mathematics

Article

A Machine Learning Approach for Improving Wafer Acceptance Testing Based on an Analysis of Station and Equipment Combinations

Chien-Chih Wang * and Yi-Ying Yang

Department of Industrial Engineering and Management, Ming Chi University of Technology, New Taipei City 24303, Taiwan
* Correspondence: ieccwang@mail.mcut.edu.tw; Tel.: +886-2-2908-9889

Abstract: Semiconductor manufacturing is a complex and lengthy process. Even with their expertise and experience, engineers often cannot quickly identify anomalies in an extensive database. Most research into equipment combinations has focused on the manufacturing process's efficiency, quality, and cost issues. There has been little consideration of the relationship between semiconductor station and equipment combinations and throughput. In this study, a machine learning approach that allows for the integration of control charts, clustering, and association rules were developed. This approach was used to identify equipment combinations that may harm production processes by analyzing the effect on Vt parameters of the equipment combinations used in wafer acceptance testing (WAT). The results showed that when the support is between 70% and 80% and the confidence level is 85%, it is possible to quickly select the specific combinations of 13 production stations that significantly impact the Vt values of all 39 production stations. Stations 046000 (EH308), 049200 (DW005), 049050 (DI303), and 060000 (DC393) were found to have the most abnormal equipment combinations. The results of this research will aid the detection of equipment errors during semiconductor manufacturing and assist the optimization of production scheduling.

Keywords: DRAM manufacturing; statistical quality control; clustering; associative analysis; case study

MSC: 62P30

check for updates

Citation: Wang, C.-C.; Yang, Y.-Y. A Machine Learning Approach for Improving Wafer Acceptance Testing Based on an Analysis of Station and Equipment Combinations. *Mathematics* **2023**, *11*, 1569. https://doi.org/10.3390/math11071569

Academic Editors: Ravil Muhamedyev, Evgeny Nikulchev and Stelios Psarakis

Received: 2 January 2023
Revised: 1 March 2023
Accepted: 22 March 2023
Published: 23 March 2023

1. Introduction

Semiconductor manufacturing is a very competitive and capital- and technology-intensive industry. As the demand for electronic components that are light and thin continues to increase, the size of the wafers used for integrated circuits increases and the line width per component decreases, thus allowing an increase in chip density, making it necessary to optimize quality to meet customer demand continuously [1–3]. The manufacturing of semiconductor wafers typically consists of more than 500 steps and takes two months to complete, and each of these steps must be closely monitored to ensure that the error associated with each step remains within the allowable limit [4,5]. Several uncertainties exist in the manufacturing process, such as those pertaining to people, machines, materials, methods, environments, and measurements [6]. Therefore, semiconductor manufacturing faces more constraints and difficulties than other industries in relation to quality control and yield improvement. The high degree of precision demanded by semiconductor manufacturing and the low tolerance for defects mean that the efficiency of the manufacturing process can be assessed only after all steps in the production process have been completed [7–9]. Thus, ensuring process stability requires more resources to be invested in process management and data monitoring.

This study is based on the practical issues raised by DRAM manufacturers in Taiwan. DRAM is commonly used in modern computers, smartphones, and other electronic devices.

Dynamic Random Access Memory (DRAM) is a volatile memory that stores data temporarily and requires constant refreshing to maintain data integrity. The capacitors in DRAM cells store electrical charge to represent binary bits of data, with the presence or absence of charge indicating a 1 or 0 value, respectively. However, due to the nature of capacitors, they tend to leak charge over time, leading to memory loss. As a result, DRAM cells require periodic refreshing to maintain their data integrity. This process involves reading and rewriting the data in each memory cell, effectively recharging the capacitors to ensure the correct values are stored [10,11]. DRAM manufacturing is a complex and precise process requiring significant expertise. Each step must be carefully controlled to ensure the final product meets specific design specifications and precision to produce reliable, high-quality memory modules.

