Prediction of Per-Batch Yield Rates in Production Based on Maximum Likelihood Estimation of Per-Machine Yield Rate

Philip F. E. Adipraja[1][0000-0002-1289-166X], Chin-Chun Chang[2] [0000-0002-4566-8683], Wei-Jen Wang[*1][0000-0002-7036-3101], Deron Liang[1][0000-0001-6538-7128]

[1] Department of Computer Science and Information Engineering, National Central University, Taoyuan City 32001, Taiwan.

[2] Department of Computer Science and Engineering, National Taiwan Ocean University, Keelung City 20224, Taiwan.

*corresponding author: wjwang@csie.ncu.edu.tw

Abstract

The demand for high-quality customized products compels manufacturers to adopt batch production. To better plan their production cost, most manufacturers must predict their production yield rate. However, existing approaches to predicting the yield rate of batch production systems are either missing or require prohibitively large data sets and long computational times. Based on first pass yield, one approach to predicting the per-batch yield rate involves using the per-machine yield rate in a given production run. However, for most manufacturers, the actual per-machine yield rates are unknown and might be affected by multiple factors. Therefore, we propose an approach to predict the per-batch yield rate based on an estimated per-machine yield rate. By using data from a so-called T-company, the proposed approach could yield next-week predictions of per-batch yield rates at an average accuracy of 91.86% and could do so for 5 consecutive weeks with an average accuracy of more than 90%. Additionally, to validate our method, we conducted simulations to generate per-machine yield rates and batch data of sizes similar to T-company data. The average accuracy of the estimated per-machine yield rates was 92.06%.

Keywords: batch yield rate prediction, EM algorithm, machine yield rate estimation, manufacturing process

This is the preprint or submission version before being reviewed. The final or published version could be found in https://doi.org/10.1016/j.jmsy.2021.11.015

1. Introduction

The dynamic nature of market demand compels most manufacturers to offer high-quality customized products [1], and those products are usually manufactured in batches [2]. Production planning for batch systems requires considerable effort because it involves numerous products, production stations, and machines. In this context, it is crucial for manufacturers to have advance knowledge of the yield rates of their products so that they can better plan their production costs. Furthermore, it allows manufacturers to adjust their parameters and estimate and evaluate their production costs [3–5].

Many approaches to predicting yield rate in manufacturing use either *macro yield modeling* or *micro yield modeling* [6]. However, research on the application of these approaches to batch production systems has not been conducted. Furthermore, previously developed approaches are potentially unsuitable because they require prohibitively large data sets and long computational times. Although manufacturers can plan which machines to use in the production sequence of future product batches, doing so based on first pass yield (FPY) [7], one simple approach to predicting the per-batch production yield rate involves using the actual yield rates of all machines associated with the production of a given batch. In this approach, the actual per-machine yield rate can be calculated based on the number of observed defective products generated by a particular machine, and the defective products can only be observed using quality inspection devices. However, it is expensive to use quality inspection devices for all production machines [8]. Consequently, manufacturers may reduce the number of inspection devices if possible and attempt to balance between the number of inspection devices and the ability to control production quality [9]. Therefore, in practice, it is impossible to determine the actual per-machine yield rate given a limited number of inspection devices.

Most manufacturers find it difficult to estimate the actual per-machine yield rate because it is affected by multiple factors, such as production process drift, the environment, machine condition, machine misconfiguration, and machine age [9–12]. To solve this problem, in this study, we propose a simple micro-yield-modeling approach to predict the per-batch production yield rate by using an estimated per-machine yield rate. In this case, we use the maximum likelihood method to estimate the per-machine yield rate in a production process. The proposed approach requires relatively simple information, such as the production paths of products, number of raw products, and number of defective products detected by each inspection device to estimate the per-machine yield rate. Thereafter, to predict a batch yield rate, the proposed approach only requires the production paths of the products in the batch.

The major challenge associated with the proposed approach is that the initial machine yield rate is unknown. Many options to this problem are available, including the expectation-maximization (EM) algorithm and the Newton–Raphson method [13]. Although iterations converge faster in the Newton–Raphson method, its computational cost is higher. In addition, Springer and Urban [14] compared the EM algorithm to other alternatives and demonstrated that the EM algorithm is faster and incurs no significant overhead. Therefore, our proposed approach uses the EM algorithm, a widely used algorithm that demonstrably leads to convergence [15].

In this study, we aimed to predict the production yield rate by estimating the per-machine yield rate. To validate this approach to production yield rate prediction, we used time-series data from a so-called T-company. The data covered up to 70 weeks of real-world production that were collected using sensors and by human operators; a week of data covered thousands of batches of products and more than 200 machines in each week's data. Our approach can predict the subsequent week's production yield rate at an average accuracy of 91.86% and can predict the subsequent 5 weeks' production yield rate at an average accuracy of >90%. However, because actual data on machine yield rate are limited and difficult to obtain, we performed simulations to validate the performance of the proposed approach by comparing the setup per-machine yield rate and the estimated per-machine yield rate. In the simulations, we used different production sequence lengths, number of batches, number of production machines, and ratio of inspection machines. Subsequently, we calculated how close the estimated per-machine yield rate could be to the actual per-machine yield rate. The results indicated that by using a data size similar to that of the actual data set for T-company, our approach provided good estimates at an average accuracy of 92.06%.

The remainder of this paper is organized as follows. Section 2 introduces the existing approaches to production yield rate prediction. Section 3 describes the maximum likelihood estimation of the per-machine yield rates based on the observation of defective products. Section 4 details our proposed approach to predicting the production yield rate based on the estimated per-machine yield rates. Section 5 describes the experiment and simulation design and results and discusses the implications of our approach. Finally, the last section concludes the paper and proposes directions for future research.

2. Related Works

Different manufacturers may implement different approaches to predict the production yield rate. In practice, a manufacturer could choose from several approaches that are either based on *macro yield modeling* or *micro yield modeling* [6]. For semiconductor manufacturers, the macro yield modeling approach considers only large *a priori* factors, whereas the micro-yield-modeling approach considers detailed information on different classes of defect categories, layouts, and process variations of circuit design. However, many of the approaches used in this area tend to predict a single production yield rate [16–18]. Therefore, when applied to batch production systems, the prediction accuracy of these approaches could be poor.

Some existing approaches utilize time-series data as the input to predict the production yield rate. Chen and Chiu [18] proposed an approach based on time-series production data that uses an interval fuzzy number-based fuzzy collaborative forecasting (IFN-based FCF) scheme to predict the production yield rate. Their approach performed well at a mean absolute percentage error (MAPE) of <2.17%. Although this approach requires simple data, such as time-series data on product yield rate, it requires human experts to construct the fuzzy yield forecast. Therefore, manufacturers that offer multiple customized products in large numbers of batches may find this approach to be excessively effort intensive.

Jun et al. [16] proposed a micro approach to constructing a model to predict any defect in the production process. This approach requires several variables related to 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 time-series data and predict 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 yields by using the predictions to reduce the occurrence of defects, and it improved the yield by approximately 8.7% in a continual process. Although the production yield rate can be predicted using this approach, the computational costs are prohibitively high because deep learning is used [19] and the process must be executed at every equipment inspection run. Moreover, it is difficult to directly reapply this approach to other domains, such as batch production systems, because it demands the use of a particular statistical model.

