyor machine can be realized by measuring the pressure difference of the two sides of the particle filter. That is, if P_1 and P_2 denote pressures of the chamber and the pump, respectively, then the pressure difference is $|P_1 - P_2|$ (Fig. 2). In practice, the pressure difference is monitored continuously by an industrial sensor; and its value changes irregularly with time, because the particle filter ages with time so that it is stuck with more and more particles as time goes by. A real instance of the data sensed by the industrial sensor is shown in Fig. 3. Fig. 3(a) gives the plots of all pressure difference values of a complete life cycle of a particle filter, which went through 16 epitaxy processes of epi wafers of different quantity and different sizes, because the products ordered by customers are usually of small quantity and of large diversity in the modern LED industry. Note that the data of the 16 processes is framed and labelled by red serial numbers in Fig. 3(a), in which the time length of each process differs. In practice, based on previous experiences, if the pressure difference exceeds a threshold (i.e., the dotted horizontal line in Fig. 3(a)), the pneumatic conveyor machine would perform abnormally to manufacture defect epi wafers.

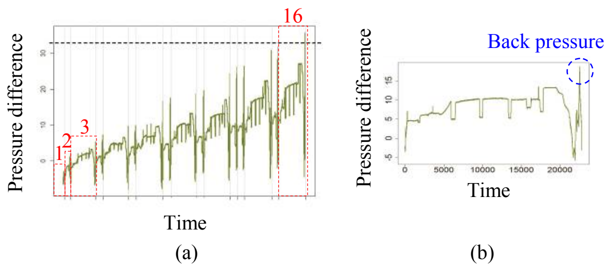


Fig. 3. The raw sensor data. (a) Plots of all pressure differences of a complete life cycle of a particle filter, consisting of 16 processes of epi wafers. (b) Detailed plots of pressure differences of the third process in the life cycle.

Another phenomenon in Fig. 3 should be noticed. When a process is completed (i.e., the process of one lot of same-size epi wafers is completed, also called a *run* throughout the rest of this paper), the pump stops working such that some air may flow back from the pump to the chamber. This phenomenon is called a back pressure, as shown in the dotted cycle in Fig. 3(b), in which the pressure difference originally falls down but suddenly leads to a peak value. Back pressure is a normal physical phenomenon, but its sudden peak value may exceed the threshold. This phenomenon becomes serious as the particle filter ages with time, i.e., the back pressure value becomes larger as time goes by as shown in Fig. 3(a). Especially, the last (i.e., 16th) back pressure exceeds the threshold; then, an alarm signal will be sent. Once the maintenance staff receive the alarm, they will replace the particle filter immediately.

B. Grey forecasting model

Grey forecasting [6] is based on concept of mechanical systems to construct a forecasting model to analyze patterns of changes of system states. It enjoys the advantage of requiring only small samples (generally, the required sample number is 4 to 10); while the conventional methods (e.g., time series and regression) require a large number of samples. It is based on the concept of mechanical systems to construct a grey model denoted by GM(i, N) to analyze patterns of changes of system states, in which an i-th order differential equation with N variables is used to fit the forecasting model. Grey forecasting has lots of applications, e.g., forecasting fashion retailing [7], forecasting wind power [8], engineering prediction [9], and forecasting values of agricultural imports [10].

This study will apply the GM(1, N) model, i.e., using the 1st order differential equation with N variables, detailed as follows. Consider the regression problem with one dependent variable x_0 and (N – 1) independent variables x_1, x_2, …, x_{N-1}. Suppose that n samples of those variables are observed. Let $X_j^{(0)}$ denote the time series for variable x_j, and $X_j^{(0)}(t)$ denote its t-th entry. Represent all the sample as a series $X_j^{(0)} = \left\{X_j^{(0)}(1), X_j^{(0)}(2), ..., X_j^{(0)}(n)\right\}$, for j = 0, 1, …, N – 1.

Generate a first-order accumulated generating operation (AGO) series $X_j^{(1)}$ for each $X_j^{(0)}$ as follows: $X_j^{(1)}$ = { $X_j^{(1)}(1)$, $X_j^{(1)}(2)$, …, $X_j^{(1)}(n)$ } = { $X_j^{(0)}(1)$, $\sum_{k=1}^{2} X_j^{(0)}(k)$, …, $\sum_{k=1}^{n} X_j^{(0)}(k)$ }. Then, the grey model is expressed as the following first-order grey differential equation:

$$\frac{\mathrm{d}X_0^{(1)}(t)}{\mathrm{d}t} + b_0 X_0^{(1)}(t) = b_1 X_1^{(1)}(t) + b_2 X_2^{(1)}(t) + ... + b_{N-1} X_{N-1}^{(1)}(t) \tag{1}$$

