Journal of Manufacturing Systems 62 (2022) 249–262

Contents lists available at ScienceDirect

Journal of Manufacturing Systems

journal homepage: www.elsevier.com/locate/jmansys

Technical Paper

Prediction of per-batch yield rates in production based on maximum likelihood estimation of per-machine yield rates

Philip F.E. Adipraja [a], Chin-Chun Chang [b], Wei-Jen Wang [a,*], Deron Liang [a]

[a] Department of Computer Science and Information Engineering, National Central University, Taoyuan City, 32001, Taiwan
[b] Department of Computer Science and Engineering, National Taiwan Ocean University, Keelung City, 20224, Taiwan

ARTICLE INFO

Keywords:
Batch yield-rate prediction
EM algorithm
Machine yield-rate estimation
Manufacturing process

ABSTRACT

The demand for high-quality customized products compels manufacturers to adopt batch production. With the ability to accurately estimate batch production yield rates in advance, manufacturers can effectively plan the batch production process and control the production risk based on the estimated values. The per-batch production yield rates can be directly predicted by multiplying the accurately estimated per-machine yield rates corresponding to a batch. Unfortunately, for most manufacturers, the actual per-machine yield rates are difficult to estimate owing to a variety factors. Moreover, per-batch yield-rate prediction has received little attention because recent studies only focused on yield-rate prediction methods for single/continuous production systems. To address this, we propose an expectation-maximization-based approach to predict per-batch yield rates by estimating the per-machine yield rates. Based on the data from T-company, the proposed method could predict the per-batch yield rates for the subsequent week with an average accuracy of 91.86 %, and for five consecutive weeks with an average accuracy of more than 90 %. To further evaluate the performance of the proposed method with different batch production patterns, we conducted simulations to obtain the average accuracy of the estimated per-machine yield rates. In the simulations, the average prediction accuracy of the per-batch yield rates was 91.29 % in the batch production pattern, as in the case of T-company (∼250 machines and ∼1000 batches per week), and it increased as the number of batches increased.

1. Introduction

The dynamic nature of market demand compels most manufacturers to offer high-quality customized products [1,2]. To address this challenge, manufacturers should respond swiftly to changes in customer's demands [3]. Since batch production enables mass production of high-variety customized products [4], it has been widely used by many manufacturers all over the world [5]. The major challenge of batch production involves good production planning because of the management complexity, including the factors of numerous products, production stations, and machines. One way to better plan batch production and control risk is to estimate the yield rate of their products accurately. This also allows manufacturers to adjust their parameters and estimate as well as evaluate their production [6–9]. In the event that manufacturers find it difficult to meet the demands of the customer, manufacturers could look for another strategy [10], such as adaptive sourcing, which could achieve business outcomes while controlling risk. As a result, the ability to estimate the yield rate of a batch production system is important to the modern manufacturing industry.

In batch-production manufacturing, manufacturers should plan the machines to be used for each batch product. They should also predict the yield rate of each batch that is manufactured without performing rework or correction, which is called first-pass yield (FPY) [11]. According to FPY theory, the yield rate of a production process is equal to the product of all the machine yield rates involved. As a result, the per-batch production yield rates can be directly predicted by accurately estimating the yield rate of each production machine (per-machine yield rates), and subsequently using the estimated per-machine yield rates to calculate future per-batch yield rates. However, the actual per-machine yield rates are generally difficult to estimate because they are affected by multiple factors, such as production process drift, environment, and machine condition, misconfiguration, and age [12–15]. Accordingly, related studies are scarce, and the problem remains challenging.

Although studies on per-batch production yield-rate prediction are scarce, several yield-rate prediction methods for single/continuous production systems have recently been proposed [16–20]. These

* Corresponding author.
E-mail address: wjwang@csie.ncu.edu.tw (W.-J. Wang).

https://doi.org/10.1016/j.jmsy.2021.11.015
Received 15 April 2021; Received in revised form 24 November 2021; Accepted 25 November 2021

methods can be classified into *macro-yield-modeling* or *micro--yield-modeling* [21]. Table 1 provides a detailed comparison of macro-yield-modeling and micro-yield-modeling. In macro-yield-modeling, only large a priori factors are considered (e.g. time-series data of product yield rates). Therefore, these methods may ignore the factors of particular products or conditions of the manufacturing environment at the time. Conversely, in micro-yield-modeling, more detailed information is considered, such as different classes of defect categories, layouts, and process variations of circuit design. Most recent studies employ micro-yield-modeling and adopt deep-learning methods [16–18] to predict the yield rates since they have demonstrated outstanding performance in smart manufacturing [22] such as fault diagnosis [23] and defect detection [24]. Moreover, there are a few macro-yield-modeling approaches that only focus on time series data and utilize fuzzy forecasting methods to predict the yield rates [19,20]. Unfortunately, these methods cannot be reused directly in batch production systems, in which the number of produced pieces is relatively small, and the number of product types is large compared with the corresponding numbers in single/continuous production systems. Therefore, the amount of the accumulated production data for each batch product may not be sufficient to construct an efficient yield-rate prediction model, particularly when deep-learning technologies are adopted.

To address the challenge of small-volume, large-variety production in a batch production system, we propose a simple macro-yield-modeling approach to predict the per-batch production yield rates by using the estimated per-machine yield rates. In this approach, the actual per-machine yield rates can be calculated based on the number of observed defective products manufactured by a particular machine. However, defective products can only be observed using quality inspection devices, which are expensive to use for all production machines [25]. Consequently, manufacturers may reduce the number of inspection devices if possible, and they should achieve a balance between the number of inspection devices and the ability to control production quality [2,15]. Therefore, in practice, it is difficult to determine the actual per-machine yield rates given a limited number of inspection devices.

In this study, we developed a novel approach based on the expectation-maximization (EM) algorithm to estimate the per-machine yield rates in a production process. We assume that the per-machine yield rates do not change in the short term, and therefore the per-machine yield rates can be used to predict future per-batch yield rates. As a result, one or two weeks of production data are appropriate to predict the per-batch yield rate for subsequent weeks. The rationale for using the EM-based algorithm is that it requires relatively simple information, including production paths, the number of raw products, and the number of defective products detected by each inspection device, to estimate the per-machine yield rates. As a result, the proposed method can be applied to a manufacturer, even though a large number of sensors to detect the production status or machine production conditions, such as temperature, humidity, and vibrations, are not available. It should be noted that the EM algorithm has several limitations, including slow convergence and convergence to local optima [26–28]. To overcome these limitations, we used constrained least-squares to obtain appropriate initial values, which later improved the maximization of the likelihood function. Additionally, the proper parameters for the stopping criterion, such as the maximum number of iterations and the threshold value for MSE, are set based on our preliminary experiments. The results suggested that approximately 60–100 iterations and an MSE threshold of 0.001 are sufficient to stop the algorithm.

Table 1
Comparisons on the proposed method and recent related studies.

	Micro-Yield-Modeling [16–18]	Macro-Yield-Modeling [19,20]	Macro-Yield-Modeling Proposed Method
Required Data	Detailed information, e.g. *different classes of defect categories, layouts, and process variations of circuit design*	large a priori factors such as *product yield rates* (time series).	large a priori factors for each batch as follows: *production paths, number of raw products, and number of observed defective products*
Method	Using Deep learning	Using fuzzy collaborative forecasting	Maximum likelihood (EM-Based algorithm)
Manufacturing type	Single/Continuous flow	Single/Continuous flow	Batch production system
Limitations	Required prohibitively large data sets and long computational time	Required a particular statistical model from human expert for each manufacture flow.	Unsuitable for continuous flow

To validate the efficiency of the proposed method, we used time-series data from T-company. The datasets contain up to 70 weeks of real-world production data that were collected through sensors and human operators; a week of the data contains thousands of batches of products and more than 200 machines. We used one-week data to build the prediction model and then used the model to predict the per-batch yield rate in the subsequent weeks. The results demonstrate that our approach can predict the per-batch yield rates of the subsequent week with an average accuracy of 91.86 %, and the production yield rates of the subsequent five weeks with an average accuracy of more than 90 %.

The real data from T-company only indicate that the proposed method achieves good prediction accuracy with this particular batch production pattern. To further understand whether the proposed method is effective with different batch production patterns, we used simulations. For each simulation, we set the per-machine yield rates along with the batch production parameters, and then simulated the batch production process. Subsequently, we used the proposed EM-based algorithm to obtain the estimated per-machine yield rates. In general, if the estimated per-machine yield rates approach closer to the pre-set per-machine yield rates (ground truth), better prediction accuracy is achieved. Therefore, we used the average accuracy of the estimated per-machine yield rates to evaluate the performance of the proposed method with different batch production patterns. The results indicated that, by using a batch production pattern similar to that of the actual data set for T-company (∼250 machines and ∼1000 batches per week), the average accuracy of our approach for per-machine yield-rate estimates is 91.84 %, with a minimum accuracy of 62.51 %, a maximum accuracy of 100 %, and a standard deviation of 6.12 %. Furthermore, the average accuracy of our approach for per-batch yield-rate estimates is 91.29 %, with a minimum accuracy of 90.54 %, a maximum accuracy of 92.08 %, and a standard deviation of 0.19 %. The accuracy increased as the number of batches increased.

The proposed method has several managerial implications. First, manufacturers can plan the number of additional production pieces for each production batch using the proposed method. This facilitates the control of production cost as well as risk. Second, our approach uses only production data without an excessive number of parameters. Therefore, it has less management overhead than other recent deep-learning methods. It is conceivable that manufacturers with limited resources can easily implement this approach. The contributions of this study are as follows.

- The proposed method can predict per-batch yield rates with satisfactory accuracy, as verified using real-world datasets in our experiments.

- The proposed method uses only common production data, and therefore it is considered more lightweight and inexpensive compared with recent deep-learning methods.

The remainder of this paper is organized as follows. In Section 2, we review existing approaches of production yield-rate prediction. In Section 3, we describe the maximum likelihood estimation of per-machine yield rates based on the observation of defective products. In Section 4, we detail the proposed method to predicting the per-batch yield rates based on the estimated per-machine yield rates. In Section 5, we discuss the experiment and simulation design, as well as the implications of our approach. Finally, the last section concludes the paper and proposes directions for future research.