DRAM products are manufactured in batches and processed using photo, etching, diffusion, and thin film. There are three main quality issues associated with DRAM. The first is the quality of the design. Based on the TRIZ theory and verification experiments, Wang and Lo tried to optimize the quality of the leakage current. The study found that using different angles of bungee ion implantation can improve the gate-derived drain leakage current. Moreover, this was further verified by the experimental design when the tilt angle of 21 degrees can reach the industry's expected improvement goals [12]. The second issue is the manufacturing quality. Six Sigma, Design of Experiments, and SPC techniques are the main methods used to identify critical causes and to optimize the manufacturing process [13–18]. Real-time process-control methods include statistical process-control and engineering process-control methods. In contrast to these methods, post-mortem methods require data, such as the results of wafer acceptance testing (WAT), parameter analysis, and wafer map analysis, to be available after all the manufacturing processes have been completed. When the yield of a batch is too low, data collected during the manufacturing process are used to determine if a machine anomaly caused the failure. In the case of semiconductor data, engineers often cannot quickly identify the possible causes of abnormalities or the factors causing poor product quality due to a large amount of information involved and its complexity. Engineers also tend to rely a lot on their experience and can sometimes miss problems hidden in data if these problems do not conform to their preconceived ideas. Fault detection based on machine learning models and deep learning models is also used to analyze sensor data and detect abnormalities during the early stages of processing [19–22]. The third issue is the inspection quality. In addition to using specialized measuring equipment, this study focused on detecting defects using machine vision.

Wafer acceptance testing is the most common method of determining semiconductor yield [23,24]. A lack of analytical tools and statistical concepts often prevents engineers from quickly identifying the possible causes of anomalies in large amounts of complex engineering data or from summarizing the characteristics of poor-quality products. Thus, yield management has become increasingly concerned with providing engineers with a basis for problem-solving by converting large amounts of engineering data into valuable knowledge through effective analysis. In the production of semiconductors, the testing of wafer pins is a time-consuming part of post-production quality checks; however, it does not improve the quality of the product. Therefore, how to effectively manage and improve the yield rate, or even how to predict the final yield early enough so that remedial action can be taken and the time and cost involved in subsequent testing reduced, is the main reason for studying yield management. A wafer acceptance test is the last control point before pin testing for a completed wafer. Therefore, by analyzing the parameters of the wafer acceptance test that significantly impact wafer pin test yields, we can help wafer manufacturers to predict yield levels and identify problems early, analyze and improve quickly, and reduce the amount of variability in their products. This will also reduce the time and money spent on wafer pin testing and improve customer satisfaction.

There are many reasons for the failure of semiconducting devices, among which the effect on the yield of the combination of equipment used in the production process can be considered a stochastic effect. Most research into equipment combinations has focused on the manufacturing process's efficiency, quality, and cost. Suh et al. use a genetic algorithm to optimize the layout of a fabrication facility, particularly in the context of Fab facilities, which require careful consideration of material handling and flow [25]. As part of the semiconductor manufacturing process, equipment combinations refer to which machines are selected for production and how they are scheduled. According to practical experience, equipment combinations significantly impact semiconductor manufacturing and can directly affect the process's efficiency, quality, and cost. In addition to improving product yields and quality and reducing costs, Ghasemi et al. propose a machine scheduling method that optimizes a manufacturing process to minimize process time and cost [26]. Uzsoy et al. propose a new method for maximizing machine resources and improving productivity and yield by adjusting the scheduling and assignment in the mask manufacturing process [27].

In semiconductor manufacturing, equipment combinations were once considered a random factor, and few studies examined the relationship between yield and equipment combinations or station groupings. However, as semiconductor manufacturing technology advances, the effect of equipment combinations on yield can be quantified and predicted. Some studies have shown that equipment interaction and interference during manufacturing can increase defective rates and that good equipment pairing can reduce this effect [28,29]. Therefore, choosing the right combination of equipment is essential to ensure high quality and yield. This paper focused on a DRAM semiconductor factory in Taiwan. The aim was to apply a machine-learning approach that allows for the integration of control charts and clustering, and association rules were developed. This approach was used to identify equipment combinations that may harm production processes by analyzing the effect of the equipment combinations used in wafer acceptance testing (WAT) on Vt parameters and thus improve product quality.