Many manufacturers offer high-quality customized products [1] and commonly use large numbers of batches in the manufacturing process. In this case, although many existing approaches can predict a single instance of the production yield rate, to the best of our knowledge, no approach can predict the per-batch yield rate.

3. Maximum Likelihood Estimation of Per-Machine Yield Rate

To accelerate the manufacture 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) [2] for future reference, where the BoO contains operational information, such as the sequence of stations for each batch.

Fig. 1 illustrates how a batch of products can be processed with any machine in a station described in the BoO at the time of manufacturing. However, a manufacturer can use their machines for many purposes. Any machine can be used by only one BoO, and the other machines are used by one or several BoOs in any of their sequences. Moreover, a batch may use the very same machine more than once in its sequence, if that batch requires some stations to be revisited.

Although a manufacturer may process many batches in a single day, each machine can only handle a single batch at a time, including the inspection equipment. Therefore, if the inspection equipment detects any defective products, we can assume that any of the previous machines, including the current machine, may have generated the defective pieces. For example, in Figure 2, for each defective piece observed in the *j*th manufacturing step, all previous machines, including the current one (at the *k*th step), are suspicious machines, where $k \leq j$. Consequently, it becomes possible to estimate the yield rate of each production machine based on any observed defective pieces in each manufacturing process for each *i*th batch.

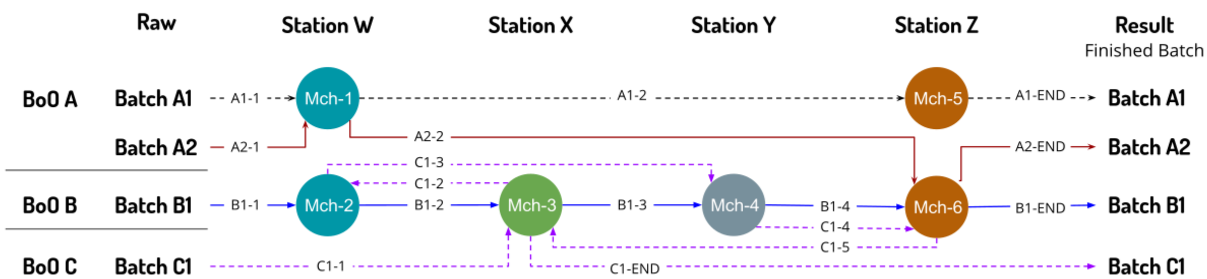

Fig. 1 Example involving three bills of operation (A to C) that use four stations (W to Z) to produce four batches (A1 to C4).

The total number of defective products in a batch is equal to the sum of 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 through (let the variable be θ). Accordingly, based on FPY theory, we designed a likelihood function for each manufacturing step, which we use to estimate the per-machine yield rate, as encapsulated in (1) to (5).

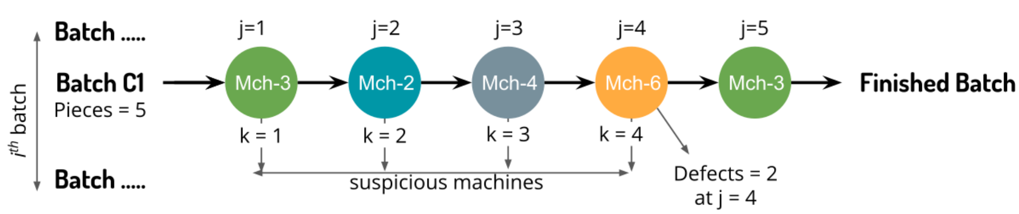

Fig. 2 Illustration of suspicious machines that are generating defective pieces.

Let $P(Y,Z|\theta)$ be the likelihood that the sets Y (observed variable of the defective piece) and Z (indicator variable for the machine that generates the defective piece) occur given θ (the set of machine yield rates). $P(Y,Z|\theta)$ is calculated from all batches in the manufacturing process, as given in (1), where y_i and z_i are the elements of Y and Z, respectively. In addition, let I be the number of batches in the manufacturing process and let $P(y_i,z_i|\theta)$ be the likelihood function of each *i*th batch.

$$P(Y,Z|\theta) = \prod_{i=1}^{I} P(y_i, z_i|\theta) \tag{1}$$

First, let N_i be the number of processed pieces in the *i*th batch. Subsequently, equation (2) can be used to calculate $P(y_i, z_i|\theta)$ based on the condition of each *n*th piece (of the *i*th batch in the manufacturing process) observed to be defective (value of 1) or in good condition (value of 0), represented as y_{in}. In this case, y_{in} affects both F_1 and F_2.

Let P_{ik} be the probability that a product piece will be good when using the machine associated with the *k*th manufacturing step of the *i*th batch (yield rate of a machine in the *k*th manufacturing step of the *i*th batch); let J_{in} be the number of manufacturing steps completed in the processing of the *n*th piece of the *i*th batch. Subsequently, based on the defect rate of the production machines (represented as $1 - P_{i,k}$), F_1 denotes the likelihood function of the defect rate if the *n*th piece of the *i*th batch is observed to be defective in a manufacturing process, as shown in (3). In each manufacturing process, the likelihood is the defect rate of the current manufacturing process $(1 - P_{i,k})$ multiplied by the yield rate of previous production machines $(\prod_{s=1}^{k-1} P_{i,s})$. F_1can be used to estimate the indicator variable of the *n*th piece of the *i*th batch observed to be defective due to the machine used in the *k*th step, and it is denoted z_{ink}. If the *n*th piece is observed to be defective ($y_{in} = 1$) in the *j*th step of the manufacturing process, F_1 will provide an interval value between 0 and 1.

Moreover, based on the yield rates of the production machines, F_2 is the likelihood function of the yield rate if the *n*th piece of the *i*th batch is observed to be a good piece in a manufacturing process, as shown in (4). The calculation is straightforward, involving multiplication with all the production machine yield rates denoted by each P_{ik} that was used in a specific batch. Therefore, if the *n*th piece is observed to be good ($y_{in} = 0$) in the *j*th step of the manufacturing process, the value of F_1 is 1 and the value of F_2 is between 0 and 1.

$$P(y_i, z_i|\theta) = \prod_{n=1}^{N_i} \big(F_1(i,n) \times F_2(i,n)\big) \tag{2}$$

where

$$F_1(i,n) = \left(\prod_{k=1}^{J_{i;n}} \left((1-P_{ik})\prod_{s=1}^{k-1} P_{is}\right)^{z_{in;k}}\right)^{y_{in}} \quad (3)$$

$$F_2(i,n) = \left(\prod_{k=1}^{J_{i;n}} P_{ik}\right)^{1-y_{in}} \quad (4)$$

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 still assign an available machine at each station in the station sequence to the batch during production. Hence, in (3) and (4), the yield rate of each machine (P_M) is mapped using the *i;k* index, which gives us the notation $P_{M(i;k)}$. However, this notation makes the equation less readable and more complex. Hence, 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 written in (5).

Given that our objective is to estimate the yield rate of each machine, the likelihood in (5) can be used to estimate the defective pieces generated by each machine, which can subsequently be used to estimate the yield rate of each machine.

$$P(Y,Z|\theta) = \prod_{i=1}^{I}\prod_{n=1}^{N_i}\left(\left(\prod_{k=1}^{J_{i;n}}\left((1-P_{ik})\prod_{s=1}^{k-1}P_{is}\right)^{z_{in;k}}\right)^{y_{in}} \times \left(\prod_{k=1}^{J_{i;n}}P_{ik}\right)^{1-y_{in}}\right) \quad (5)$$

4. Proposed Method

Since our approach is related to the unknown or missing machine yield rate data, we propose a new EM-based algorithm to solve this problem. Using the algorithm, the estimated per-machine yield rate could be used to make the predictions.