where b_0 is the grey developmental coefficient; and b_1, b_2, …, b_{N-1} are associated coefficients corresponding to their respective associated series. By approximating $\mathrm{d}X_0^{(1)}(t)/\mathrm{d}t$ as $X_0^{(0)}(t+1)$, $X_0^{(1)}(t)$ as $(X_0^{(1)}(t) + X_0^{(1)}(t+1))/2$, and each $X_j^{(1)}(t), j = 1, 2, ..., N-1$ as $X_j^{(1)}(t+1)$, the least-square solution to parameters of the above grey model is as follows:

$$[b_0,\ b_1,\ ...,\ b_{N-1}]^{\mathrm{T}} = (B^{\mathrm{T}} B)^{-1} B^{\mathrm{T}} Y \tag{2}$$

where

$$B = \begin{bmatrix} -(X_0^{(1)}(1) + X_0^{(1)}(2))/2 & X_1^{(0)}(2) & ... & X_{N-1}^{(0)}(2) \\ -(X_0^{(1)}(2) + X_0^{(1)}(3))/2 & X_1^{(0)}(3) & ... & X_{N-1}^{(0)}(3) \\ \vdots & \vdots & \vdots & \vdots \\ -(X_0^{(1)}(n-1) + X_0^{(1)}(n))/2 & X_1^{(0)}(n) & ... & X_{N-1}^{(0)}(n) \end{bmatrix}; \tag{3}$$

$$Y = [X_0^{(0)}(2),\ X_0^{(0)}(3),\ ..., X_0^{(0)}(n)]^{\mathrm{T}}. \tag{4}$$

Hence, the forecasting series is computed as follows:

$$\begin{aligned} \hat{X}_0^{(1)}(t) &= \left(X_0^{(0)}(1) - \frac{1}{b_0} \sum_{i=1}^{N-1} b_i \cdot X_i^{(1)}(t) \right) \cdot e^{-a(t-1)} \\ &\quad + \frac{1}{b_0} \sum_{i=1}^{N-1} b_i \cdot X_i^{(1)}(t). \end{aligned} \tag{5}$$

Then, based on the following first-order inverse accumulated generating operation (IAGO) on $\hat{X}_0^{(1)}$, the forecasting value for the dependent series $\hat{X}_0^{(0)}$ can be obtained.

$$\hat{X}_0^{(0)}(1) = \hat{X}_0^{(1)}(1) = X_0^{(0)}(1); \tag{6}$$

$$\hat{X}_0^{(0)}(t+1) = \hat{X}_0^{(1)}(t+1) - \hat{X}_0^{(1)}(t) \text{ for } t = 1, 2, ... \tag{7}$$

C. Related works on CBM

Previous works on applying CBM of components to forecasting broken components can be classified into two categories: model-based method and data-driven method. Model-based method is based on data samples to establish a mathematical model to find the aging trend or RUL of machine components for forecasting. The work in [11] modeled the deterioration of complex industrial assets as a multiple dependent deterioration path process to provide their CBM policies. The work in [12] developed an improved rolling grey forecasting method to conduct the forecasting of CBM.

Data-driven method [13] is based on the method of data training to analyze the trend or model of the concerned objects. The data-driven methods based on artificial neural networks (ANN) are the most popular. The work in [14] incorporated ANN and fuzzy theory to forecast the health conditions of machines. The works in [15] applied statistical-time features and ANN to detect bearing faults. The work in [16] based on ANN, and incorporated regression analysis and least square regression for condition monitoring and fault diagnosis of sensors in power plants. The work in [17] established a dynamic parsimonious fuzzy neural network method, and showed that their method performed better than other methods on various benchmarks.

The work in [18] designed a nonlinear multivariate statistical process monitoring system for dryers of paper machines, to detect moisture of products to judge whether the machine is broken. The work in [19] suggested to construct a proportional hazards model for the CBM problem, in which the health condition index of the machine is represented by a transition probability matrix. Since it is not easy to construct a transition probability matrix for the health condition of machines, Markov models can be incorporated, e.g., the works in [20] and [21] respectively constructed the mixture of Gaussians hidden Markov models and the hidden Markov model to forecast the lifetime of motor bearings. To enhance the function of monitoring machine conditions in industrial wireless sensor networks, the work in [22] used computationally intensive classifiers in computationally weak sensor network nodes.