2. Materials and Methods

In relation to manufacturing, machine learning approaches can be used to identify bottlenecks in research and development, design, production, and sales. In this study, data-based machine learning technology was developed to determine the impact of different equipment combinations on the quality of DRAMs at different stages of the manufacturing process. These results could then be used to improve production schedules. A model for recommending equipment combinations was produced by combining control charts and machine-learning techniques. Figure 1 shows a flowchart that describes the proposed method.

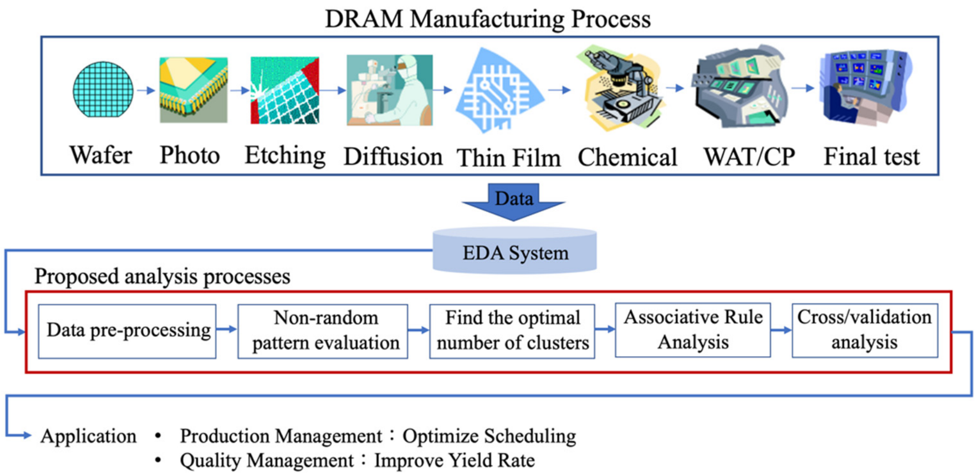

Figure 1. The analysis framework flowchart.

In the production of DRAM devices, the wafers are collected by the engineering data-analysis system, the equipment used in each process, and the final WAT values are determined. In this study, the data were pre-processed and missing values and anomalies were deleted directly. Based on a discussion with the senior engineer, the missing values are due to the fact that the wafer sampling process uses lot number sampling instead of a full inspection, which means that the measured value is not recorded in the EDA system if the data belong to a non-sampled lot number. As a result of system conversion problems, extreme values, such as a yield value of −99, are generated during the capture process, which will be considered abnormal. In the second step, a run chart technique was used to check whether the data had a trend or a non-random trend in the cluster. Run charts and corresponding statistical testing techniques are used to determine whether there are any trends or clusters in the data. If any trends were identified, the situation was discussed with an engineer, and the process improved until no non-random trends could be identified. If any trends were identified, the situation was discussed with an engineer, followed by root cause analysis techniques implemented to improve the process until no non-random trends were detected. The third step was to find the optimal number of bins. The optimal number of bins within a reasonable control range was determined using the individual control-boundary clustering method and the k-times standard deviation binning method. The next step was to use the determined optimal number of clusters to find the best association rule using the association-rule analysis method. The final step consisted of cross-validation of the results. The recommended equipment combinations were discussed with senior engineers to confirm the correctness of the equipment at each station. This was carried out to improve the accuracy of the analysis. Cross-validation is determined by comparing the equipment with fine-tuning and calibration recorded in the system with the equipment resulting from this study.

2.1. Run Chart Technique

A run chart displays the evolution of data during a process and can be used to identify the causes of variations in the data [30]. In this study, the Vt values of equipment at different stations, which corresponded to one or more time-variation and process relationships, were analyzed. To determine the presence of non-random trends in WAT measurements, we used a run chart to analyze the Vt values, perform data preprocessing to determine the presence of trends or clustering in the data, eliminate noise in the data, and highlight the effect of equipment combinations on Vt values.