Our proposed approach aims to estimate the machine yield rate by iterating the EM algorithm until the most convergent result is obtained. The overall procedure of our approach is shown in the activity diagram in Fig. 3. To give a brief route of our explanation, first, in step (0), we need to preprocess and clean the raw data, and we then guess the initial machine yield rates in step (1). Step (7) is the objective of each iteration, which improves the estimation of the yield rate of each machine based on the variables of the expected number of good pieces and the expected number of defective pieces generated by that machine. Both variables are estimated based on the observed defective pieces in step (2). Due to the fact that both variables can be calculated independently, we use the parallelism (bar) symbol after step (2). Hereinafter, while the expected number of good pieces could be calculated by steps (4), (5), and (6) sequentially, the expected number of defective pieces could be calculated by step (3). Next, the bottom bar symbol indicates that we need to use both variables in step (7). Finally, when the EM iteration stops in step (8), the future batch yield rates could be predicted using the prior estimated machine yield rates. We will explain the details of the algorithm in the following subsections.

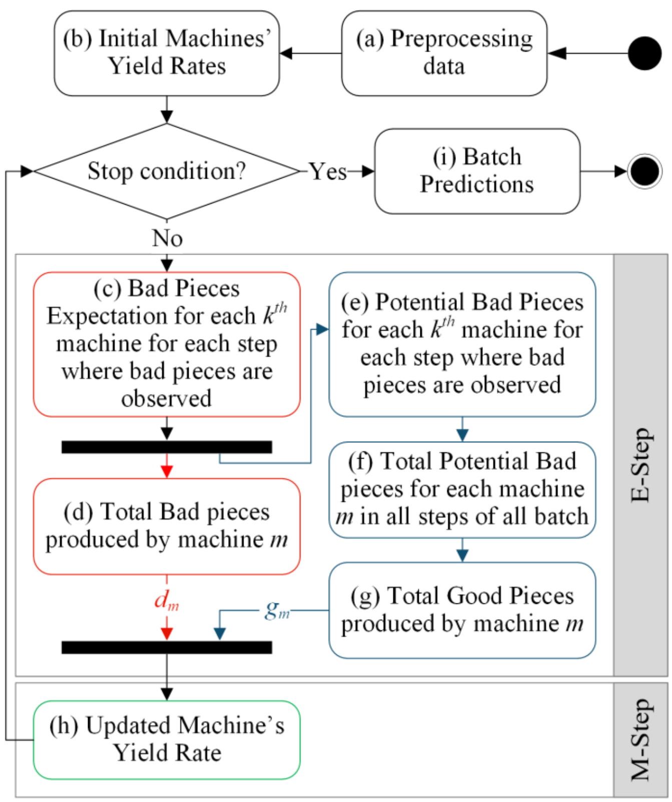

Fig. 3 Activity diagram of the proposed approach to estimate machine yield.

4.1. Preprocessing

Because the raw production data provided by most manufacturers is not clean, data preprocessing and cleaning are required, as indicated by step (a) in Fig. 3. Several steps must be executed to initialize the calculation, such as preprocessing and cleaning 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 those sessions into one session. Finally, an example of the required preprocessed based on Fig. 1 is summarized in TABLE I.

TABLE I
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
…	…	…	…	…

4.2. Initial Yield Rates

Initially, because the machine yield rates are unknown, we must guess the yield rates of all the machines used in the manufacturing process, as indicated by step (b) in Fig. 3. Apart from random guessing, many approaches can be used to guess the yield rates. For example, we can use constrained least-squares to generate the initial machine yield rates based on the batch yield rates, as follows:

$$\arg\min_{\theta} \sum_{i=1}^{I} \left(\sum_{k=1}^{J_i} q_{ik} - \ln l_i \right)^2 \tag{6}$$

subject to $q_{ik} \leq 0$. In equation (6), let J_i be the number of manufacturing steps for processing the *i*th batch; q_{ik} the natural logarithm of P_{ik}; θ is the set $\{q_{ik} | (1 \leq i \leq I) \wedge (1 \leq k \leq J_i)\}$; and l_i is the yield rate of the *i*th batch of products.

4.3. M-Step

The yield rates θ are the initial input for E-Step of the proposed approach. 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 generated among all the pieces generated by that machine. Therefore, the objective of each iteration can be expressed as follows:

$$Pr(m) = \frac{g_m}{g_m + d_m} \tag{7}$$

where $Pr(m)$ is the probability of obtaining good pieces by using machine *m* (yield rate of machine *m*). In addition, let g_m be the expected total number of good pieces generated by machine *m* in every manufacturing step, and let d_m be the expected total number of defective pieces generated by machine *m* in every manufacturing step.

4.4. E-Step

To obtain d_m and g_m, we must estimate the number of defective pieces due to each suspicious machine where defective pieces are observed, as indicated by step (c) in Fig. 3. Because we have no prior knowledge about the per-machine yield rates, we assume that the yield rates of all the machines are close to 0.999. However, according to the principle of likelihood, when the inspection equipment observes defective pieces, all the previous machines, including the current machine, that are used to process that batch are suspicious. Therefore, we expect the shared probability of defective pieces to be distributed among the suspicious machines based on their yield rates. A particular machine could have a lower estimated yield rate if multiple batches yield many defective pieces after using that machine. Consequently, if any defective pieces are observed in each manufacturing step of each batch, we must first estimate the likelihood that each particular machine has caused that defect (z_{ink}) by using (5). Hence, we have (8) to calculate the likelihood that the *k*th manufacturing step of the *i*th batch causes the *n*th piece to be defective (value of 1) with the given θ, which is represented as $P(z_{ink} = 1|y_{in}, \theta)$.

$$P(z_{ink} = 1|y_{in}, \theta) = \begin{cases} 0; & if\ y_{in} = 0 \\ \dfrac{(1 - P_{ik})(\prod_{s=1}^{k-1} P_{is})}{\sum_{t=1}^{J_{i;n}}\left((1 - P_{it})(\prod_{u=1}^{t-1} P_{iu})\right)}; & if\ y_{in} = 1 \end{cases} \tag{8}$$

For each *n*th piece observed to be defective in the *j*th manufacturing step, each particular *k*th manufacturing step is probably suspect. Therefore, let $E[z_{ink}]$ be the expectation that a particular *k*th manufacturing step causes the *n*th piece of the *i*th batch to be defective, which is calculated using (9).

$$E[z_{ink}] = 0 * P(z_{ink} = 0|y_{ink}, \theta) + 1 * P(z_{ink} = 1|y_{ink}, \theta)$$
$$E[z_{ink}] = P(z_{ink} = 1|y_{ink}, \theta) \tag{9}$$

For each *j*th manufacturing step of the *i*th batch, more than one piece may be observed to be defective. Consequently, several observed defective pieces have the same $E[z_{ink}]$. Subsequently, because we must estimate the number of defective pieces due to each suspicious machine, the likelihood estimation in (9) could be multiplied with the number of defective pieces observed in that manufacturing step, as in (10).

$$e_{ijk} = E[z_{ink}] \times b_{ij} \tag{10}$$

where e_{ijk} is the expected number of defective pieces generated in the *k*th manufacturing step when any defective pieces of the *i*th batch are observed in *j*th manufacturing step; $b_{ij}$is the number of defective pieces observed in the *i*th batch in the *j*th manufacturing step.