III. Problem Description

In an LED epitaxy plant, the particle filter of a pneumatic conveyor machine ages with time, so that the pressure difference between the chamber and the pump (Fig. 2) has an increasing back pressure phenomenon (Fig. 3(b)). Once the pressure difference exceeds the threshold, the machine starts to perform abnormally to manufacture defect products. This study focuses on CBM of a particle filter, to forecast the precise time of replacing the particle filter in advance to avoid its malfunction and reduce maintenance costs. More specifically, reminding that a life cycle of a particle filter consists of multiple processes, an alarm signal will be sent to maintenance mangers when the pressure difference is forecasted to exceed the threshold at the next process. Once a manager receives the alarm, he/she replaces the particle filter after the current process is completed. By doing so, the times of unexpected shutdown so as to affect the whole scheduling could be reduced, and so could the risk of increasing production cost due to manufacturing defect products.

Our forecasting problem can be regarded as a regression problem: $x_0 = f(x_1, x_2)$, where each variable has a state when a process of the machine is completed; the dependent variable x_0 is a binary variable: $x_0 = 1$ if the particle filter is replaced before this process; otherwise, $x_0 = -1$. Industrial sensors are used to be collected the values of many factors (e.g., temperature, moisture, and type of the chemical air to be used) that could affect the pressure difference. Then, the principal components analysis (PCA) is conducted to determine two PCA variables, denoted by x_1 and x_2 in the above regression equation.

If a false alarm is sent, a normal particle filter is still replaced so that the total cost is increased. Hence, the objective of the concerned problem is to minimize the cost of penalizing false alarms, which is expressed as follows:

$$\begin{aligned}&\text{Minimize } c(\hat{X}_0^{(0)})\\&=\sum_{t=1}^{n}\left|X_0^{(0)}(t)-\hat{X}_0^{(0)}(t)\right|\cdot\left(0.5\,|\,\hat{X}_0^{(0)}(t)-1\,|+50\,|\,\hat{X}_0^{(0)}(t)+1\,|\right)\end{aligned} \tag{8}$$

where $X_0^{(0)} = (X_0^{(0)}(1), X_0^{(0)}(2), ..., X_0^{(0)}(n))$ is the real time series of dependent variable x_0 during the life cycle of a particle filter; $\hat{X}_0^{(0)} = (\hat{X}_0^{(0)}(0), \hat{X}_0^{(0)}(1), ..., \hat{X}_0^{(0)}(n))$ is the forecasted series; n is number of processes of the life cycle.

In Equation (8), if the forecasted value is correct (i.e., $X_0^{(0)}(t) = \hat{X}_0^{(0)}(t)$), then the cost for the t-th process is zero. Otherwise, consider two possible cases: 1) if $\hat{X}_0^{(0)}(t) = 1$ but $X_0^{(0)}(t) = -1$, then the cost is 200; 2) if $\hat{X}_0^{(0)}(t) = -1$ but $X_0^{(0)}(t) = 1$, then the cost is 2. That is, the penalty cost is larger when the filter with no need to be replaced is replaced.

IV. The Proposed Approach

This study proposes a hybrid grey forecasting and harmony search approach for CBM of a particle filter with small data samples. The proposed approach incorporates techniques of data mining and grey forecasting, and hence, enjoys both of their advantages, in which association rules or key hidden information could be found via data mining; and precise parameters of the grey forecasting model can be found via the proposed approach, so as to generate the optimal curve to fit the aging trend of a particle filter.

The flowchart of the proposed approach is given in Fig. 4. The input of the problem includes the real data of replacing particle filters or not (i.e., x_0 in the regression equation) and the two major PCA values called PCA 1 and PCA 2 (i.e., x_1 and x_2 in the regression equation) of each process in each particle filter life cycle as described in Section III. The proposed approach for the concerned regression problem is a supervised learning approach, and hence, 75% of particle filter cycles are used as the training data, while the others are used as the testing data.

Then, the training data is preprocessed by *data filtering* and *clustering analysis*. Grey forecasting is suitable for the data with small samples. However, the samples of replacing particle filters accounts for a relatively small ratio of the whole data samples. Hence, the *data filtering* cuts the samples of not replacing filters (i.e., those samples with $x_0 = 1$). On the other hand, since it is hard to find a general pattern for the aging behavior of a particle filter, the *clustering analysis* divides all particle filter life cycles into multiple clusters, each of which is regarded as an aging behavior pattern.

Then, a hybrid grey forecasting and harmony search approach is used to construct a grey forecasting model for each cluster of particle filter life cycles. In data testing, the proposed approach checks which cluster in the training data is the most similar to the concerned testing data. Specifically, the cluster in the training data with the shortest distance between its mean coordinate and the mean coordinate of the testing data is the so-called *similar cluster*. Then, the grey model of the similar cluster is used to forecast the trend of the testing data (time series), and hence, some precision measures can be computed for performance comparison.