2. Related work

It is essential and beneficial for manufacturers to accurately estimate the yield rates of their production. In terms of yield-rate prediction, manufacturers that adopt single/continuous production systems usually have a large homogeneous dataset for their products. This is because a single production system is associated with a set of machines, which are organized to manufacture a single type of product [29]. By contrast, manufacturers that adopt batch production systems usually have large heterogeneous datasets for their products. These systems handle a large variety of products that require different sets of production stations and machines [4]. Typically, each batch is considered independent and may not be associated with other batches. Thus, each batch production process usually generates a small dataset. Regardless of using single/continuous or batch production, it is possible to estimate the production yield rate by using actual per-machine yield rates. Estimating the actual per-machine yield rates is a difficult challenge in practice [12–15].

Even though per-batch production yield-rate prediction has received little attention, several recent studies have been concerned with the yield-rate prediction for single/continuous production [16–20], as summarized in Table 1. These studies can be classified as either macro-yield modeling or micro-yield modeling [21]. The former considers only large *a priori* factors, whereas the latter considers detailed information, such as different classes of defect categories, layouts, and process variations of circuit design.

Many recent macro-yield-modeling approaches utilize time-series data as input to predict the production yield rate. Chen and Chiu [19] proposed an approach based on time-series production data that uses an interval-fuzzy-number-based fuzzy collaborative forecasting scheme to predict the DRAM yield rate. Their approach performed well, with a mean absolute percentage error (MAPE) of less than 2.17 %. Chen and Wu [20] proposed a similar approach, which predicts the DRAM yield rate using fuzzy collaborative forecasting, and only requires time-series production data. Although this approach requires simple data, human experts should construct the fuzzy yield forecast. Therefore, manufacturers that offer multiple customized products in large numbers of batches may find this approach excessively effort-intensive.

Most recent studies based on micro-yield-modeling adopt deep-learning methods. Jun et al. [16] proposed constructing a model that predicts defects in the production process. This approach requires several variables from the production process, such as temperature, humidity, and other production variables. Initially, each product piece is labeled as either defective or good through machine learning. Subsequently, a recurrent neural network is used to analyze the time-series data and predict the feature data. Finally, a machine learning algorithm is used to classify each piece based on the previous steps. This approach could be used to improve future yield by using predictions to reduce the occurrence of defects. The authors reported that this approach can improve yield by approximately 8.7 % in a continual process. Two similar approaches have recently been proposed [17,18]. With the abundance of monitored data obtained from the manufacturing process, the production yield rate can be predicted using these approaches. However, the computational cost is prohibitively high because deep learning is used [30]. Most importantly, it is quite difficult to directly reuse these approaches in other domains and manufacturing processes. This is because important manufacturing parameters should be identified, and the prediction model should be reconstructed and justified.

Most large-scale manufacturers may invest in a large number of suitable sensors to obtain production data [9,31], including actual per-machine yield-rate data. They can then develop a per-batch yield-rate prediction system based on existing deep-learning approaches. Unfortunately, small- and medium-scale manufacturers may have limited resources for obtaining actual per-machine yield-rate data. Therefore, they should use other inexpensive and lightweight approaches. Compared with existing approaches, our approach can provide highly accurate predictions without requiring an excessive amount of resources. Accordingly, all manufacturers can easily use it.

The EM algorithm is the core technique of the proposed yield-rate prediction approach. The challenge of using the EM algorithm is that it has several limitations, such as slow convergence and convergence to local optima [27,28]. It has been demonstrated that the initial values of the EM parameters may lead to slow convergence [27,28,32,33]. The EM may also stop at some point before reaching the optimal likelihood. Therefore, it is suggested that the appropriate initial values of the EM parameters be determined. It is also suggested that an adequate number of iterations be determined to obtain the maximum likelihood. Several prior studies have suggested several ways to mitigate the limitations of the EM algorithm. For example, Shireman et al. [34] compared five techniques for obtaining starting values: random starting values, the K-means clustering technique, the iteratively constrained EM technique, the agglomerative hierarchical clustering, and the sum scores technique. Their simulations demonstrated that the technique involving random values is recommended if analyses should be run quickly, and the iteratively constrained EM technique is preferable to obtain the best results. In the proposed method, we use constrained least squares to obtain the proper initial values. Furthermore, we determined a stopping criterion based on the closest result to the maximum likelihood. In addition, our mathematical model uses Bernoulli trials, which are well suited to the EM algorithm [35]. This is because a closed-form solution for the parameter is available at the M-step.

3. Maximum-likelihood estimation of per-machine yield rates

The notations used in this paper are defined in Tables 2 and 3. Specifically, Table 2 contains all notations with previously known values

Table 2
List of notations with previously known values.

Notation	Description
I	The total number of batches in the manufacturing process
N_i	The starting number of pieces in the ith batch
J_{in}	The number of manufacturing steps before the nth piece being discarded as defective or fully completing the process of the ith batch
y_{in}	The condition of nth piece (of the ith batch in the manufacturing process) observed to be defective (value of 1) or in good condition (value of 0),
C_i	The number of manufacturing steps required to completely process the ith batch
l_i	The actual yield rate of the ith batch (ground truth).
b_{ij}	The number of defective pieces observed in the jth manufacturing step of the ith batch.
S_m	The set of machine indexes in (i,j,k) indexes (as tuple elements) of all batches; which are the machines in kth manufacturing step of the ith batch (where defective pieces are observed in the jth manufacturing step)
m_{ik}	The machine used in the kth manufacturing step of the ith batch
f_{im}	The number of good pieces at the end of the manufacturing process of the ith batch that uses machine m
r_{im}	How many times machine m is used in the manufacturing process of the ith batch.

Table 3
List of notations with unknown initial values.

Notation	Description
$P(Y,Z\|\theta)$	The complete likelihood function for Y (set of the piece conditions at each step of all batches) and Z (set of the indicator variable of a machine causing defective pieces in each manufacturing step in all batches) conditioned on the machine yield rates θ.
z_{ink}	The indicator variable ($\in \{0,1\}$) of the nth piece of the ith batch observed to be defective due to the machine used in the kth manufacturing step. For example, if $z_{ink} = 1$, then the nth piece of the ith batch is defective due to this kth machine.
$F_D(i,n)$	The likelihood function of the defect rate if the nth piece of the ith batch is observed to be defective in a manufacturing process
$F_G(i,n)$	The likelihood function of the yield rate if the nth piece of the ith batch is observed to be a good piece in a manufacturing process
P_{ik}	The yield rate of a machine in the kth manufacturing step of the ith batch (the probability that a piece of product will be good when using the machine associated with the kth manufacturing step of the ith batch).
q_{ik}	The natural logarithm of P_{ik}
θ	The set $\{q_{ik}\|(1 \le i \le I) \wedge (1 \le k \le J_i)\}$
$Pr(m)$	The yield rate of machine m (the probability of obtaining good pieces by using machine m).
d_m	The total expected number of defective pieces generated by machine m in every manufacturing step in all batches.
g_m	The total expected number of good pieces generated by machine m in every manufacturing step in all batches.
$P(z_{ink} = 1\|y_{in}, \theta)$	The likelihood that the kth manufacturing step of the ith batch causes the nth piece to be defective (value of 1) with the given θ
$E[z_{ink}]$	The expectation that kth manufacturing step causes the nth piece of the ith batch to be defective
e_{ijk}	The expected number of defective pieces generated in the kth manufacturing step when any defective pieces of the ith batch are observed in jth manufacturing step
h_{ijk}	The expected number of good pieces produced by the kth manufacturing step, in which these pieces are later observed as defective in the jth manufacturing step (an inspection step) for the ith batch. In other words, the pieces become defective in one of the subsequent manufacturing steps after the kth manufacturing step, and then they are observed as defective in the jth manufacturing step
x_m	The total number of potential defective pieces generated by machine m in every manufacturing step in all batches.
$\overline{acc}$	The average accuracy of per-batch yield-rate prediction.

that can be extracted from real-world data, whereas Table 3 contains all notations, the initial values of which are unknown and are to be estimated later. In our study, we make several assumptions that are summarized as follows:

- We assume that the actual per-machine yield rates do not change in a short period.
- We assume that all defects are caused by production machines; in the real world, a defective product may be affected by several factors, including the quality of raw materials and the human operator of the production process [36].
- We assume that defective pieces are observed at a particular manufacturing step with inspection equipment, and the defects could have occurred before this step. This is because inspections can only be conducted in some manufacturing steps. Therefore, when a piece becomes defective in a manufacturing step, it will go through subsequent manufacturing steps until found by the next inspection step. This assumption can affect the per-machine yield rate estimation for a manufacturing sequence with few inspection devices. With this assumption, the proposed method is very likely to overestimate the per-machine yield rate of the manufacturing step with actual low yield and to underestimate those of the other manufacturing steps with actual good yields. However, the estimation accuracy improves when the number of batches increases.
- We assume that the observed defective pieces in a step will be removed, and only the good pieces (including the unobserved defective pieces) will be processed in the next step. After identifying the defective intermediate components or final products, manufacturers generally utilize a smart approach to determine whether the defective pieces should be reworked or disposed [37]. However, this assumption is based on FPY [11], which removes the pieces from the manufacturing process right after it is observed as defective.

To accelerate the production of large quantities of products, manufacturers may divide the manufacturing process into several jobs called *batches*. Then, each batch is tied to a *batch number* based on its *bill of operation* (BoO) [4] for future reference, where the BoO contains operational information, such as the sequence of stations for each batch.

Fig. 1 shows the processing of a batch of products by any machine in a station described in the BoO at the time of manufacturing. However, manufacturers can use their machines for many purposes. Any machine can be used by one or several BoOs in any of their sequences. Moreover, a batch may use the same machine more than once in its sequence if that batch requires some stations to be revisited.

Manufacturers use quality inspection equipment to control the production of each batch [38]. However, manufacturers attempt to balance the ability to control production quality and to control production cost [15]. Therefore, the inspections can only be carried out in a few manufacturing steps. Fig. 1 shows an example in which several stations and machines are involved in the production process, and the inspections could only be carried out at Station W and Station Z. Consequently, we expect no defective pieces to be observed by machines at other stations. Let us consider batch C1 in this example, as shown in Fig. 2. The defective pieces can only be observed in Mch-2 of Station W and Mch-6 of Station Z. It should be noted that the final manufacturing step is assumed to have the inspection equipment for product quality control.