To obtain d_m, we can combine the expected number of defective pieces for each machine *m* in each manufacturing process from all the batches, as indicated by step (d) in Fig. 3. The machine *m* is used in several manufacturing steps, which are registered in the set of *(i,j,k)* indexes of machine *m*, as in (11). Accordingly, we can determine the expected number of defective pieces for each machine *m*, as in (12).

$$S_m = \{(i, j, k) | \forall m_{ik} = m\} \tag{11}$$

$$d_m = \sum_{(i,j,k) \in S_m} e_{ijk} \tag{12}$$

In (11) and (13), let S_m be the set of *(i,j,k)* indexes (as tuple elements) of all batches, which are the machines in each *k*th step of the manufacturing process of the *i*th batch of the products in which defective pieces are observed or detected when the *j*th step is completed. Let m_{ik} be the *m*th machine that is used to complete the *k*th manufacturing step of the *i*th batch.

Meanwhile, g_m can be determined by combining the number of observed good pieces at the end of the manufacturing process of the *i*th batch by using a particular machine (represented as f_i) and the total number of good pieces generated by that machine that will be defective in the subsequent manufacturing process (represented as x_m). However, to obtain x_m, we must first calculate the potential number of defective pieces based on e_{ijk} in each manufacturing step, as in step (e) in Fig. 3, which can be expressed as follows:

$$h_{ijk} = \begin{cases} 0; & if\ j = k \\ h_{ij;k+1} + e_{ijk}; & if\ j > k \end{cases} \tag{13}$$

where h_{ijk} is the number of potential defective pieces generated in the *k*th manufacturing step of the *i*th batch of products, in which defective pieces are observed or detected in the *j*th manufacturing step. Subsequently, we can combine the h_{ijk} of each machine in each manufacturing process of all the batches, as in step (f) in Fig. 3, which can be expressed as follows:

$$x_m = \sum_{(i,j,k) \in S_m} h_{ijk} \tag{14}$$

Subsequently, because x_m and f_i are known, we can sum them into 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 in the sequence. Therefore, it can be written as

$$g_m = x_m + \left(\sum_{i=1}^{I} f_i \cdot r_{im} \right) \tag{15}$$

where r_{im} is how many times machine *m* is used in the manufacturing process of the *i*th batch.

4.5. Stop Condition

The EM algorithm continues to iterate until a stopping condition is fulfilled. Two stopping conditions can be used, namely a convergence threshold and the maximum number of iterations. When the difference between the yield rates in the current iteration and those in the previous iteration calculated using the Euclidean distance [20] 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, the threshold of 0.001, which will not exert any significant influence in further iterations, is deemed to be sufficient to stop the algorithm from iterating.

4.6. Prediction of Production Yield Rate

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 is a reliable basis for predicting the production yield rate. Consequently, the result of our approach can be used to predict the production yield rate, as in step (i) in Fig. 3. Based on the FPY of each batch, the yield rate of a batch is equal to the product of all the machines used by the batch. Therefore, the average accuracy of all the batches in a particular week can be expressed as follows:

$$\overline{acc} = \frac{\sum_{i=1}^{I}\left(1 - \left|\left(\prod_{j=1}^{J_i} P_{ij}\right) - l_i\right|\right)}{I} \tag{16}$$

where $\overline{acc}$ represents the average accuracy of our approach's result in estimating the yield rate of all batches one period or one-week data.

In this study, we used the data of only one particular week to predict the production yield rate in the following weeks, up to 5 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. Experimental Results and Discussions

We used data from T-company (70 weeks of data) to evaluate the prediction performance of our proposed approach, as described in subsection 5.1. To determine why the proposed approach achieved such a high accuracy, we conducted several simulations, as described in subsection 5.2. Thereafter, we discuss the proposed approach based on the experimental and simulation results.

TABLE II
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	1,075	747	154.3
Defective pieces in a batch	0	163,072	1,140	3325.0
Number of steps in a batch	1	59	21	7.0
Inspection steps in a batch	1%	9%	8%	1%

5.1. Experimental Results Obtained Using T-Company Data

In our study, we used a large, real-world, 70-week data set from T-company on its manufacturing. The statistics of these real-world data are summarized in TABLE II. To evaluate the performance of our approach, we used 1-week data to construct our prediction model; subsequently, we used the model to predict the consequent average batch production yield rate for the following five weeks.

As shown in Fig. 4, the predicted batch yield rates for week one had an average accuracy of 91.86%. Moreover, the results indicated that the proposed approach could yield predictions for weeks 2 to 4 (next month) at an average accuracy >90%. However, the predictions for weeks 3 and beyond were more uncertain because of a sharp increase in the standard deviation.

According to Fig. 5, the plots for the five predictions (weeks 1 to 5) were extremely similar, which explained why the prediction accuracy values in Fig. 4 were extremely similar. Moreover, the prediction accuracy for week 5 could be as low as 79% in some rare cases. Thus, the standard deviation of the prediction for the consequent fifth week was larger. In subsection 5.3, we discuss why the prediction accuracy changed.

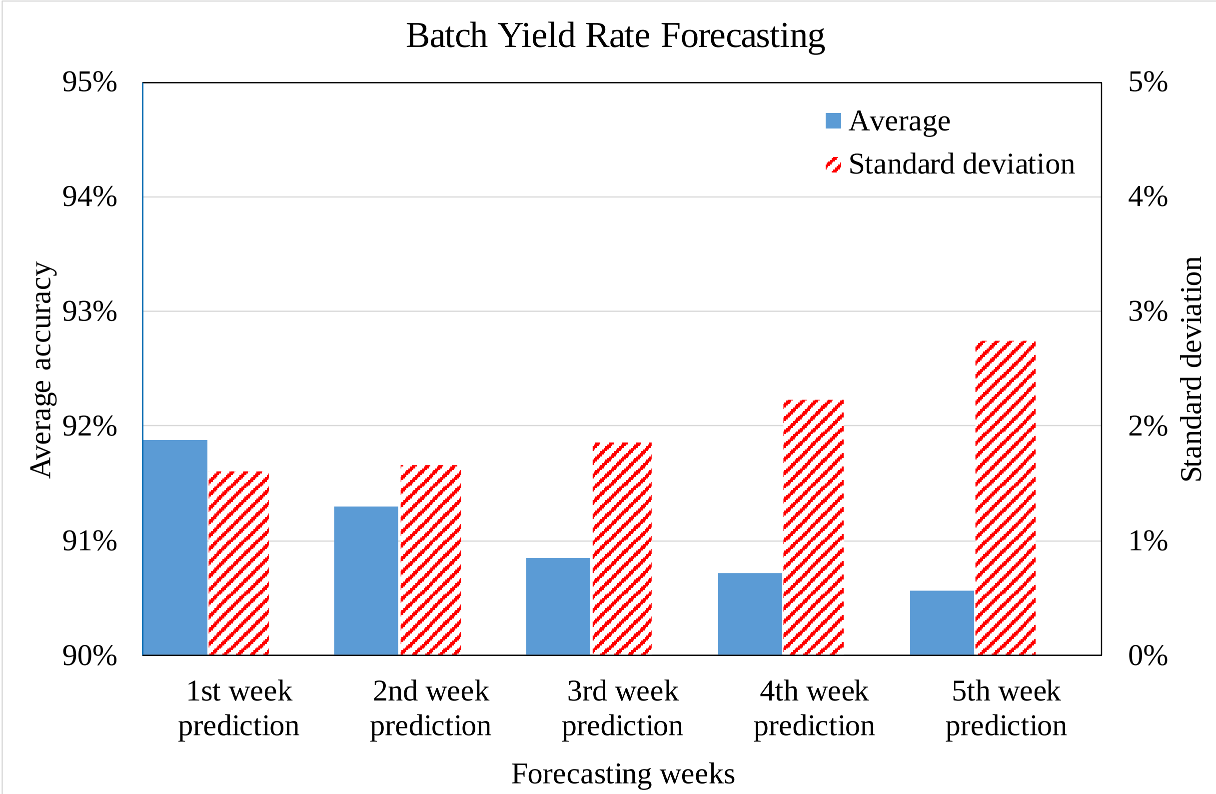

Fig. 4 Average accuracy of T-Company's batch yield rate prediction when using the proposed approach.