The rest of this section gives details of data preprocessing and the hybrid grey forecasting and harmony search approach.

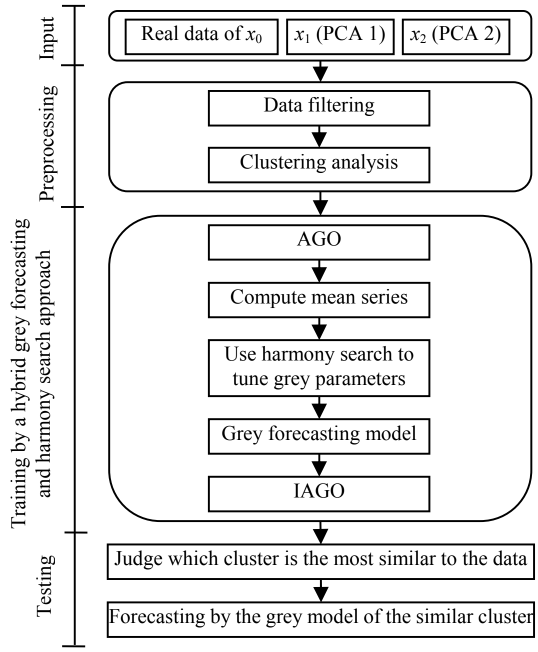


Fig. 4. Flowchart of the proposed approach.

A. Data preprocessing

Data is preprocessed by data filtering and clustering analysis.

1) Data filtering. Taking PCA 1 and PCA 2 values as a (x, y)-coordinate, plots of all data of a real instance are given in Fig. 5(a), in which each plot representing a process of replacing a particle filter is colored in red and is called a *key plot* in this paper; while the others are in blue. From Fig. 5(a), there is a serious imbalance between the two types of plots, which is harmful to forecasting. Since number of blue plots is much greater than that of red plots, we filter out a part of blue plots to reduce interference of data noises (e.g., missing values or measurement errors) and increase influences of key plots. In this instance, all the plots with PCA1 less than 0 are not key plots, and are removed to reduce the total data size (Fig. 5(b)).

2) Clustering analysis. We observe that different life cycles of particle filters have different behaviors, and hence, we conduct clustering analysis on the data after data filtering. This study considers two clustering methods: *k-means clustering* and *life cycle-based clustering*. Given a clustering number k, the *k-means clustering* determines k clusters according to the distances among plots. For example, the result after k-means clustering when $k = 3$ is shown in Fig. 5(c), in which plots of the same cluster are marked in the same color. On the other hand, *life cycle-based clustering* considers all plots (processes) of the same particle filter life cycle as the same cluster, i.e., number of particle filter life cycles is equal to that of clusters. For example, Fig. 5(d) shows the clustering result of 4 particle filter life cycles, in which distribution of each life cycle differs a lot.

B. A hybrid grey forecasting and harmony search approach

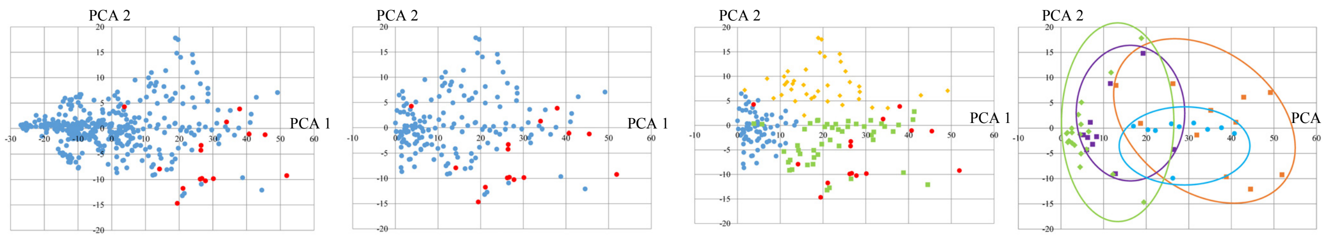
In the grey model as detailed in Subsection II-B, precision of the model is determined by two key parameter types, i.e., grey development parameter b_0 and grey associated parameters b_1, b_2, …, b_{N-1}. The grey development parameter b_0 mainly reflects the development trend of dynamic process of system states, which can judge speed of system development and some characteristics. The grey associated parameters represent additional effects external to the system. The greater an associated parameter, the greater some abstract external effect.