Although a manufacturer may process many batches daily, each machine can only handle a single batch at a time, including the inspection equipment. Accordingly, for each defective piece observed by a machine in the jth manufacturing step, all machines in the previous manufacturing steps, including the current one, are suspects (represented as the kth manufacturing step). That is, the suspect machines are those that may have produced defective pieces. For example, as shown in Fig. 2, inspection devices are installed on Mch-2 and Mch-6. In this case, Mch-2 as a machine in the second manufacturing step ($j = 2$) detects three defective pieces, which may have been generated by any previous machines, including the current one. Therefore, Mch-3 and Mch-2 (represented as the kth suspect machines) constitute the set of suspect machines for the observed defects. In another example, Mch-6 in the fourth manufacturing step ($j = 4$) also detects one defective piece. In this case, Mch-3, Mch-2, Mch-4, and Mch-6 constitute the set of suspect machines for the observed defects. Thus, it is possible to estimate the yield rate of each production machine based on the observed defective pieces in each manufacturing process for each batch.

In this study, we assume that the observed defective pieces in a step are removed, and only the good pieces will be processed in the next step. Accordingly, the total number of defective products in a batch is equal to the sum of the observed defective pieces in each manufacturing process. Therefore, in each manufacturing process, the number of defective pieces can be estimated from the yield rate of each machine through which the product passes (let the variable be θ). Accordingly, based on the FPY theory, we designed a likelihood function for each manufacturing step, which is used to estimate the per-machine yield rates, as shown in (4). Meanwhile, we provide Eqs. (1)–(3) to explain Eqs. (4) step by step.

There are several z_{ink} that affect each y_{in}, given machine yield rates θ. For example, if y_{in} observed as good pieces ($y_{in} = 0$), the corresponding z_{ink} will all be 0; otherwise, one of the corresponding z_{ink} will be 1. Accordingly, let Y be the set of y_{in} of all batches (ranging from 1 to I), and Z be the set of z_{ink} in each manufacturing step in all batches (ranging from 1 to I). Then, the complete likelihood function for Y and Z condi-

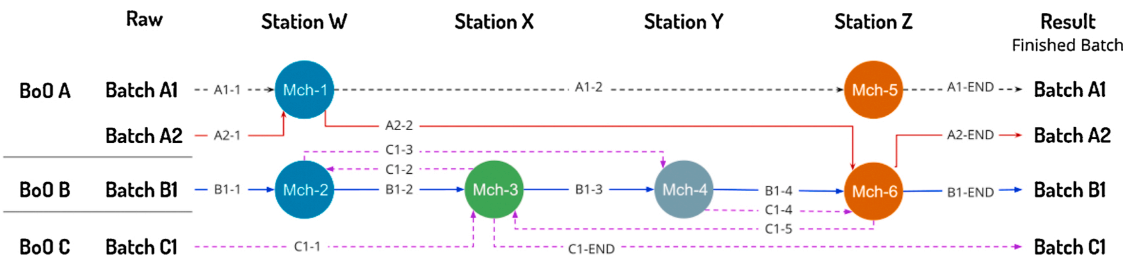

Fig. 1. Example involving three bills of operation (A to C) that use four stations (W to Z) to produce four batches (A1 to C1).

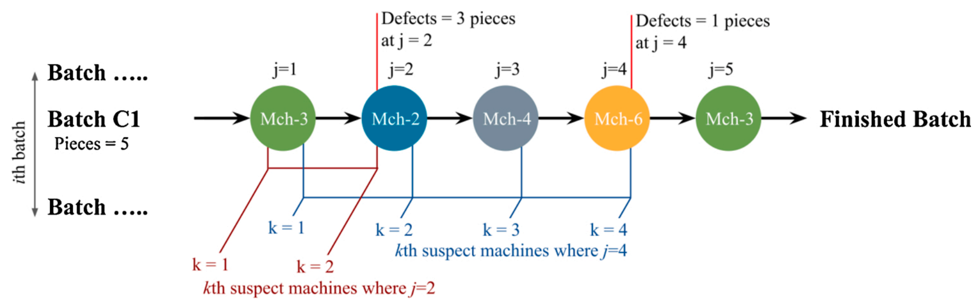

Fig. 2. Illustration of suspect machines generating defective pieces.

tioned on the machine yield rates θ can be represented as $P(Y,Z|\theta)$. Based on this, we can define $Y = \{y_{in} \mid i,n \in \mathbb{N}\ \&\ 1 \le i \le I\ \&\ 1 \le n \le N_i\}$ and $Z = \{z_{ink} \mid i,n,k \in \mathbb{N}\ \&\ 1 \le i \le I\ \&\ 1 \le n \le N_i\ \}$. Then, $P(Y,Z|\theta)$ can be calculated based on each piece in each batch of the manufacturing process, as follows:

$$P(Y,Z|\theta) = \prod_{i=1}^{I}\prod_{n=1}^{N_i}\left(F_D(i,\ n)^{y_{in}} \times F_G(i,\ n)^{1-y_{in}}\right), \tag{1}$$

where the $F_D(i,\ n)$ is the likelihood function of the defect rate, whereas the $F_G(i,\ n)$ is the likelihood function of the yield rate. The $F_D(i,\ n)$ and $F_G(i,\ n)$ are defined as follows:

$$F_D(i,\ n) = \prod_{k=1}^{J_{in}}\left((1-P_{ik})\prod_{s=1}^{k-1}P_{is}\right)^{z_{ink}}, \tag{2}$$

$$F_G(i,\ n) = \prod_{k=1}^{J_{in}}P_{ik}. \tag{3}$$

Several product pieces are manufactured in each batch. Therefore, the likelihood for each batch is calculated based on the observed condition of each piece. The observed condition of a product piece (defective or in good condition) is denoted as y_{in}. Therefore, based on (1), if the nth piece is observed to be good ($y_{in} = 0$) in the jth step of the manufacturing process, then the calculation is subject to $F_D(i,\ n)$. Otherwise, if the nth piece is observed to be defective ($y_{in} = 1$) in the jth step of the manufacturing process, then the calculation is subject to $F_G(i,\ n)$.

Subsequently, F_D is calculated based on the defect rate of the production machines (represented as $1-P_{ik}$), as shown in (2). In each manufacturing process, the likelihood is the defect rate of the current manufacturing process $(1-P_{ik})$ multiplied by the yield rate of the previous production machines $(\prod_{s=1}^{k-1}P_{is})$. F_D can be used to estimate the indicator variable of the nth piece of the ith batch observed to be defective owing to the machine used in the kth step, and the indicator variable is denoted by z_{ink}. Meanwhile, F_G is calculated based on the yield rates of the production machines, as shown in (3). The calculation is straightforward, involving multiplication of the production yield rates (denoted by P_{ik}) of all the machines that were used in a specific batch.

As shown in Fig. 1, because the BoO only specifies the production plan that uses a station sequence to process each batch, the manufacturer must assign an available machine at each station in the station sequence to the batch during production. In (2) and (3), the yield rate of each machine (P_M) is mapped using the $i;k$ index, which yields the notation $P_{M(i;k)}$. However, this notation complicates the equation. Therefore, we simplify $P_{M(i;k)}$ as P_{ik} to improve readability. Finally, the most detailed version of our equation of $P(Y,Z|\theta)$ is given in (4), as follow:

$$P(Y,Z|\theta) = \prod_{i=1}^{I}\prod_{n=1}^{N_i}\left(\left(\prod_{k=1}^{J_{in}}\left((1-P_{ik})\prod_{s=1}^{k-1}P_{is}\right)^{z_{ink}}\right)^{y_{in}} \times \left(\prod_{k=1}^{J_{in}}P_{ik}\right)^{1-y_{in}}\right). \tag{4}$$

Accordingly, given that our objective is to estimate the yield rate of each machine, the complete likelihood function in (4) can be used to estimate the defective pieces produced by each machine; this can subsequently be used to estimate the yield rate of each machine. Finally, another reason for using EM is that it is well suited for solving the mathematical model of the problem in (4).

4. Proposed method

Because our approach is related to unknown or missing per-machine yield-rate data, we propose a new EM-based algorithm whereby the estimated per-machine yield rates can be used to make predictions. In the proposed method, the per-machine yield rates are estimated by iterating the EM algorithm until convergence is attained. The overall procedure of our approach is shown in the activity diagram in Fig. 3. Specifically, first, in step (a), we should preprocess and clean the raw data, and then guess the initial per-machine yield rates in step (b). Step (h) improves the estimation of the yield rate of each machine. It is calculated based on the expected number of good pieces and the expected number of defective pieces produced by each machine. Both variables are estimated based on the observed defective pieces in step (c). Because both variables can be calculated independently, we use the parallelism (bar) symbol after step (c). The expected number of good pieces can be sequentially calculated by steps (e), (f), and (g), and the expected number of defective pieces can be calculated by step (d). The bottom bar symbol indicates that we should use both variables in step

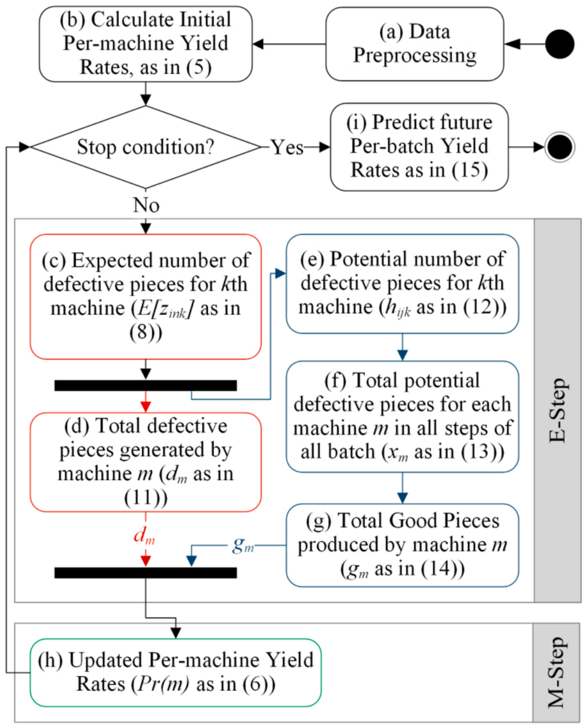

Fig. 3. Activity diagram of the proposed method to predicting the per-batch yield rates.

(h). Finally, when the EM iteration stops in step (i), the future per-batch yield rates can be predicted using the previously estimated per-machine yield rates.