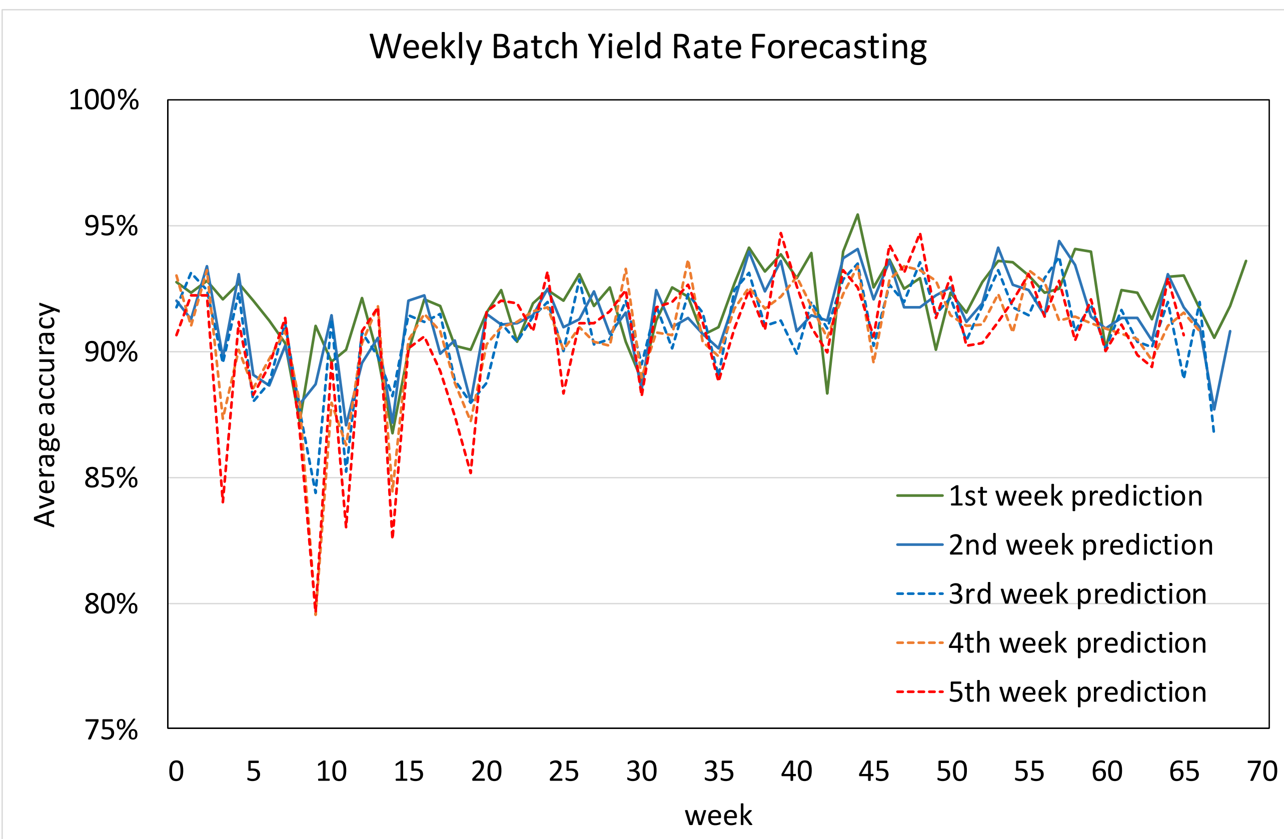

Fig. 5 Average accuracy of each prediction of T-Company data.

5.2. Simulations of Machine Yield Rate

For most manufacturers, including T-company, actual per-machine yield rate data are unavailable because it is impractical to use quality inspection equipment in all production sequences. Therefore, we performed simulations with generated per-machine yield rates and production data. In these simulations, we used production data to estimate the per-machine yield rate. Thereafter, we performed a sensitivity analysis to evaluate the performance of our approach for each manufacturing data set of a given size. We used several variables in the experiments, which were the set of machines (50, 250, and 500 machines), set of batches (500–5000 batches in increments of 500, and a special set of 10 batches at minimum), set of average batch steps (20 and 30 steps), and set of average inspections ratio (10% and 30% of batch steps).

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 under point A in TABLE III to determine the yield rate of each machine. Subsequently, we installed inspection devices on 30% of the production machines, and these inspection devices could be turned off by manufacturers during production for whatever reason.

After obtaining the machine data, we could generate batch attributes, such as the number of batch steps, number of raw pieces, and number of inspections in a batch, as points B, C, and D, respectively, in TABLE III. 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 must involve the inspection machines.

Second, we generated the number of defective pieces in each step. For each step, we randomized the yield rate of the 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.

TABLE III
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	1000	20000	10000	2000
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	88%	100%	90%	2%

* based on the simulation combination

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 given by point E in TABLE III. This finding supported by studies in which the observation accuracy had a tolerance of 10%–30% [21,22], meaning 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 0% tolerance, and all the remaining defective pieces were then observed.

Finally, each combination was used in our approach to estimate the yield rate of each machine. However, because we had to run the approach 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 result as our final result.

5.2.2. Simulation results

Fig. 6 presents the results of our simulation experiments, which demonstrate the good performance of our approach in estimating the yield rates of the machines in different simulation combinations. A manufacturer that uses 50 machines for production with an average of 20 batch steps may be required to produce fewer than 500 batches to obtain an accuracy higher than 90%; a production run with at least approximately 1,000 batches may be required to obtain the best accuracy, as indicated by point A in Fig. 6. However, when a manufacturer uses approximately 250 machines for production with the same number of batch steps, the manufacturer may be required to produce at least approximately 4,000 batches to obtain the best accuracy, as indicated by point B in Fig. 6. Based on the operational chart point E in Fig. 6, which is similar to the data model of T-company, our approach could provide a good estimate of the per-machine yield rate at an average accuracy of 92.06%.