Recent works proposed various grey forecasting methods. Most of those works applied metaheuristic algorithms to find effective grey parameters, and showed outperformance of the grey model based on the found parameters. This study proposes a novel hybrid grey forecasting and harmony search approach, in which grey parameters are computed by the geometric selective harmony search (GSHS) [23], which has been shown to perform better than conventional metaheuristic algorithms, e.g., GA and PSO. Harmony search algorithm (HSA) is a metaheuristic algorithm inspired from improvisation of musicians. The GSHS is an improved HSA that integrates the crossover and mutation schemes in GA.

The algorithm of the proposed hybrid grey forecasting and geometric selective harmony search approach (Grey-GSHS) is given in Algorithm 1, consisting of the following main steps:

1) Parameter initialization. Line 1 of Algorithm 1 initializes the following parameters for GSHS: harmony memory size (*HMS*), harmony memory considering rate (*HMCR*), pitch adjusting probability (*PAR*), number of iterations (*NI*), tournament size (*ts*), mutation step (*ms*).

2) Construct harmony memory HM and evaluate costs. Note that the original grey forecasting model finds grey parameters by the least-square method (LSM) in Equation (4), which is an approximated solution. Instead of the LSM, this study applies the GSHS to find the optimal grey development parameter b_0 and two grey associated parameters b_1 and b_2 which are associated with PCA 1 and PCA 2, respectively. In the GSHS,



(a) Original data (b) After data filtering (c) After k-means clustering (d) Life cycle-based clustering

Fig. 5. (a) Original data. (b) After data filtering (cutting all plots with PCA 1 < 0). (c) After k-means clustering when k = 3. (d) Life cycle-based clustering (showing only 4 particle filter life cycles).

each harmony is encoded as (b_0, b_1, b_2). The GSHS works on a harmony memory (*HM*) matrix consisting of *HMS* harmonies and their costs as follows:

$$HM = \begin{bmatrix} b_0^1 & b_1^1 & b_2^1 & c(b_0^1, b_1^1, b_2^1) \\ b_0^2 & b_1^2 & b_2^2 & c(b_0^2, b_1^2, b_2^2) \\ \vdots & \vdots & \vdots & \vdots \\ b_0^{HMS} & b_1^{HMS} & b_2^{HMS} & c(b_0^{HMS}, b_1^{HMS}, b_2^{HMS}) \end{bmatrix}$$

where (b_0^i, b_1^i, b_2^i) is the *i*-th harmony and c(b_0^i, b_1^i, b_2^i) is its cost, for i = 1, 2, …, *HMS*. To compute cost of a harmony, Line 2 of Algorithm 1 first calculates AGO series. Then, Line 3 applies the three parameters in each harmony to Equations (5)–(7) to obtain the forecasted series $\hat{X}_0^{(0)}$, and then substitutes it into Equation (8) to obtain cost of the harmony.

Algorithm 1 Grey-GSHS

1: Initialize the parameters for GSHS
2: Calculate AGO series
3: Construct the harmony memory *HM* and evaluate costs
3: $gn = 1$
4: **while** $gn \le NI$ **do**
5: **for** $j = 0$ to 2 **do**
6: **if** $rand(0, 1) < HMCR$ **then**
7: Use tournament selection to select two harmonies b_j^{new1} and b_j^{new2} from *HM*
8: $b_j^{new} = \alpha \cdot b_j^{new1} + (1 - \alpha) \cdot b_j^{new2}$
9: **if** $rand(0, 1) < PAR$ **then**
10: $b_j^{new} = b_j^{new} \pm ms \cdot rand(0, 1) \cdot b_j^{new}$
11: **end if**
12: **else**
13: $b_j^{new} = min_j + rand(0, 1) \cdot (max_j - min_j)$
14: **end if**
15: **end for**
16: The new harmony ($b_0^{new}, b_1^{new}, b_2^{new}$) replaces the worst harmony if its cost c($b_0^{new}, b_1^{new}, b_2^{new}$) is lower
17: **end while**
18: Apply the three grey parameters in the best harmony ($b_0^{best}, b_1^{best}, b_2^{best}$) of the *HM* matrix to the grey model
19: Calculate the IAGO series

3) Generate new harmonies. Lines 4–17 of Algorithm 1 is the main loop of the GSHS which iteratively generates a new harmony to improve the *HM* matrix. Lines 5 – 15 generates each parameter of the new harmony. If a random real number from [0, 1] (i.e., *rand*(0, 1)) is less than *HMCR* (Line 6), then the parameter is generated based on *HM* (Lines 7–11); otherwise, it is generated randomly in Line 13, in which min_j and max_j are the minimal and maximal value in the *j*-th column of the *HM* matrix, respectively. When *rand*(0, 1) < *HMCR*, tournament selection is used twice to select out two harmonies b_j^{new1} and b_j^{new2} from *HM* in Line 7, and Line 8 generates b_j^{new} by linear combination of b_j^{new1} and b_j^{new2}. Furthermore, if a random real number from [0, 1] is less than parameter *PAR* (Line 9), parameter b_j^{new} is adjusted within a small range *ms* in Line 10.