The hyperparameters in this approach are the number of iterations and the convergence threshold. We limit the stopping criterion to a threshold of 0.001. This implies that the result of current iteration is very close to that of the previous iteration. Fig. 4 shows that a maximum number of iterations of approximately 60–100 are sufficient to obtain the optimal value (near the convergence threshold). The details of the algorithm are explained in the following subsections.

4.1. Preprocessing

Because the raw production data provided by most manufacturers are not clean, data preprocessing and cleaning are required, as indicated by step (a) in Fig. 3. Several steps must be executed to preprocess and clean the data. First, because the problem involves estimating the yield rate of each machine, we can exclude manufacturing process data related to manual or human labor. Second, because the quality inspection equipment is only installed in several machines, the observed defective pieces are set to zero for other machines, where it is impossible to observe defective pieces. Third, a manufacturing process may be divided into two or more sessions, which may result in the addition of several distinct datasets related to the same manufacturing process. To address this problem, we must merge these sessions into one session. Finally, an example of the required preprocessing based on Fig. 1 is summarized in Table 4.

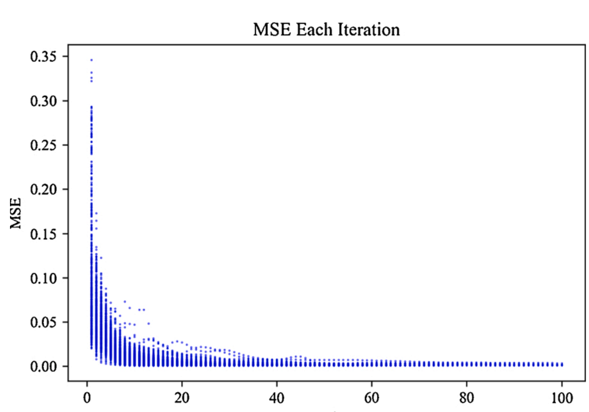

Fig. 4. MSE for each iteration in preliminary experiments.

4.2. Initial per-machine yield rates

Because the initial per-machine yield rates in the EM algorithm are unknown, we must estimate the yield rates of all the machines used in the manufacturing process, as indicated by step (b) in Fig. 3. In addition to random guessing, several approaches can be used to estimate the yield rates. For example, we can use constrained least squares to generate the initial per-machine yield rates based on the per-batch yield rates, as follows:

$$\arg\min_{\theta} \sum_{i=1}^{I} \left(\sum_{k=1}^{C_i} q_{ik} - \ln l_i \right)^2 \tag{5}$$

subject to $q_{ik} \leq 0$.

4.3. M-Step

The yield rates θ are the initial input for the E-step of the proposed method. Subsequently, the output of E-Step is used to improve the yield-rate estimation in M-Step, which is the objective of each iteration, as indicated by step (h) in Fig. 3. The yield rate of a particular machine is the percentage of good pieces among all pieces generated by the machine. Therefore, the objective of each iteration can be expressed as follows:

$$Pr(m) = \frac{g_m}{g_m + d_m}. \tag{6}$$

4.4. E-Step

To obtain d_m and g_m, we must estimate the number of defective pieces produced by suspect machines, as indicated by step (c) in Fig. 3. The proposed method is based on the principle of likelihood. In the real world, the yield rates of most machines used in production should be close to 0.999. However, when inspection equipment observes defective pieces, all the previous machines, including the current machine, which are used to process a batch, are suspects. Therefore, we expect the shared probability of defective pieces to be distributed among the suspect machines based on their yield rates. A particular machine could have a lower yield-rate estimation if multiple batches have many defective pieces after this machine is used. Consequently, if any defective pieces are observed in each manufacturing step of each batch, we must first estimate the likelihood that each particular machine is responsible for

Table 4
Example of required production data.

Batch Number	Production Sequence	Machine	# of processed pieces	# of observed defective pieces
Batch A1	1	Mch-1	10	0
	2	Mch-5	10	1
Batch A2	1	Mch-1	32	0
	2	Mch-6	32	3
Batch B1	1	Mch-2	100	0
	2	Mch-3	100	5
	3	Mch-4	95	0
	4	Mch-6	95	9
…	…	…	…	…

that defect (z_{ink}) by using (4). Hence, we use (7) to calculate the likelihood of the kth manufacturing step causing the nth piece of the ith batch to be defective, which is represented as $P(z_{ink} = 1|y_{in}, \theta)$.

$$P(z_{ink} = 1|y_{in}, \theta) = \begin{cases} 0; & \text{if} y_{in} = 0 \\ \dfrac{(1 - P_{ik})\left(\prod_{s=1}^{k-1} P_{is}\right)}{\sum_{t=1}^{J_{in}}\left((1 - P_{it})\left(\prod_{u=1}^{t-1} P_{iu}\right)\right)}; & \text{if} y_{in} = 1 \end{cases} \tag{7}$$

For each piece observed to be defective in the jth manufacturing step, each particular manufacturing step is probably suspected. Therefore, the expected number of defective pieces produced by each suspect machine can be calculated as follows:

$$\begin{aligned} E[z_{ink}] &= 0 * P(z_{ink} = 0|y_{in}, \theta) + 1 * P(z_{ink} = 1|y_{in}, \theta) \\ E[z_{ink}] &= P(z_{ink} = 1|y_{in}, \theta) \end{aligned} \tag{8}$$

For each manufacturing step of the ith batch, more than one piece may be observed to be defective. Consequently, several observed defective pieces have the same $E[z_{ink}]$. Subsequently, we must estimate the expected number of defective pieces for each suspect machine. Therefore, the likelihood estimation in (8) can be multiplied by the number of defective pieces observed in that manufacturing step, as follows:

$$e_{ijk} = E[z_{ink}] \times b_{ij}. \tag{9}$$

Then, to obtain d_m, we can combine e_{ijk} for each machine m in each manufacturing process from all batches, as indicated by step (d) in Fig. 3. Commonly, a machine m is used in several manufacturing steps, which are registered in the set of (i,j,k) indexes of machine m, as in (10). Accordingly, we can determine d_m as in (11).

$$S_m = \{(i, j, k)|\forall m_{ik} = m\}, \tag{10}$$

$$d_m = \sum_{(i,j,k)\in S_m} e_{ijk}. \tag{11}$$

It should be noted that g_m is associated with f_{im} and x_m. The value of f_{im} can be extracted from the data. However, to obtain x_m, we must first calculate the number of potential defective pieces in each manufacturing step, as in step (e) in Fig. 3; this can be expressed as follows:

$$h_{ijk} = \begin{cases} 0; & \text{if} j = k \\ h_{ij;k+1} + e_{ijk}; & \text{if} j > k \end{cases} \tag{12}$$

Next, we can combine the h_{ijk} of each machine in each manufacturing process of all batches to obtain x_m, as in step (f) in Fig. 3, as follows:

$$x_m = \sum_{(i,j,k)\in S_m} h_{ijk}. \tag{13}$$

Subsequently, as the values of x_m and f_{im} are known, we can calculate g_m, as in step (g) in Fig. 3. g_m is related to a production sequence in which a particular machine may be used more than once. Therefore,

$$g_m = x_m + \left(\sum_{i=1}^{I} f_{im} \cdot r_{im}\right). \tag{14}$$

4.5. Stop condition

The EM algorithm continues to iterate until a stopping condition is satisfied. Two stopping conditions can be used: a convergence threshold and the maximum number of iterations. The difference between the per-machine yield rates in the current iteration and those in the previous iteration is calculated using the mean square error (MSE) [28]. When the MSE is lower than the convergence threshold, the stopping condition is met; otherwise, if the number of iterations reaches the defined maximum number of iterations, the algorithm is stopped. In our study, a threshold of 0.001 difference is used to stop the search for a better solution. Based on our observation in preliminary experiments, a maximum number of iterations of approximately 60–100 is deemed sufficient to stop the algorithm from iterating. This does not have any significant influence in further iterations, as shown in Fig. 4.

4.6. Prediction of per-batch yield rates

At the beginning of our approach, for each week's data, we must estimate the yield rate of each machine. We assume that the per-machine yield rates estimated using the production data are a reliable basis for predicting the production yield rates. Consequently, the result of our approach can be used to predict the production yield rates, as shown in step (i) in Fig. 3. Based on the FPY, the yield rate of a batch is equal to the product of all the machine yield rates used by the batch, which is expressed as $\left(\prod_{j=1}^{C_i} P_{ij}\right)$. In this case, P_{ij} is the per-machine yield-rate estimation of the jth manufacturing step of the ith batch. To calculate the average prediction accuracy of all batches (denoted by $\overline{acc}$), we need l_i which is the actual yield rate of the ith batch. Therefore, the average prediction accuracy of all batches in a particular week can be expressed as follows:

$$\overline{acc} = \frac{\sum_{i=1}^{I}\left(1 - \left|\left(\prod_{j=1}^{C_i} P_{ij}\right) - l_i\right|\right)}{I}. \tag{15}$$

In this study, we used data from only one week to predict the production yield rates in the following weeks, up to five consecutive weeks. However, in the future, the production process may use a machine that is not used in the current week and does not have a yield-rate estimate. Although we can use older production data for that particular machine, to simplify the experiment, we ignore this case and exclude future production batches that use any machine with no yield-rate estimation in the current week.

5. Experiments and discussion

We used data from T-company (70-week data) to evaluate the prediction performance of the proposed method, as described in Subsection 5.1. To explain the high accuracy of the proposed method, we conducted several simulations, as described in Subsection 5.2. Then, we discuss the proposed method based on the experimental and simulation results.

5.1. Experimental results using T-Company data

In our study, we used a large, real-world, 70-week dataset from T-company, containing information regarding its manufacturing process. The statistics of these real-world data are summarized in Table 5. To evaluate the performance of our approach, we used 1-week data to construct our prediction model, which was subsequently used to predict the average per-batch yield rates for the following five weeks.

As shown in Fig. 5, the predicted per-batch yield rates for week 1 have an average accuracy of 91.86 %. Moreover, the results indicate that the proposed method can do predictions for weeks 2–4 (the next month) with an average accuracy of over 90 %. However, the predictions for weeks 3 and beyond are more uncertain because of a sharp increase in

Table 5
Statistics of T-Company data for up to 70 weeks.