For normally distributed data, Z_1 statistic can be used to check for non-random clustering [31]. If the *p*-value, which is equal to cdf(Z_1), where cdf denotes the cumulative probability of Z_1, is less than a specified value, a tendency for clustering is indicated. Z_1 is given by

$$Z_1 = \frac{R - 1 - (2mn\,/\,N)}{\sqrt{\frac{2mn(2mn-N)}{N^2(N-1)}}} \tag{1}$$

Non-random patterns can also be checked for using the test statistic Z_2 [31]. In this case, *p*-value = cdf(Z_2), where cdf denotes the cumulative probability of Z_2. Again, a *p*-value that is less than a specified value indicates a trend. Z_2 can be calculated as

$$Z_2 = \frac{V - \frac{2N-1}{3}}{\sqrt{\frac{16N-29}{90}}} \tag{2}$$

where V = number of runs up or down and N = total number of points.

2.2. Confirmation of the Number of Clusters

A control chart can be used to perform the clustering of WAT values. In this study, we used the individual control chart–boundary clustering method and the k-times standard deviation clustering method to determine the optimal number of clusters within a reasonable control range. The reason for using individual control charts was that each wafer is measured and divided into three groups according to the control limits. An individual control chart can be used to determine the mean and variance of a set of values. The following steps are used to build an individual control chart [32].

Step 1. Calculate the moving range (MR). The MR is calculated by calculating the distance between two adjacent points:

$$MR_i = |x_i - x_{i-1}| \tag{3}$$

Step 2. Calculate the average and the average of the moving range:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i}{k} \text{ and } \overline{MR} = \frac{MR_i}{k-1} \tag{4}$$

Here, x_i is the observed value of the ith sample and k is the number of samples used to construct the control chart.

Step 3. Construct the individual control charts:

$$\text{individual control chart} = \begin{cases} UCL = \bar{x} + 3\frac{\overline{MR}}{d_2} \\ CL = \bar{x} \\ LCL = \bar{x} - 3\frac{\overline{MR}}{d_2} \end{cases} \tag{5}$$

where, d_2, D_3, and D_4 are control chart constants [32].

The k-times standard deviation grouping method is based on the concept of a control chart that divides the data into k clusters by using the mean as the center point together with the upper and lower k-times standard deviation techniques. The determination of the number of clusters is limited by whether there is a significant difference between the groups in the delineated data. The determination of the optimal grouping, k, is made using Analysis of Variance (ANOVA) with the null hypothesis (H_0) and alternative hypothesis (H_1) as follows.

$$\begin{cases} H_0 : \text{There are no significant differences between the k clusters.} \\ H_1 : \text{There are significant differences between the k clusters.} \end{cases} \tag{6}$$

If the p-value for the above test is less than or equal to a statistically significant level, this means there is a significant difference between the clusters that can be used for the association-rule analysis. If the p-value is greater than this significant level, H_0 is not rejected, as this means that there are no significant differences between the clusters and that the clusters without significant differences between them should be merged. The differences between the clusters should be checked repeatedly until they are found to be significant; the merging of the groups should then be stopped. Following this, the association-rule analysis can be performed.

2.3. Association-Rule Analysis

The Apriori algorithm has been shown to be extremely useful for discovering previously unknown relationships in data sets by finding rules and associations between attributes [33–35]. Association-rule analysis is used in the manufacturing industry not only for marketing, inventory management, and storage analysis, but also for failure analysis, process capability analysis, etc. Important factors affecting manufacturing processes and yields can be extracted so that relevant parameters can be adjusted and equipment used more efficiently, thus improving yields and productivity and reducing production costs.

In the manufacture of DRAM devices, more than one piece of equipment is used in each part of the production process. Since there is a back-and-forth relationship between the different stages of the production process, in this study, we used the association-rule algorithm to establish a relationship between the equipment combinations used at different stages and the results of the WAT. This allowed us to develop an association rule for the equipment combinations that will produce better-quality products.