4) Update the HM matrix. Line 16 of Algorithm finds the worst harmony in the *HM* matrix, and considers whether its cost is worse than cost of the new harmony ($b_0^{new}, b_1^{new}, b_2^{new}$). If true, replace the worse harmony by the new one.

5) Forecasting. Line 18 of Algorithm 1 applies the three grey parameters in the best harmony ($b_0^{best}, b_1^{best}, b_2^{best}$) of *HM* matrix to the grey model to obtain the forecasted AGO series $\hat{X}_0^{(1)}$ from Equation (5). Then, Line 19 applies the IAGO to generate the forecasted series $\hat{X}_0^{(0)}$ from Equations (6) and (7).

V. Implementation and Experimental Results

This section evaluates performance of the proposed approach to a real instance in an LED epitaxy plant. This instance has the data of 20 particle filters (i.e., 20 life cycles). Since the propose approach is based on supervised learning, 15 cycles (with 402 processes) are used as the training data, and the remaining 5 cycles (with 185 processes) are used as the testing data for performance evaluation.

We compare the experimental performance of the following four grey forecasting approaches:

- *Grey*: The conventional grey forecasting approach;
- *Grey_GSHS_1*: The Grey-GSHS approach without any data filtering and clustering analysis;
- *Grey_GSHS_2*: The Grey-GSHS approach using data filtering and *k*-means clustering;
- *Grey_GSHS_3*: The Grey-GSHS approach using data filtering and life cycle-based clustering.

The proposed approach is implemented in C++ programming language. All experiments run on a PC with an Intel i7-3770 CPU 3.40GHz and 16-GB RAM.

For evaluating forecasting performance, it is common to apply the following four measures of confusion matrix to judge if the process of replacing particle filters is forecasted correctly (i.e., a correct alarm is sent): True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP and TN are measures for correct number of alarms from the positive and negative aspects. Conversely, FP and FN are measures for incorrect number of alarms from positive and negative aspects. Furthermore, the above four measure constitute the following five common performance indices: true positive rate (TPR) = TP / (TP + FN); false positive rate (FPR) = FP / (FP + TN); accuracy = (TP + TN) / (TP + TN + FP + FN); precision = TP / (TP + FP); f-measure = 2 TP / (2TP + FP + FN).

In the rest of this section, the parameter setting for the

Grey-GSHS approach is discussed, and the cut value of PCA 1 used in data filtering and number of clusters *k* used in *k*-means clustering are analyzed in detail. Finally, the experimental results of four forecasting approaches are evaluated.

A. Parameter Setting

By referring to [23] and lots of experimental trials, some parameters for the proposed Grey-GSHS approach are given as follows, i.e., harmony memory size (*HMS*) = 100, harmony memory considering rate (*HMCR*) = 0.8, pitch adjusting probability (*PAR*) = 0.1, number of iterations (*NI*) = 15000, tournament size (*ts*) = 2, and mutation step (*ms*) = 0.5.

The cut value of PCA 1 used in data filtering (i.e., all the plots with PCA 1 value less than the cut value are removed) has great influence on the performance. Hence, we conduct the experimental analysis under 4 different cut values of PCA 1 in Table I. From Table I, the case when the cut value is set to 3 has the best forecasting performance (see precision = 8.16%, in which among 49 alarms, 4 alarms are correct, but 45 alarms are incorrect). Hence, the cut value is set to 3.

TABLE I
ANALYSIS ON THE CUT VALUE OF PCA 1 USED IN DATA FILTERING

	The cut values of PCA 1			
	0	1	2	3
TPR	80.00%	80.00%	80.00%	80.00%
FPR	49.44%	27.78%	26.11%	25.00%
accuracy	51.35%	72.43%	74.05%	75.14%
precision	4.30%	7.41%	7.84%	8.16%
f-measure	8.16%	13.56%	14.29%	14.81%
TP	4	4	4	4
TN	91	130	133	135
FN	1	1	1	1
FP	89	50	47	45

In *k*-means clustering, setting a different number of clusters affects the performance. To find an appropriate cluster number, we conduct experimental analysis using 4 different cluster numbers (i.e., *k* = 2, 3, 4, 5) in Table II. From Table II, the case when *k* = 4 has the best results (see *precision* = 17.39%, in which among 23 alarms, 4 alarms are correct, but 19 alarms are incorrect. Hence, the number of clusters *k* used in *k*-means clustering is set to 4.