Description	Min	Max	Mean	Stdev
Number of machines used	194	250	222	10.2
Number of batches	287	1075	747	154.3
Defective pieces in a batch	0	163,072	1140	3,325.0
Number of steps in a batch	1	59	21	7.0
Inspection steps in a batch	1 %	9 %	8 %	1 %

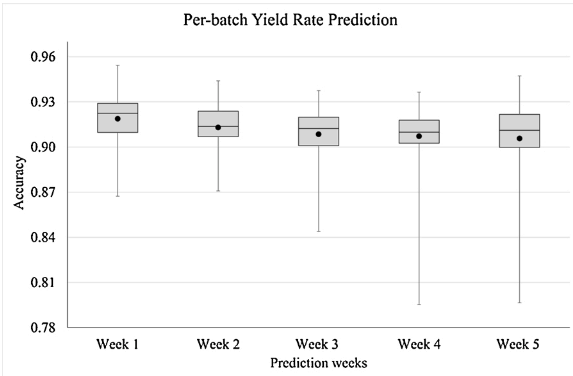

Fig. 5. The prediction accuracy of T-Company's per-batch yield-rate.

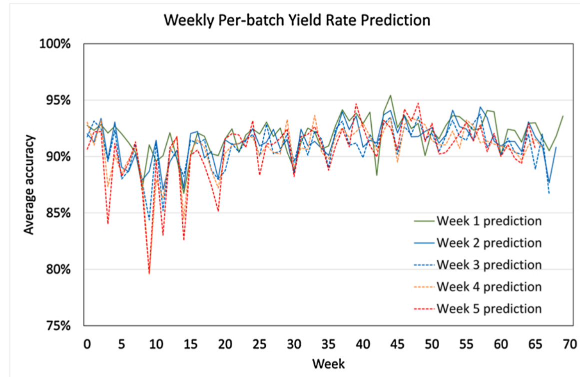

Fig. 6. Average accuracy of each prediction of T-Company data.

the standard deviation. The accuracy distribution of the per-batch prediction could be seen in Table 6.

As seen in Fig. 6, the plots for the five predictions (weeks 1–5) are highly similar, explaining the high similarity of the prediction accuracy values in Fig. 5. Moreover, the prediction accuracy for week 5 is as low as 79 % in some rare cases. Thus, the standard deviation of the prediction for the fifth week is larger. In Subsection 5.3, we will discuss the prediction accuracy changes.

5.2. Simulations of per-machine yield rates

For most manufacturers, including T-company, data regarding actual per-machine yield rates are unavailable because it is impractical to use quality inspection equipment in all production sequences. Therefore, we performed simulations using generated per-machine yield rates and production data. In these simulations, we used production data to estimate the per-machine yield rates. Subsequently, we performed a sensitivity analysis to evaluate the performance of our approach for each manufacturing dataset of a given size. We used several variables in the experiments: the set of machines (50, 250, and 500 machines), the set of batches (500–5000 batches in increments of 500, and a special set of 10 batches at minimum), the set of average of 20 batch steps, and the set of average inspection ratios of 10 % in each batch. The batch production paths and the number of pieces, including defective pieces, were included in each batch dataset. Therefore, the datasets can be classified as follows in terms of size: small (<=1500 batches), medium (>1500 and <=3500 batches), and large (>3500 batches).

5.2.1. Machine and batch generation

At the start of the simulation, the per-machine yield rates were generated; then, the batches were generated based on the machine data. In one set of machine simulations, we used a random normal distribution along with the distribution parameters given in row A in Table 7 to determine the yield rate of each machine. Subsequently, we installed inspection devices on 30 % of the production machines; these devices could be turned off by manufacturers during production for whatever reason.

After obtaining the machine data, we generated batch attributes, namely, the number of batch steps, number of raw pieces, and number of inspections in a batch, as indicated in rows B, C, and D, respectively, in Table 8. After generating the number of steps, we randomly assigned inspection machines to several steps (based on the combinations explained in Subsection 5.2); the remaining steps featured random "normal machines." In this case, a machine could be used more than once to process a particular batch. Then, the sequence of the machine in the batch steps was shuffled, but the last step involved the inspection machines.

Table 6
Distribution of the prediction accuracy using T-company dataset.

Statistics Description	Week 1	Week 2	Week 3	Week 4	Week 5
Minimum	86.73 %	87.08 %	84.39 %	79.53 %	79.65 %
Mean	91.86 %	91.30 %	90.85 %	90.71 %	90.56 %
Maximum	95.42 %	94.40 %	93.75 %	93.65 %	94.73 %
Standard Deviation	1.60 %	1.66 %	1.86 %	2.23 %	2.74 %

Table 7
Distribution parameters in each experiment combination.

#	Description	Min	Max	Mean	Stdev
A	Machine yield rates*	0	1	0.99	0.1
B	Number of raw pieces	1	20,000	2000	1500
C	Number of batch steps**	5	50	20 or 30	7
D	Number of inspection machines in each batch ***	2 machines	batch length	10 % or 30 % of batch length	2
E	Observation accuracy of inspection equipment	85 %	100 %	90 %	2 %

* A includes extra 10 % randomness for the yield of each machine.
** Average of 30 batch steps is only used in Table VIII.
*** Average of 30 % inspection is only used in Table VIII.

Table 8
Average accuracy of the per-machine yield-rate estimation on a manufacturer that uses approximately 500 machines.

Batches	10 % inspection		30 % inspection	
	20 batch steps	30 batch steps	20 batch steps	30 batch steps
10	25.04 %	37.81 %	26.01 %	38.58 %
500	88.85 %	85.19 %	89.43 %	86.84 %
1000	90.15 %	86.17 %	90.87 %	87.72 %
3000	94.06 %	89.31 %	94.09 %	90.42 %
5000	94.19 %	89.64 %	94.34 %	91.36 %

Second, we generated the number of defective pieces in each step. Specifically, we randomized the yield rate of each machine by approximately ±10 % of its original yield rate and calculated the number of defective pieces from the current number of raw pieces and the modified machine yield rate. Subsequently, the remaining good pieces were carried forward to the next batch step.

Third, to simulate a real-world manufacturing environment, the number of defective pieces in each step was hidden and accumulated until the subsequent inspection. The batch steps with the inspection machines were the only steps in which the defective pieces could be observed, as indicated in row E in Table 8. This is supported by studies in which the observation accuracy had a tolerance of 10 %–30 % [39,40], implying that only 70 %–90 % of the accumulated defective pieces were observed; the remaining unobservable defective pieces were accumulated for the next inspection. In this case, the last inspection machine was set to have a 0 % tolerance, and all remaining defective pieces were then observed.

Finally, each combination was used in our approach to estimate the yield rate of each machine. However, because the procedure should be applied 10 times for each combination, the machine and batch data were regenerated. For this reason, we expected to obtain marginally different results in each run, and we used the average as the final result.

5.2.2. Simulation results

Fig. 7 shows the results of our simulation experiments, which demonstrate the good performance of our approach in estimating the yield rates of machines in different simulation combinations. The three lines in Fig. 7 represent simulations with approximately 20 batch steps and 30 % of the inspection machines. A manufacturer that uses 50 machines for production may be required to produce fewer than 200 batches to obtain an accuracy higher than 90 %; a production run with at least approximately 1000 batches may be required to obtain the best accuracy, as indicated by point A in Fig. 7. However, when a manufacturer uses approximately 250 machines for production, the manufacturer may be required to produce at least approximately 1500 batches to obtain the best accuracy, as indicated by point B in Fig. 7. Based on the operational chart point D in Fig. 7, which is similar to the data model of T-company, our approach could provide a good estimate of the per-machine yield rates with an average accuracy of 91.84 %. Meanwhile, a manufacturer with 500 machines and an average of 20 batch steps in its production may be required to produce at least approximately 1000 batches to obtain an accuracy higher than 90 %, whereas at least approximately 3000 batches to obtain the best accuracy, as indicated by point C in Fig. 7. The accuracy distribution of the per-machine yield-rate estimation on a manufacturer that uses approximately 20 batch steps and 10 % inspections, could be seen in Appendix B.

To better understand the performance of the proposed method, we performed extra simulations with different batch lengths (20 and 30) and inspection ratios (10 % and 30 % of batch steps in a batch). Based on the simulation result on a manufacturer that uses up to 250 machines for production, the average accuracy of per-machine yield rates is less affected by the average number of batch steps and inspection machines. However, when a manufacturer uses a large number of machines (approx. 500 machines), these two factors greatly affect the average accuracy of per-machine yield-rate estimation shown in Table 8. The simulation result shows that, a manufacturer that uses approximately 500 machines with an average of 30 batch steps leads to a decrease in accuracy compared to the one with an average of 20 batch steps. However, by using approximately 30 % of the inspection machines in each batch, accuracy increases compared with that when only approximately 10 % of the inspection machines in each batch are used, as shown in Table 8. The accuracy distribution of the per-machine yield-rate estimation on a manufacturer that uses approximately 500 machines, could be seen in Appendix C.

5.2.3. Validation

To further validate the simulation results in Fig. 7, the estimated per-machine yield rates were used to predict the per-batch yield rates. The validation is based on the first assumption in section 3, that the actual per-machine yield rates do not change next week. Therefore, we generated 5 validation datasets for each simulation combination in subsection 5.2.1.

The results in Fig. 8 shows that, most simulations with only 10 batches dataset give poor prediction accuracy (close to 0 %). This is very reasonable considering that many machines do not have the estimated yield rates. Meanwhile, a manufacturer that uses 50 machines for production may require a dataset with at least 100 batches to obtain a good accuracy of per-batch yield rates prediction (>90 %). Whereas, a manufacturer that uses approximately 250 machines for production may require a dataset with at least 500 batches to obtain a good prediction accuracy (>90 %). Furthermore, point A in Fig. 8 also indicates the simulation result using a dataset pattern similar to T-Company provide a good per-batch yield-rate prediction, which has average accuracy of around 91.29 %. This confirms the prediction results with real-world data of T-company shown in Fig. 5. However, a manufacturer that uses approximately 500 machines may require a dataset with at least 1000 batches to obtain a good prediction accuracy (>90 %). The distribution of the prediction accuracy using the simulation dataset could be seen in Appendix D.