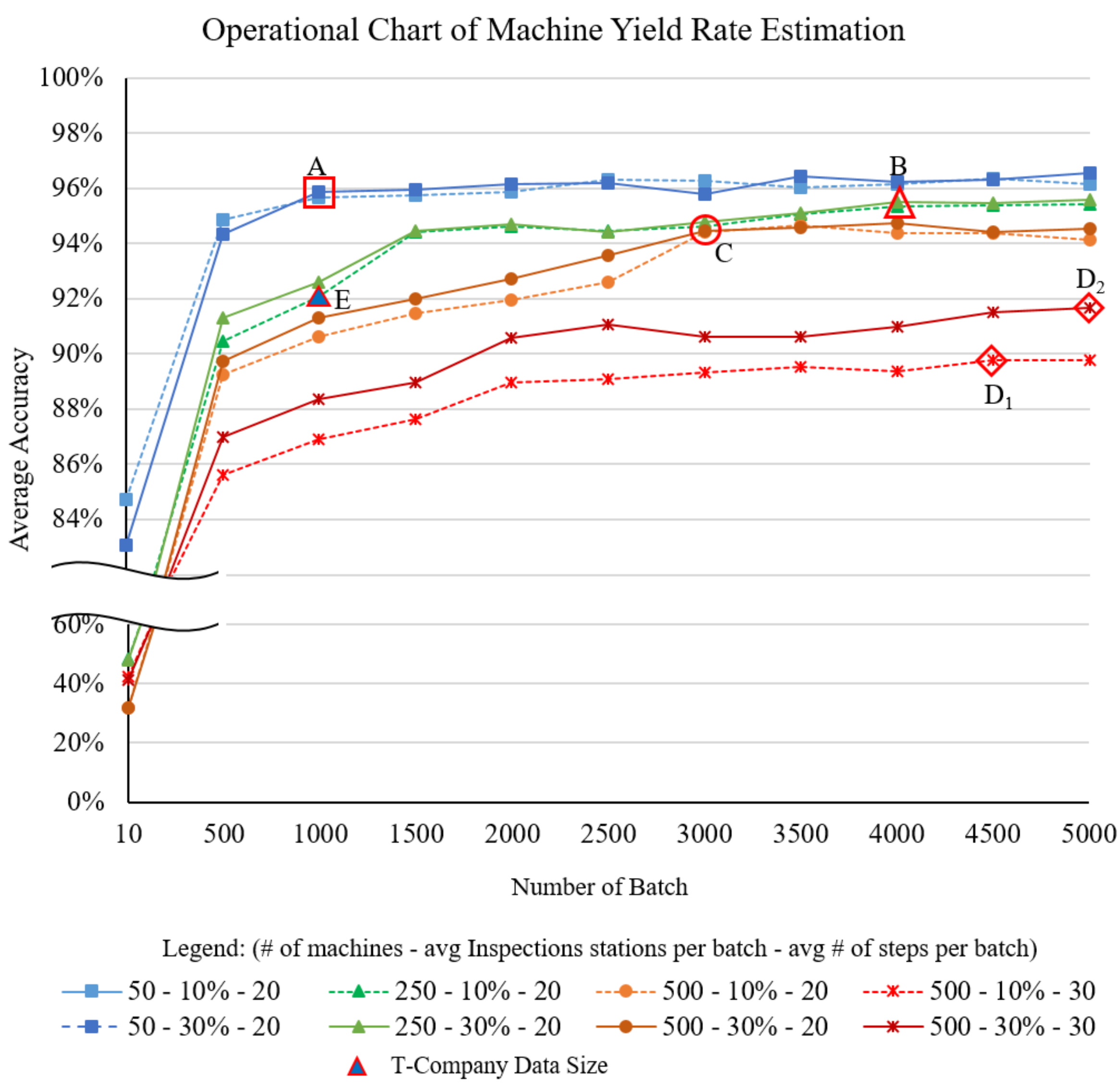

Fig. 6 Operational chart of the proposed approach to estimate the per-machine yield rates.

A manufacturer with 500 machines and an average of 20 batch steps in their production may be required to produce at least approximately 1,000 batches to obtain an accuracy higher than 90%, as indicated by point C in Fig. 6. At the same number of machines, an average of 30 batch steps led to a decrease in accuracy, as indicated by both points D in Fig. 6. However, by using approximately 30% of the inspection machines in each batch, as indicated by point D_2 in Fig. 6, the accuracy increased compared with that when using only approximately 10% of the inspection machines in each batch, as indicated by point D_1 in Fig. 6.

5.3. Discussion

According to the experiments in section 5.1, our approach accurately predicted batch yield rates. In the case of T-company, although our approach could generate predictions for periods longer than 1 month ahead, we suggest that predictions should be made for only 1–2 weeks (or periods) ahead because the standard deviation increases when predictions are made for periods further into the future. This may be due to the following reasons. First, the machines used in future weeks may be different, which may not be used in the week where the machine yield rates were estimated. Hence, fewer batches can be predicted. Second, the yield rates of some machines estimated in a particular week may be affected by machine wear and tear in the following weeks, which may lower their actual yield rates. Conversely, they may undergo maintenance work in the following week, which may improve their actual yield rates. Third, in rare cases, our approach predicts batch yield rate at an accuracy that is as low as 79%. In this study, we observed that this was caused by the processing of a few batches in the week when the per-machine yield rates were estimated, which resulted in less accurate estimates of per-machine yield rate. This observation is supported by Fig. 7, in which the fluctuation of the plot of the number of batches shows similarity to the plot of week 5 prediction in Fig. 5.

In this study, although we could not compare the prediction results for the production data of different manufacturers, the simulation of machine yield rates could be used as a benchmark when our proposed approach is used. 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 data set was more random if more machines were used in production. This makes our approach less accurate. In this case, we suggest for more batches to be used to obtain accurate estimates, but doing so could be difficult in practice. Moreover, with the same number of machines, long sequences of each batch (longer batch steps) reduces 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.

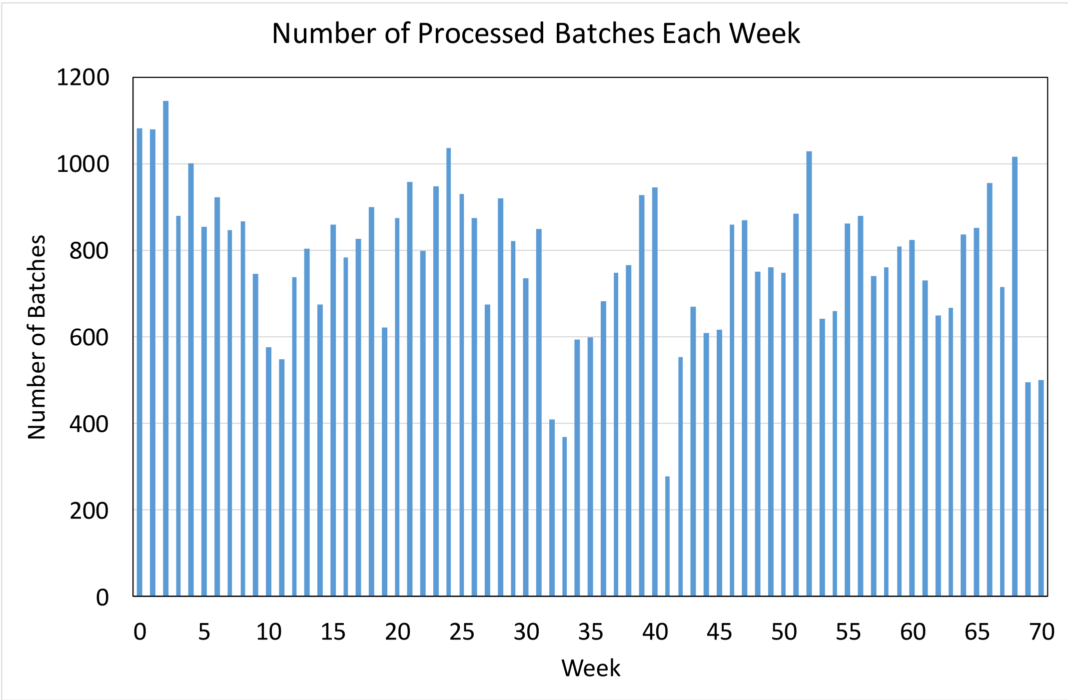

Fig. 7. The number of batches each week on T-Company data

The reduction in accuracy caused by a long sequence can be alleviated by increasing the number of batches and the density of inspection stations, which, however, increases production cost. Therefore, managers must balance between estimation accuracy (through having more inspection stations) and production cost [23]. Although the number of machines, number of batches, and 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. Accordingly, based on the simulation results presented in subsection 5.2.2, our approach works well for small- and medium-scale manufacturers, who use 500 or fewer machines and have an average batch sequence with 30 or fewer steps.