TABLE II
ANALYSIS OF NUMBER OF CLUSTERS USED IN K-MEANS CLUSTERING

	k-means			
	k = 2	*k* = 3	*k* = 4	*k* = 5
TPR	80.00%	80.00%	80.00%	80.00%
FPR	18.33%	20.56%	10.56%	11.11%
accuracy	81.62%	79.46%	89.19%	88.65%
precision	10.81%	9.76%	17.39%	16.67%
f-measure	19.05%	17.39%	28.57%	27.59%
TP	4	4	4	4
TN	147	143	161	160
FN	1	1	1	1
FP	33	37	19	20

B. Experimental results and performance comparison

This section compares the experimental results of four grey forecasting approaches (i.e., Grey, Grey_GSHS_1, Grey_GSHS_2, and Grey_GSHS_3) in Table III. From Table III, the Grey_GSHS_3 has the best performance in all indices, followed by Grey_GSHS_2 and Grey_GSHS_1, and the Grey has the worst performance.

In comparison of experimental results between Grey and Grey_GSHS_1, we find that GSHS is helpful to compute the grey development coefficient and grey associated coefficients, so that its forecasting performance is better. Additionally, from the results for Grey_GSHS_1, Grey_GSHS_2, and Grey_GSHS_3, we find that data preprocessing can effectively increase the predictive ability of grey forecasting; and life cycle-based clustering performs better than *k*-means clustering.

TABLE III
EXPERIMENTAL RESULTS OF FOUR GREY FORECASTING APPROACHES

	Grey	Grey_GSHS_1	Grey_GSHS_2	Grey_GSHS_3
TPR	80.00%	80.00%	80.00%	100.00%
FPR	38.89%	18.33%	12.78%	2.78%
accuracy	61.62%	81.62%	87.03%	97.30%
precision	5.41%	10.81%	14.81%	50.00%
f-measure	10.13%	19.05%	25.00%	66.67%
TP	4	4	4	5
TN	110	147	157	175
FN	1	1	1	0
FP	70	33	23	5

VI. CONCLUSION

This study has proposed a novel hybrid grey forecasting and geometric selective harmony search approach for using small data samples of particle filters of a pneumatic conveyor machine in an LED epitaxy plant to conduct the CBM for particle filters. Data is preprocessed by data filtering and clustering analysis to reduce data noises and increase influence of key data used in later grey forecasting. Experimental results of a real instance show that the proposed approach performs well. Although this study is designed for LED epitaxy plants, it can be applied to detecting rare events in other fields (e.g., identifying pathologic cells [24]) or extension to other problems with imbalanced data samples (e.g., [25]).

REFERENCES

[1] H. Gao, S. Kwong, B. Fan, and R. Wang, "A hybrid particle-swarm tabu search algorithm for solving job shop scheduling problems," *IEEE Trans. Ind. Informat.*, vol. 10, no. 4, pp. 2044–2054, Nov. 2014.

[2] S. Nguyen, M. Zhang, M. Johnston, and K. C. Tan, "Automatic design of scheduling policies for dynamic multi-objective job shop scheduling via cooperative coevolution genetic programming," *IEEE Trans. Evol. Comput.*, vol. 18, no. 2, pp. 193–208, Apr. 2014.

[3] M. Foumani, I. Gunawan, and K. Smith-Miles, "Notes on feasibility and optimality conditions of small-scale multifunction robotic cell scheduling problems with pickup restrictions," *IEEE Trans. Ind. Informat.*, vol. 11, no. 3, pp. 821–829, Jun. 2015.

[4] O. Fink, E. Zio, and U. Weidmann, "A classification framework for predicting components' remaining useful life based on discrete-event diagnostic data," *IEEE Trans. Rel.*, vol. 64, no. 3, pp. 1049-1056, Sep. 2015.

[5] J. B. Ali, B. Chebel-Morello, L. Saidi, S. Malinowski, and F. Fnaiech, "Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network," *Mech. Syst. Signal Pr.*, vol. 56–57, pp. 150–172, May 2015.

[6] J.-L. Deng, "Control problems of grey systems," *Syst. Control Lett.*, vol. 1, pp. 288–294, Mar. 1982.

[7] M. Xia and W. K. Wong, "A seasonal discrete grey forecasting model for fashion retailing," *Knowl.-Based Syst.*, vol. 57, pp. 119–126, Feb. 2014.