To see the performance of the proposed prediction method on different lengths of batch steps, we analyzed the simulation results of approximately 20 batch steps for the per-machine yield-rate estimation. The details of the analytical results of our simulation could be seen in Table 9. The dataset is divided into two groups, the simple batch group (<20 batch steps) and the complex batch group (>20 batch steps). The data for 20 batch steps is excluded in this analysis. Based on the analytical results, the complex batch group has a larger relative error and a smaller accuracy. In other words, the proposed method performs better on the simple group. Meanwhile, our prediction tends to underestimate the future batch yield rate in both groups, as seen in Table 9. Similar results could be seen in a similar simulation, which has the same settings as in Table 7 except that the randomness of the per-machine yield rate to be 1 %. The results confirm that the proposed method performs better on the simple group. The difference of this simulation result is that, the prediction accuracy could reach close to 99 % with 1 % randomness, while it could only reach 89 %–92 % with 10 % randomness. The details of the simulation result with 1 % randomness could be seen in Appendix E.

5.3. Discussions

According to the experiments described in Section 5.1, our approach accurately predicted the per-batch yield rates. In the case of T-company,

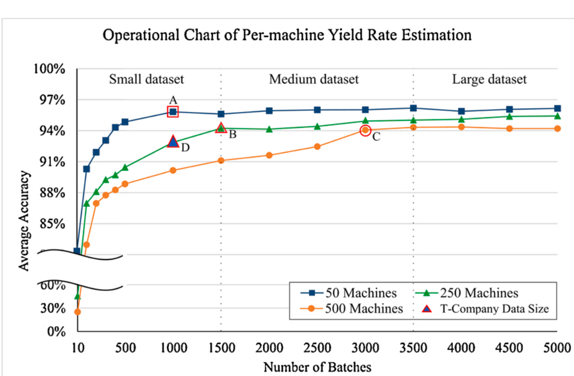

Fig. 7. Operational chart of the proposed method to estimating the per-machine yield rates.

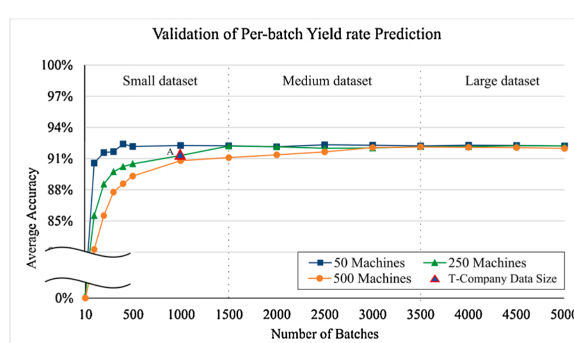

Fig. 8. Validation of the proposed method to predicting the per-batch yield rates.

Table 9
Average simulated batch yield-rate and predicted batch yield rate regarding batch complexity in A manufacturer that uses approximately 50 machines.

Batches	< 20 batch steps					> 20 batch steps				
	Actual YR	Pred YR	Diff*	Rel Error**	Acc***	Actual YR	Pred YR	Diff*	Rel Error**	Acc***
10	0.5185	0.4699	0.0486	9.37 %	90.63 %	0.3551	0.2840	0.0711	20.03 %	79.97 %
500	0.5306	0.4858	0.0448	8.45 %	91.55 %	0.3645	0.3269	0.0376	10.32 %	89.68 %
1000	0.5435	0.5008	0.0427	7.87 %	92.13 %	0.3812	0.3453	0.0359	9.42 %	90.58 %
3000	0.5622	0.5214	0.0408	7.26 %	92.74 %	0.4013	0.3663	0.0350	8.72 %	91.28 %
5000	0.5470	0.5035	0.0435	7.96 %	92.04 %	0.3814	0.3447	0.0367	9.63 %	90.37 %

[*] Differences = Actual Yield Rate – Predicted Yield Rate.
[**] Relative error = Differences / Actual Yield Rate.
[***] Accuracy = 1 - Relative error.

our approach could generate predictions for periods longer than one month. However, we suggest that predictions should be made for only one to two weeks (or periods) ahead. This is because the standard deviation increases when predictions are made for periods further into the future. This standard deviation increase may be due to the following reasons. First, some machines used in future weeks may not be used in the week in which the per-machine yield rates were estimated. Hence, fewer batches can be predicted. Second, the per-machine yield rates may be affected by wear and tear in the following weeks. Therefore, the actual per-machine yield rates may be lower than those estimated for a certain day. Conversely, some machines may be maintained in the following week, thus increasing their actual yield rate. Third, in rare cases, our approach predicts the per-batch yield rates at accuracy as low as 79 %. In this study, we observed that this was caused by the processing of a few batches in the week for which the per-machine yield rates were estimated, resulting in less accurate estimates of per-machine yield rates. This observation is supported by Fig. 9, in which the fluctuation of the plot of the number of batches is similar to the plot of the week-5 prediction in Fig. 6.

To further understand whether the proposed method is effective for different batch production patterns, we conducted simulations to measure the average accuracy of the estimated per-machine yield rates. Then, the simulation of per-machine yield rates could be used as a benchmark for the proposed method. The prediction results of T-company's cases were good because the simulation indicated that the estimated average per-machine yield rates were very close to the average actual per-machine yield rates, where the estimation error was approximately 8 %. However, our simulation indicated that the randomness of the dataset increased when more machines were used in production. This makes the proposed method less accurate. In this case, we suggest that more batches be used to obtain accurate estimates, but this could be difficult in practice. Moreover, with the same number of machines, long sequences of each batch (longer batch steps) reduce the accuracy of our approach. This is due to the effect of the principle of likelihood on our E-step, as explained in Subsection 4.4, which may lead to underestimation for a greater number of machines when defective pieces are observed.

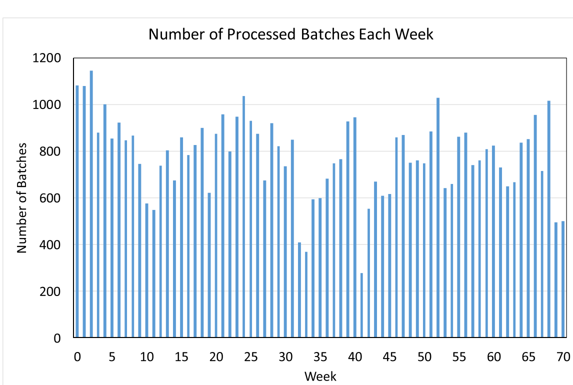

Fig. 9. Number of batches for each week in T-Company data.

The accuracy reduction caused by a long sequence can be alleviated by increasing the number of batches and the density of inspection stations; however, this increases production costs. Therefore, managers must balance estimation accuracy (by having more inspection stations) and production cost [41]. Although the number of machines, the number of batches, and the average batch steps are the most important variables in our approach, the results indicate that the density of inspection stations should increase as the average number of batch steps increases.

Based on the simulation results presented in Subsection 5.2.2, our approach has a rather low accuracy for small datasets (small number of batches). However, it is highly accurate for medium and larger datasets (medium and large number of batches). In addition, based on the simulation results, our approach is effective for small- and medium-scale manufacturers, which use 500 or fewer machines and have an average batch sequence with 30 or fewer steps.

In subsection 5.2.3, per-batch yield-rate prediction is used to validate the simulation result of per-machine yield-rate estimation. It shows that a very small dataset gives poor prediction accuracy. It is due to that the dataset is insufficient to be able to estimate the per-machine yield rates. Therefore, the estimation result can hardly be used to predict the per-batch yield rates. It is reasonable that the prediction accuracy increases as the number of batches in the dataset increases. However, based on the simulation results in Fig. 8, the highest accuracy that could be reached when predicting per-batch yield rates is approximately 92 %. This confirms the prediction result of per-batch yield rates for week 1 shown in Fig. 5.

Furthermore, there are practical considerations for the proposed method. First, from an economic perspective, using the proposed method to predict the per-batch yield rates, manufacturers can plan their production in general, including their production costs. Second, from a technical perspective, our approach uses simple production data that require fewer parameters. Accordingly, most manufacturers can easily implement the proposed method with fewer resources.

6. Conclusions and directions for future research

The proposed method is a useful solution for predicting the per-batch yield rates based on per-machine yield-rate estimations. In our experiments, we used one-week data from T-company to build the prediction model, then used the model to predict the pre-batch yield rate in the five subsequent weeks. We validated the proposed prediction approach with a total of 70 weeks of data from T-company. The average accuracy of our approach is 91.86 % for the subsequent week, and consistently over 90 % for five consecutive weeks (over one month). However, it is suggested that the per-batch yield rates be predicted for only one or two weeks. To determine whether the proposed method is effective for different batch production patterns, we used simulations. It was demonstrated that by using a generated batch production pattern similar to that of T-company, our approach could provide a good estimation of the per-machine yield rates. As a result, the average prediction accuracy of the per-batch yield rates could reach 91.29 %. Thus, the simulation results confirmed the prediction accuracy of the per-batch yield rates using a real-world dataset.

The number of machines, number of batches, and average number of batch steps are the most important variables for estimating the per-machine yield rates. The simulation results indicated that, the minimum number of batches per week should be 100, 500, and 1000 batches for the manufacturers with approx. 50, 250, and 500 machines, respectively. In addition, our approach is appropriate for any manufacturer with 500 or fewer machines with an average of 30 or shorter batch sequences. The proposed method is reasonable and could be used by manufacturers to predict per-batch yield rates. Thus, manufacturers can better plan their batch production and control risk. In addition, our lightweight approach uses only production data without an excessive number of parameters. Therefore, it may reduce management overhead, as it requires fewer resources and is simpler to use. Hence, manufacturers with limited resources can easily implement this approach.

However, our approach has several limitations related to batch specifications and single/continuous production systems (or the manufacturing of a similar type). First, our approach estimates the per-machine yield rates based on the station sequence. Although this is accurate, it may provide unexpected results for fixed sequences of machine stations because the machines at the beginning of the sequence are necessarily considered suspects. Hence, their yield rates may be lower than those of the machines in the final sequence. This implies that unexpected results may be obtained in flow-shop manufacturing (or similar). Second, based on the design of our approach, any suspect machine is assigned a share of defective pieces based on its current yield rate, regardless of its step or station number. These shares of defective pieces in each batch are summed into a single value for each machine. For this reason, our approach requires the processing of production data based on batch specifications. We found that the machines used to process different batch specifications may provide different batch yield rates. Finally, this study opens the way to further investigate per-batch yield rates so that these two limitations may be addressed in future research.