6. Conclusions and Directions for Future Research

The proposed approach is a useful solution to predict the production yield rate, based on the per-machine yield rate estimations. Using the T-company data, while our approach could provide next week's prediction of production yield rate with an average accuracy of 91.86%, and continuously over 90% for 5 consecutive weeks (over 1 month), it is suggested to predict the batch yield rates for only one or two weeks. Besides, the EM Algorithm proved to be the current best solution in our approach. Since it helps to calculate the unknown machine yield rate based on the observed defective pieces in the production data. In our experiment, by using generated data similar to the T-company data size, our approach could provide a good estimation of the per-machine yield rates, which has an average accuracy of 92.06%.

On the other hand, based on our simulation results, our approach provides a good result for any manufacturers with 500 or less machines with an average of 30 or shorter batch sequence. In this case, the number of machines, the number of batches, and the average of batch steps are the most important variables to estimate the per-machine yield rate. Our proposed approach may be one of the reasonable approaches that could be used by manufacturers, to obtain their estimated per-machine yield rate. Hereinafter, by using our approach, the manufacturer could predict the production yield rate, to get better prepared for their production, and analyze its costs.

Our lightweight approach uses only production data without too many parameters. Hence, manufacturers with limited resources can implement this approach with ease. However, our approach has several limitations related to batch specifications and flow-shop manufacturing (or manufacturing of a similar type). First, our approach estimates the machine yield rate based on their station sequence. Although the approach is accurate, it may provide unexpected results for fixed sequences of machine stations because the machines at the start of the sequence are necessarily considered suspicious. Hence, their yield rates may be lower than those of the machines in the final sequence. This means that unexpected results may be obtained in flow-shop manufacturing (or similar). Second, based on the design of our approach, any suspicious 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 up 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. We aim to address these two limitations in future research.

Acknowledgments

This work was supported by The Ministry of Science and Technology of Taiwan (Grant No: 108-2221-E-008-030-MY3 and 108-2221-E-008-032-MY3)

References

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

[2] Jianxin Jiao, M.M. Tseng, Qinhai Ma, Yi Zou, Generic Bill-of-Materials-and-Operations for High-Variety Production Management, Concurr. Eng. 8 (2000) 297–321. https://doi.org/10.1177/1063293X0000800404.

[3] T.-S. Li, C.-L. Huang, Z.-Y. Wu, Data Mining using Genetic Programming for Construction of a Semiconductor Manufacturing Yield Rate Prediction System, J. Intell. Manuf. 17 (2006) 355–361. https://doi.org/10.1007/s10845-005-0008-7.

[4] V.R. Joseph, H. Adya, Prediction of Yield in a Multiproduct Batch Production Environment, Qual. Eng. 14 (2002) 153–159. https://doi.org/10.1081/QEN-100106894.

[5] S.J. Bae, J.Y. Hwang, W. Kuo, Yield prediction via spatial modeling of clustered defect counts across a wafer map, IIE Trans. 39 (2007) 1073–1083. https://doi.org/10.1080/07408170701275335.

[6] Tory Chen, M.-J.J. Wang, A fuzzy set approach for yield learning modeling in wafer manufacturing, IEEE Trans. Semicond. Manuf. 12 (1999) 252–258. https://doi.org/10.1109/66.762883.

[7] R.R. Mohan, K. Thiruppath, R. Venkatrama, S. Raghuraman, Quality Improvement through First Pass Yield using Statistical Process Control Approach, J. Appl. Sci. 12 (2012) 985–991.

https://doi.org/10.3923/jas.2012.985.991.

[8] I. Tirkel, G. Rabinowitz, D. Price, D. Sutherland, Wafer fabrication yield learning and cost analysis based on in-line inspection, Int. J. Prod. Res. 54 (2016) 3578–3590. https://doi.org/10.1080/00207543.2015.1106609.

[9] M. Chincholkar, J.W. Herrmann, Estimating manufacturing cycle time and throughput in flow shops with process drift and inspection, Int. J. Prod. Res. 46 (2008) 7057–7072. https://doi.org/10.1080/00207540701513893.

[10] P. Cawley, Structural health monitoring: Closing the gap between research and industrial deployment, Struct. Heal. Monit. 17 (2018) 1225–1244. https://doi.org/10.1177/1475921717750047.

[11] S. Patel, B.G. Dale, P. Shaw, Set-up time reduction and mistake proofing methods: an examination in precision component manufacturing, TQM Mag. 13 (2001) 175–179. https://doi.org/10.1108/09544780110385528.

[12] X. Mehmeti, B. Mehmeti, R. Sejdiu, The equipment maintenance management in manufacturing enterprises, IFAC-PapersOnLine. 51 (2018) 800–802. https://doi.org/10.1016/j.ifacol.2018.11.192.

[13] C. Herzet, X. Wautelet, V. Ramon, L. Vandendorpe, Iterative Synchronization: EM Algorithm Versus Newton-Raphson Method, in: 2006 IEEE Int. Conf. Acoust. Speed Signal Process. Proc., IEEE, n.d.: pp. IV-393-IV–396. https://doi.org/10.1109/ICASSP.2006.1660988.

[14] T. Springer, K. Urban, Comparison of the EM algorithm and alternatives, Numer. Algorithms. 67 (2014) 335–364. https://doi.org/10.1007/s11075-013-9794-8.

[15] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, New York, 2006.

[16] J. Jun, T.-W. Chang, S. Jun, Quality Prediction and Yield Improvement in Process Manufacturing Based on Data Analytics, Processes. 8 (2020) 1068. https://doi.org/10.3390/pr8091068.

[17] E.-D. Sandru, E. David, G. Pelz, Pre-Silicon Yield Estimation using Machine Learning Regression, in: 2019 26th IEEE Int. Conf. Electron. Circuits Syst., IEEE, 2019: pp. 103–104. https://doi.org/10.1109/ICECS46596.2019.8964997.

[18] T. Chen, M.-C. Chiu, 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.

[19] N.C. Thompson, K. Greenewald, K. Lee, G.F. Manso, The Computational Limits of Deep Learning, (2020). http://arxiv.org/abs/2007.05558.

[20] R. Dwivedi, N. Ho, K. Khamaru, M.I. Jordan, M.J. Wainwright, B. Yu, Singularity, Misspecification, and the Convergence Rate of EM, Stat. Theory. (2018). http://arxiv.org/abs/1810.00828.