[8] J. Shi, Z. Ding, W.-J. Lee, Y. Yang, Y. Liu, and M. Zhang, "Hybrid forecasting model for very-short term wind power forecasting based on grey relational analysis and wind speed distribution features," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 521–526, Jan. 2014.

[9] X.-J. Guo, S.-F. Liu, L.-F. Wu, Y.-B. Gao, and Y.-J. Yang, "A multi-variable grey model with a self-memory component and its application on engineering prediction," *Eng. Appl. Artif. Intel.*, vol. 42, pp. 82–93, Jun. 2015.

[10] Y.-S. Lee and W.-Y. Liu, "Forecasting value of agricultural imports using a novel two-stage hybrid model," *Comput. Electron. Agric.*, vol. 104, pp. 71–83, Jun. 2014.

[11] Z. Liang and A.K. Parlikad, "A condition-based maintenance model for assets with accelerated deterioration due to fault propagation," *IEEE Trans. Rel.*, vol. 64, no. 3, pp. 972–982, Sep. 2015.

[12] T.-B. Xia, X.-N. Jin, L.-F. Xi, Y.-J. Zhang, and J. Ni, "Operating load based real-time rolling grey forecasting for machine health prognosis in dynamic maintenance schedule," *J. Intell. Manuf.*, vol. 26, pp. 269–280, Apr. 2015.

[13] X. Dai and Z. Gao, "From model, signal to knowledge: A data-driven perspective of fault detection and diagnosis," *IEEE Trans. Ind. Informat.*, vol. 9, no. 4, pp. 1551–3203, Nov. 2013.

[14] C. Chen, B. Zhang, G. Vachtsevanos, and M. Orchard, "Machine condition prediction based on adaptive neuro-fuzzy and high-order particle filtering," *IEEE Trans. Ind. Electron.*, vol. 58, pp. 4353–4364, Sep. 2011.

[15] M. D. Prieto, G. Cirrincione, A. G. Espinosa, J. A. Ortega, and H. Henao, "Bearing fault detection by a novel condition-monitoring scheme based on statistical-time features and neural networks," *IEEE Trans. Ind. Electron.*, vol. 60, no. 8, pp. 3398–3407, Aug. 2012.

[16] J.-H. Chen, H.-K. Li, D. Sheng, and W. Li, "A hybrid data-driven modeling method on sensor condition monitoring and fault diagnosis for power plants," *Int. J. Elec. Power*, vol. 71, pp. 274–284, Oct. 2015.

[17] M. Pratama, M. J. Er, X. Li, R. J. Oentaryo, E. Lughofer, and I. Arifin, "Data driven modeling based on dynamic parsimonious fuzzy neural network," *Neurocomputing*, vol. 110, pp. 18–28, Jun. 2013.

[18] A. Haghani, T. Jeinsch, and S. X. Ding, "Quality-related fault detection in industrial multimode dynamic processes," *IEEE Trans. Ind. Electron.*, vol. 61, pp. 6446–6453, Nov. 2014.

[19] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mech. Syst. Signal Pr.*, vol. 20, pp. 1483–1510, Oct. 2006.

[20] D. A. Tobon-Mejia, K. Medjaher, N. Zerhouni, and G. Tripot, "A data-driven failure prognostics method based on mixture of Gaussians hidden Markov models," *IEEE Trans. Rel.*, vol. 61, pp. 491–503, Jun. 2012.

[21] O. Geramifard, J.-X. Xu, and S. K. Panda, "Fault detection and diagnosis in synchronous motors using hidden Markov model-based semi-nonparametric approach," *Eng. Appl. Artif. Intel.*, vol. 26, pp. 1919–1929, Sep. 2013.

[22] J. Neuzil, O. Kreibich, R. Smid, "A distributed fault detection system based on IWSN for machine condition monitoring," *IEEE Trans. Ind. Informat.*, vol. 10, no. 2, pp. 1118–1123, May 2014.

[23] M. Castelli, S. Silva, L. Manzoni, and L. Vanneschi, "Geometric selective harmony search," *Inform. Sciences*, vol. 279, pp. 468–482, Sep. 2014.

[24] E. Szekely, A. Sallaberry, F. Zaidi, and P. Poncelet, "A graph-based method for detecting rare events: Identifying pathologic cells," *IEEE Comput. Graph. Appl. Mag.*, vol. 35, no. 3, pp. 65–73, May-Jun. 2015.

[25] N. García-Pedrajas, J. Perez-Rodríguez, and A. de Haro-García, "OligoIS: Scalable instance selection for class-imbalanced data sets," *IEEE Trans. Cybern.*, vol. 43, no. 1, pp. 332–346, Feb. 2013.