Data availability

The data that has been used is confidential.

Declaration of Competing Interest

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

Acknowledgements

We would like to thank the anonymous reviewers for their insightful comments that helped improve this paper. This study was supported by the Ministry of Science and Technology of Taiwan, under Grants 105-2634-F-008-001, 106-2634-F-008-001-CC2, 108-2221-E-008-032-MY3, and 108-2221-E-008-030-MY3.

Appendix

A Equations list

Equation	Eq. Number
The simple version of the likelihood	1
$P(Y,Z\|\theta) = \prod_{i=1}^{I}\prod_{n=1}^{N_i}\left(F_D(i,n)^{y_{in}} \times F_G(i,n)^{1-y_{in}}\right)$	
where	
$F_D(i,n) = \prod_{k=1}^{J_{in}}\left((1-P_{ik})\prod_{s=1}^{k-1}P_{is}\right)^{z_{ink}}$	2
$F_G(i,n) = \prod_{k=1}^{J_{in}}P_{ik}$	3
Complete Likelihood Function	4
$P(Y,Z\|\theta) = \prod_{i=1}^{I}\prod_{n=1}^{N_i}\left(\left(\prod_{k=1}^{J_{in}}\left((1-P_{ik})\prod_{s=1}^{k-1}P_{is}\right)^{z_{ink}}\right)^{y_{in}} \times \left(\prod_{k=1}^{J_{in}}P_{ik}\right)^{1-y_{in}}\right)$	
Constrained Least square	5
$\arg\min_{\theta}\sum_{i=1}^{I}\left(\sum_{k=1}^{C_i}q_{ik} - \ln l_i\right)^2$ subject to $q_{ik} \le 0$	
Updating per-machine yield rates	6
$Pr(m) = \frac{g_m}{g_m + d_m}$	
Likelihood of kth manufacturing steps generates defective pieces	7
$P(z_{ink}=1\|y_{in},\theta) = \begin{cases} 0; & \text{if} y_{in}=0 \\ \frac{(1-P_{ik})\left(\prod_{s=1}^{k-1}P_{is}\right)}{\sum_{t=1}^{J_{in}}\left((1-P_{it})\left(\prod_{u=1}^{t-1}P_{iu}\right)\right)}; & \text{if} y_{in}=1 \end{cases}$	
$E[z_{ink}] = 0*P(z_{ink}=0\|y_{in},\theta) + 1*P(z_{ink}=1\|y_{in},\theta)$ $E[z_{ink}] = P(z_{ink}=1\|y_{in},\theta)$	8
Estimation number of the defective pieces	9
$e_{ijk} = E[z_{ink}] \times b_{ij}$	
Set of *(i,j,k)* indexes	10
$S_m = \{(i,j,k)\|\forall m_{ik} = m\}$	
	11

(*continued on next page*)

(*continued*)

Equation	Eq. Number
Total expectation number of defective pieces $d_m = \sum_{(i,j,k)\in S_m} e_{ijk}$	
The number of potential defective pieces $h_{ijk} = \begin{cases} 0; & \text{if} j = k \\ h_{ij;k+1} + e_{ijk}; & \text{if} j > k \end{cases}$	12
The total number of potential defective pieces $x_m = \sum_{(i,j,k)\in S_m} h_{ijk}$	13
The total expected number of good pieces $g_m = x_m + \left(\sum_{i=1}^{I} f_{im} \bullet r_{im}\right)$	14
The average accuracy of per-batch yield prediction $\overline{acc} = \frac{\sum_{i=1}^{I}\left(1 - \left|\left(\prod_{j=1}^{C_i} P_{ij}\right) - l_i\right|\right)}{I}$	15

B Accuracy distribution of the machine yield-rate estimation (20 batch steps and 10 % inspections)

Batches	50 Machines				250 Machines				500 Machines			
	Min	Avg	Max	Stdev	Min	Avg	Max	Stdev	Min	Avg	Max	Stdev
10	0.00 %	82.34 %	99.98 %	23.64 %	0.00 %	45.54 %	100.00 %	45.02 %	0.00 %	25.04 %	100.00 %	40.39 %
500	76.43 %	94.85 %	100.00 %	4.02 %	55.63 %	90.44 %	100.00 %	7.68 %	36.80 %	88.85 %	100.00 %	9.62 %
1000	82.64 %	95.83 %	99.99 %	3.49 %	62.51 %	91.84 %	100.00 %	6.12 %	53.55 %	90.15 %	100.00 %	7.77 %
3000	82.07 %	96.02 %	100.00 %	3.80 %	69.89 %	94.14 %	100.00 %	4.71 %	71.94 %	94.06 %	100.00 %	4.78 %
5000	86.02 %	96.16 %	99.99 %	3.86 %	79.78 %	95.41 %	100.00 %	3.85 %	73.41 %	94.19 %	100.00 %	4.60 %

C. Accuracy distribution of the per-machine yield-rate estimation (500 machines)

Batches	20 batch steps				30 batch steps			
	Min	Avg	Max	Stdev	Min	Avg	Max	Stdev
				10 % inspection				
10	0.00 %	25.04 %	100.00 %	40.39 %	0.00 %	37.81 %	100.00 %	44.67 %
500	36.80 %	88.85 %	100.00 %	9.62 %	16.87 %	85.19 %	100.00 %	14.40 %
1000	53.55 %	90.15 %	100.00 %	7.77 %	35.42 %	86.17 %	100.00 %	12.49 %
3000	71.94 %	94.06 %	100.00 %	4.78 %	54.26 %	89.31 %	100.00 %	10.31 %
5000	73.41 %	94.19 %	100.00 %	4.60 %	54.49 %	89.64 %	100.00 %	10.72 %
				30 % inspection				
10	0.00 %	26.01 %	99.99 %	40.74 %	0.00 %	38.58 %	99.99 %	44.87 %
500	39.30 %	89.43 %	100.00 %	8.86 %	36.88 %	86.84 %	99.99 %	11.66 %
1000	57.75 %	90.87 %	100.00 %	7.06 %	48.42 %	87.72 %	100.00 %	9.79 %
3000	76.53 %	94.09 %	100.00 %	4.35 %	63.27 %	90.42 %	100.00 %	7.61 %
5000	63.63 %	94.34 %	100.00 %	4.20 %	67.70 %	91.36 %	100.00 %	7.13 %

D. Distribution of the per-batch yield rate prediction accuracy using simulation dataset

Batches	50 Machines				250 Machines				500 Machines			
	Min	Avg	Max	Stdev	Min	Avg	Max	Stdev	Min	Avg	Max	Stdev
10	0.00 %	77.41 %	94.41 %	11.24 %	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
500	91.65 %	92.17 %	93.02 %	0.31 %	89.28 %	90.49 %	91.48 %	0.39 %	87.99 %	89.31 %	90.70 %	0.40 %
1000	91.55 %	92.25 %	93.00 %	0.14 %	90.54 %	91.29 %	92.08 %	0.19 %	89.94 %	90.79 %	91.66 %	0.25 %
3000	91.72 %	92.28 %	92.70 %	0.09 %	91.31 %	92.01 %	92.40 %	0.13 %	91.69 %	92.06 %	92.40 %	0.11 %
5000	91.88 %	92.22 %	92.50 %	0.07 %	91.80 %	92.21 %	92.51 %	0.08 %	91.63 %	91.97 %	92.23 %	0.09 %

* The value of "n/a" means that the dataset for building model (machine yield rate estimation) are insufficient for batch yield-rate prediction.

E Average simulated batch yield-rate and predicted batch yield rate regarding batch complexity

This simulation results have the same settings as in Table 7 except that the randomness of the per-machine yield rate to be 1 %. The actual yield rate is generated in the simulation as the ground truth. Then, the values in the columns of actual or predicted yield rate are the average of the yield rate in one-week data.

Batches	< 20 batch steps					> 20 batch steps				
	Actual YR	Pred YR	Diff*	Rel Error**	Acc***	Actual YR	Pred YR	Diff*	Rel Error**	Acc***
A manufacturer that uses approximately 50 machines										
10	0.8071	0.7721	0.0350	4.34 %	95.66 %	0.7151	0.6221	0.0930	13.01 %	86.99 %
500	0.8061	0.7991	0.0070	0.86 %	99.14 %	0.7133	0.7055	0.0077	1.09 %	98.91 %
1000	0.8101	0.8036	0.0064	0.79 %	99.21 %	0.7191	0.7106	0.0085	1.18 %	98.82 %
3000	0.8015	0.7940	0.0075	0.94 %	99.06 %	0.7061	0.6970	0.0092	1.30 %	98.70 %
5000	0.8175	0.8111	0.0064	0.78 %	99.22 %	0.7281	0.7199	0.0082	1.12 %	98.88 %
A manufacturer that uses approximately 250 machines										
10	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
500	0.8109	0.7996	0.0113	1.40 %	98.60 %	0.7187	0.7179	0.0008	0.11 %	99.89 %
1000	0.8060	0.7956	0.0104	1.29 %	98.71 %	0.7103	0.7076	0.0027	0.38 %	99.62 %
3000	0.8055	0.7972	0.0083	1.03 %	98.97 %	0.7110	0.7032	0.0079	1.11 %	98.89 %
5000	0.8125	0.8052	0.0074	0.91 %	99.09 %	0.7212	0.7148	0.0063	0.87 %	99.13 %
A manufacturer that uses approximately 500 machines										
10	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
500	0.8071	0.7977	0.0094	1.17 %	98.83 %	0.7134	0.7106	0.0028	0.40 %	99.60 %
1000	0.8067	0.7970	0.0097	1.20 %	98.80 %	0.7145	0.7131	0.0015	0.21 %	99.79 %
3000	0.8050	0.7980	0.0070	0.87 %	99.13 %	0.7112	0.7037	0.0075	1.06 %	98.94 %
5000	0.8115	0.8042	0.0073	0.90 %	99.10 %	0.7197	0.7123	0.0074	1.02 %	98.98 %

*Differences = Actual Yield Rate – Predicted Yield Rate.
**Relative error = Differences / Actual Yield Rate.
*** Accuracy = 1 - Relative error.

References

[1] Bortolini M, Galizia FG, Mora C. Reconfigurable manufacturing systems: literature review and research trend. J Manuf Syst 2018;49:93–106. https://doi.org/10.1016/j.jmsy.2018.09.005.