[21] J.E. See, Visual Inspection Reliability for Precision Manufactured Parts, Hum. Factors J. Hum. Factors Ergon. Soc. 57 (2015) 1427–1442. https://doi.org/10.1177/0018720815602389.

[22] S.F. Beckert, W.S. Paim, Critical analysis of the acceptance criteria used in measurement systems evaluation, Int. J. Metrol. Qual. Eng. 8 (2017) 23. https://doi.org/10.1051/ijmqe/2017016.

[23] M.A. Farooq, R. Kirchain, H. Novoa, A. Araujo, Cost of quality: Evaluating cost-quality trade-offs for inspection strategies of manufacturing processes, Int. J. Prod. Econ. 188 (2017) 156–166. https://doi.org/10.1016/j.ijpe.2017.03.019.

Biography

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.

Appendix

A. Notation descriptions

A.1 List of variables with known values

Notation	Description
b_{ij}	the number of defective pieces observed in the *i*th batch in the *j*th manufacturing step.
f_i	the number of observed good pieces at the end of the manufacturing process of the *i*th batch by using a particular machine
I	the number of batches in the manufacturing process
J_i	the number of manufacturing steps for processing the *i*th batch
J_{in}	the number of manufacturing steps completed in the processing of the *n*th piece of the *i*th batch
l_i	the yield rate of the *i*th batch of products.
m_{ik}	the *m*th machine that is used to complete the *k*th manufacturing step of the *i*th batch
N_i	the number of processed pieces in the *i*th batch
r_{im}	how many times machine *m* is used in the manufacturing process of the *i*th batch.
S_m	the set of *(i,j,k)* indexes (as tuple elements) of all batches, which are the machines in each *k*th step of the manufacturing process of the *i*th batch of the products in which defective pieces are observed or detected when the *j*th step is completed
y_{in}	the condition of each *n*th piece (of the *i*th batch in the manufacturing process) observed to be defective (value of 1) or in good condition (value of 0),

A.2 List of variables with unknown values or to be estimated

Notation	Description
$\overline{acc}$	the average accuracy of our approach's result in estimating the yield rate of all batches one period or one-week data.
d_m	the expected total number of defective pieces generated by machine *m* in every manufacturing step.
e_{ijk}	the expected number of defective pieces generated in the *k*th manufacturing step when

	any defective pieces of the *i*th batch are observed in *j*th manufacturing step
$E[z_{ink}]$	the expectation that a particular *k*th manufacturing step causes the *n*th piece of the *i*th batch to be defective
$F_1(i; n)$	the likelihood function of the defect rate if the *n*th piece of the *i*th batch is observed to be defective in a manufacturing process
$F_2(i; n)$	the likelihood function of the yield rate if the *n*th piece of the *i*th batch is observed to be a good piece in a manufacturing process
g_m	the expected total number of good pieces generated by machine *m* in every manufacturing step
h_{ijk}	the number of potential defective pieces generated in the *k*th manufacturing step of the *i*th batch of products, in which defective pieces are observed or detected in the *j*th manufacturing step
P_{ik}	the probability that a product piece will be good when using the machine associated with the *k*th manufacturing step of the *i*th batch (yield rate of a machine in the *k*th manufacturing step of the *i*th batch).
$P(Y, Z\|\theta)$	the likelihood that the sets Y (observed variable of the defective piece) and Z (indicator variable for the machine that generates the defective piece) occur given θ (the set of machine yield rates).
$P\left(y_i, z_i \middle\| \theta\right)$	the likelihood function of each *i*th batch.
$P(z_{ink} = 1\|y_{in}, \theta)$	the likelihood that the *k*th manufacturing step of the *i*th batch causes the *n*th piece to be defective (value of 1) with the given θ
$Pr(m)$	the probability of obtaining good pieces by using machine *m* (yield rate of machine *m*).
q_{ik}	the natural logarithm of P_{ik}
x_m	the total number of good pieces generated by that machine that will be defective in the subsequent manufacturing process
$z_{in;k}$	the indicator variable of the *n*th piece of the *i*th batch observed to be defective due to the machine used in the *k*th step
θ	the set $\{q_{ik}\|(1 \leq i \leq I) \land (1 \leq k \leq J_i)\}$

B. Equations list

Equation	Eq. Number
Simple version of likelihood $$P(Y,Z\|\theta)=\prod_{i=1}^{I}P(y_i,z_i\|\theta)$$	1
The likelihood for each batch $$P(y_i,z_i\|\theta)=\prod_{n=1}^{N_i}\big(F_1(i,n)\times F_2(i,n)\big)$$	2
$$F_1(i,n)=\left(\prod_{k=1}^{J_{i;n}}\left((1-P_{ik})\prod_{s=1}^{k-1}P_{is}\right)^{z_{in;k}}\right)^{y_{in}}$$	3
$$F_2(i,n)=\left(\prod_{k=1}^{J_{i;n}}P_{ik}\right)^{1-y_{in}}$$	4
Complete Likelihood Function $$P(Y,Z\|\theta)=\prod_{i=1}^{I}\prod_{n=1}^{N_i}\left(\left(\prod_{k=1}^{J_{i;n}}\left((1-P_{ik})\prod_{s=1}^{k-1}P_{is}\right)^{z_{in;k}}\right)^{y_{in}}\times\left(\prod_{k=1}^{J_{i;n}}P_{ik}\right)^{1-y_{in}}\right)$$	5
Constrained Least square $$\arg\min_{\theta}\sum_{i=1}^{I}\left(\sum_{k=1}^{J_i}q_{ik}-\ln l_i\right)^2$$ subject to $q_{ik}\le 0$	6
Updating machine yield rates $$Pr(m)=\frac{g_m}{g_m+d_m}$$	7
Likelihood of k^{th} manufacturing steps generates defective pieces $$P(z_{ink}=1\|y_{in},\theta)=\begin{cases}0; & if\ y_{in}=0\\ \frac{(1-P_{ik})(\prod_{s=1}^{k-1}P_{is})}{\sum_{t=1}^{J_{i;n}}\big((1-P_{it})(\prod_{u=1}^{t-1}P_{iu})\big)}; & if\ y_{in}=1\end{cases}$$	8
$$E[z_{ink}]=0*P(z_{ink}=0\|y_{ink},\theta)+1*P(z_{ink}=1\|y_{ink},\theta)$$ $$E[z_{ink}]=P(z_{ink}=1\|y_{ink},\theta)$$	9

Estimation number of the defective pieces $$e_{ijk} = E[z_{ink}] \times b_{ij}$$	10
Set of *(i,j,k)* indexes $$S_m = \{(i,j,k) \mid \forall m_{ik} = m\}$$	11
expected quantity of defective pieces $$d_m = \sum_{(i,j,k)\in S_m} e_{ijk}$$	12
the quantity of potential defective pieces $$h_{ijk} = \begin{cases} 0; & if\ j = k \\ h_{ij;k+1} + e_{ijk}; & if\ j > k \end{cases}$$	13
the quantity of all potential defective pieces $$x_m = \sum_{(i,j,k)\in S_m} h_{ijk}$$	14
expected quantity of all good pieces $$g_m = x_m + \left(\sum_{i=1}^{I} f_i \cdot r_{im}\right)$$	15
Batch yield accuracy $$\overline{acc} = \frac{\sum_{i=1}^{I}\left(1 - \left|\left(\prod_{j=1}^{J_i} P_{ij}\right) - l_i\right|\right)}{I}$$	16