[2] Lu Y, Xu X, Wang L. Smart manufacturing process and system automation – a critical review of the standards and envisioned scenarios. J Manuf Syst 2020;56:312–25. https://doi.org/10.1016/j.jmsy.2020.06.010.

[3] Da Rold C, Karamouzis F. Enhance your IT agility and grow the business by optimizing the three layers of adaptive sourcing strategy. Gartner, Inc. Research; 2014.

[4] Jiao Jianxin, Tseng MM, Ma Qinhai, Zou Yi. Generic bill-of-Materials-and-Operations for high-variety production management. Concurr Eng 2000;8:297–321. https://doi.org/10.1177/1063293X0000800404.

[5] Malik AI, Sarkar B. Disruption management in a constrained multi-product imperfect production system. J Manuf Syst 2020;56:227–40. https://doi.org/10.1016/j.jmsy.2020.05.015.

[6] Li T-S, Huang C-L, Wu Z-Y. Data mining using genetic programming for construction of a semiconductor manufacturing yield rate prediction system. J Intell Manuf 2006;17:355–61. https://doi.org/10.1007/s10845-005-0008-7.

[7] Joseph VR, Adya H. Prediction of yield in a multiproduct batch production environment. Qual Eng 2002;14:153–9. https://doi.org/10.1081/QEN-100106894.

[8] Bae SJ, Hwang JY, Kuo W. Yield prediction via spatial modeling of clustered defect counts across a wafer map. IIE Trans 2007;39:1073–83. https://doi.org/10.1080/07408170701275335.

[9] Tao F, Qi Q, Liu A, Kusiak A. Data-driven smart manufacturing. J Manuf Syst 2018;48:157–69. https://doi.org/10.1016/j.jmsy.2018.01.006.

[10] Keyvanloo M, Kimiagari AM, Esfahanipour A. A hybrid approach to select the best sourcing policy using stochastic programming. J Manuf Syst 2015;36:115–27. https://doi.org/10.1016/j.jmsy.2014.08.003.

[11] Mohan RR, Thiruppath K, Venkatrama R, Raghuraman S. Quality improvement through first pass yield using statistical process control approach. J Appl Sci 2012;12:985–91. https://doi.org/10.3923/jas.2012.985.991.

[12] Cawley P. Structural health monitoring: closing the gap between research and industrial deployment. Struct Heal Monit 2018;17:1225–44. https://doi.org/10.1177/1475921717750047.

[13] Patel S, Dale BG, Shaw P. Set-up time reduction and mistake proofing methods: an examination in precision component manufacturing. TQM Mag 2001;13:175–9. https://doi.org/10.1108/09544780110385528.

[14] Mehmeti X, Mehmeti B, Sejdiu R. The equipment maintenance management in manufacturing enterprises. IFAC-PapersOnLine 2018;51:800–2. https://doi.org/10.1016/j.ifacol.2018.11.192.

[15] Chincholkar M, Herrmann JW. Estimating manufacturing cycle time and throughput in flow shops with process drift and inspection. Int J Prod Res 2008;46:7057–72. https://doi.org/10.1080/00207540701513893.

[16] Jun J, Chang T-W, Jun S. Quality prediction and yield improvement in process manufacturing based on data analytics. Processes 2020;8:1068. https://doi.org/10.3390/pr8091068.

[17] Kim D, Kim M, Kim W. Wafer edge yield prediction using a combined long short-term memory and feed- forward neural network model for semiconductor manufacturing. IEEE Access 2020;8:215125–32. https://doi.org/10.1109/ACCESS.2020.3040426.

[18] Stich P, Wahl M, Czerner P, Weber C, Fathi M. Yield prediction in semiconductor manufacturing using an AI-based cascading classification system. In: 2020 IEEE Int. Conf. Electro Inf. Technol.; 2020. p. 609–14. https://doi.org/10.1109/EIT48999.2020.9208250.

[19] Chen T, Chiu M-C. An interval fuzzy number-based fuzzy collaborative forecasting approach for DRAM yield forecasting. Complex Intell Syst 2020. https://doi.org/10.1007/s40747-020-00179-8.

[20] Chen T-CT, Wu H-C. Forecasting the unit cost of a DRAM product using a layered partial-consensus fuzzy collaborative forecasting approach. Complex Intell Syst 2020;6:479–92. https://doi.org/10.1007/s40747-020-00146-3.

[21] Chen Tory, Wang M-JJ. A fuzzy set approach for yield learning modeling in wafer manufacturing. IEEE Trans Semicond Manuf 1999;12:252–8. https://doi.org/10.1109/66.762883.

[22] Wang J, Ma Y, Zhang L, Gao RX, Wu D. Deep learning for smart manufacturing: methods and applications. J Manuf Syst 2018;48:144–56. https://doi.org/10.1016/j.jmsy.2018.01.003.

[23] Wang P, Ananya, Yan R, Gao RX. Virtualization and deep recognition for system fault classification. J Manuf Syst 2017;44:310–6. https://doi.org/10.1016/j.jmsy.2017.04.012.

[24] Zhang Y, You D, Gao X, Zhang N, Gao PP. Welding defects detection based on deep learning with multiple optical sensors during disk laser welding of thick plates. J Manuf Syst 2019;51:87–94. https://doi.org/10.1016/j.jmsy.2019.02.004.

[25] Tirkel I, Rabinowitz G, Price D, Sutherland D. Wafer fabrication yield learning and cost analysis based on in-line inspection. Int J Prod Res 2016;54:3578–90. https://doi.org/10.1080/00207543.2015.1106609.

[26] Bishop CM. Pattern recognition and machine learning. New York: Springer; 2006.

[27] Lange K. A gradient algorithm locally equivalent to the em algorithm. J R Stat Soc Ser B 1995;57:425–37. https://doi.org/10.1111/j.2517-6161.1995.tb02037.x.

[28] Dwivedi R, Ho N, Khamaru K, Wainwright MJ, Jordan MI, Yu B. Singularity, misspecification and the convergence rate of EM. Ann Stat 2020;48. https://doi.org/10.1214/19-AOS1924.

[29] Goli A, Babaee Tirkolaee E, Soltani M. A robust just-in-time flow shop scheduling problem with outsourcing option on subcontractors. Prod Manuf Res 2019;7:294–315. https://doi.org/10.1080/21693277.2019.1620651.

[30] Thompson NC, Greenewald K, Lee K, Manso GF. The computational limits of deep learning. 2020. http://arxiv.org/abs/2007.05558.

[31] Qu YJ, Ming XG, Liu ZW, Zhang XY, Hou ZT. Smart manufacturing systems: state of the art and future trends. Int J Adv Manuf Technol 2019;103:3751–68. https://doi.org/10.1007/s00170-019-03754-7.

[32] Biernacki C, Celeux G, Govaert G. Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. Comput Stat Data Anal 2003;41:561–75. https://doi.org/10.1016/S0167-9473(02)00163-9.

[33] Xiang W, Karfoul A, Yang C, Shu H, Le Bouquin Jeannès R. An exact line search scheme to accelerate the EM algorithm: application to Gaussian mixture models identification. J Comput Sci 2020;41:101073. https://doi.org/10.1016/j.jocs.2019.101073.

[34] Shireman E, Steinley D, Brusco MJ. Examining the effect of initialization strategies on the performance of Gaussian mixture modeling. Behav Res Methods 2017;49: 282–93. https://doi.org/10.3758/s13428-015-0697-6.

[35] Zhou H, Weinberg CR. Modeling conception as an aggregated bernoulli outcome with latent variables via the EM algorithm. Biometrics 1996;52:945. https://doi.org/10.2307/2533055.

[36] Nookabadi AS, Middle JE. An integrated quality assurance information system for the design-to-order manufacturing environment. TQM Mag. 2006;18:174–89. https://doi.org/10.1108/09544780610647883.

[37] Soares JC, Sousa S, Tereso A. Industry practices on the rework of defective products: survey results. TQM J. 2020;32:1177–96. https://doi.org/10.1108/TQM-06-2019-0162.

[38] Ouzineb M, Mhada FZ, Pellerin R, El Hallaoui I. Optimal planning of buffer sizes and inspection station positions. Prod Manuf Res 2018;6:90–112. https://doi.org/10.1080/21693277.2017.1422812.

[39] See JE. Visual inspection reliability for precision manufactured parts. Hum Factors J Hum Factors Ergon Soc 2015;57:1427–42. https://doi.org/10.1177/0018720815602389.

[40] Beckert SF, Paim WS. Critical analysis of the acceptance criteria used in measurement systems evaluation. Int J Metrol Qual Eng 2017;8:23. https://doi.org/10.1051/ijmqe/2017016.

[41] Farooq MA, Kirchain R, Novoa H, Araujo A. Cost of quality: evaluating cost-quality trade-offs for inspection strategies of manufacturing processes. Int J Prod Econ 2017;188:156–66. https://doi.org/10.1016/j.ijpe.2017.03.019.

Philip F. E. Adipraja received his BA degree from the State University of Malang, Indonesia in 2011. Then in 2015, he received his M.Comp.Sc. degree from Institut Teknologi Sepuluh Nopember (ITS), Indonesia. Currently, he is working towards his Ph.D. degree in the Department of Computer Science and Information Engineering, National Central University, Taiwan. His research interests include smart manufacturing, software engineering, and intelligent transport system.

Chin-Chun Chang received the B.S. degree and the M.S. degree in computer science in 1989 and 1991, respectively, and the Ph.D. degree in computer science in 2000, all from National Chiao Tung University, Hsinchu, Taiwan. He is currently an Associate Professor in the Department of Computer Science and Engineering, National Taiwan Ocean University, Keelung, Taiwan. His research interests include computer vision, machine learning, and pattern recognition.

Wei-Jen Wang received his Ph.D. degree in Computer Science from Rensselaer Polytechnic Institute (RPI) in the United States in December 2006. He is currently an Associate Professor in the Department of Computer Science and Information Engineering, National Central University, Taiwan. His research interests include cloud computing, edge computing, Industrial IoT systems, distributed programming, high availability, and fault tolerance.

Deron Liang received his Ph.D. degree in Computer Science from the University of Maryland at College Park in the United States. He is currently a Professor and Director for Software Research Center at National Central University, Taiwan. His research interests include intelligent systems software, smart manufacturing, data analysis, software engineering, and information system security.