The measures of support, confidence, and lift, which are defined below, were used to generate the association-rule metrics [36].

- The support of A ⇒ B is calculated as the percentage of transactions in the database that contain both A and B:

$$\mathrm{support}(\mathrm{A} \Rightarrow \mathrm{B}) = P(\mathrm{A} \cap \mathrm{B}) = \frac{\text{number of transactions containing both A and B}}{\text{total number of transactions}} \quad (7)$$

- The confidence of A ⇒ B is determined by calculating the percentage of transactions in the database that contain both A and B simultaneously:

$$\mathrm{confidence}\ (\mathrm{A} \Rightarrow \mathrm{B}) = P(\mathrm{B} \mid \mathrm{A}) = \frac{P(\mathrm{A} \cap \mathrm{B})}{P(\mathrm{A})} = \frac{\text{number of transactions containing both A and B}}{\text{number of transactions containing A}} \quad (8)$$

- The lift measures the degree of independence or dependence between A and B. If a rule has a lift of 1, A and B are independent and no rule containing either event will be generated. A and B are codependent and positively correlated whenever a rule has a lift greater than 1. Generally, rules with high support and confidence are preferred in practice:

$$\mathrm{lift}(\mathrm{A}, \mathrm{B}) = \frac{P(\mathrm{A} \cup \mathrm{B})}{P(\mathrm{A})P(\mathrm{B})} \quad (9)$$

The Apriori algorithm was first proposed by Agrawal et al. [37]. Subsequently, the wide-search algorithms, Partition Apriori, DHP (standing for direct hashing and pruning), and MSApriori, were developed based on the original algorithm [38,39]. In addition, there is also the Depth-first FP growth [40]. In this study, we focused on the interpretation and application of significance rules, and the Apriori algorithm was used when there was no significant difference in the efficiency and effectiveness of the problem to be solved. The Apriori algorithm consists of the following steps.

Step 1. A candidate item set is formed by combining the (k − 1)th item set obtained from the previous iteration with the kth item set.
Step 2. The support for each candidate item set is calculated. By scanning the database, the number of transactions containing all the items from each of the candidate item sets is determined.
Step 3. A high-frequency pattern is set. A high-frequency pattern containing k items is determined from the candidate sites whose support is greater than the minimum support.
Step 4. The process is stopped if no new high-frequency pattern is found. Otherwise, the process is repeated from Step 1.

3. Case Analysis and Discussion

In this study, data from a DRAM semiconductor manufacturing plant in Taiwan were used. Because of the high production rate and the need to replace equipment in this plant, the equipment used at each processing station are not of the same brand. To confirm its quality, the equipment was tested and evaluated several times by different manufacturers before it was officially used. Furthermore, the degree of wear on the equipment has an impact on how much of the final yield has a quality that falls within the control range. As a result, the senior engineer who provided the data suggested that differences in equipment brands could be ignored. In this study, we have treated the newness and age of each piece of equipment and the brand the same.

The processing data were discussed with the senior engineers at the factory to coordinate the capture of the data fields that were required for verification. The Vt values from the EDA database were retrieved from the results of the wafer acceptance test for DRAM production yield. The senior engineers reviewed the results and then removed the stations that they considered had less effect on the results: 39 stations and 15,628 data remained for analysis. The results of partial data are shown in Table 1. There are some missing values where batch-number sampling rather than a full inspection was used for the wafer sampling. The EDA system did not record values if the captured data belonged to a non-sampled lot number.

Table 1. The results of the collected partial data.

Lot Number	Station and Equipment Number (1)	……	Station and Equipment Number (39)	Vt Value
P00000001A	000001/DC393	……	0012034/DD249	0.3012
P00000002A	000102/EH132	……	0490500/DI101	0.3212
P00000012A	001022/DD246	……	0001500/DC521	0.3211
⋮	⋮	……	⋮	⋮

In this case, a prerequisite for the use of association-rule analysis was that there was no time factor between Vt values. Therefore, a run chart was used to check for trends and patterns in in the sample data. The results of this are shown in Figure 2.

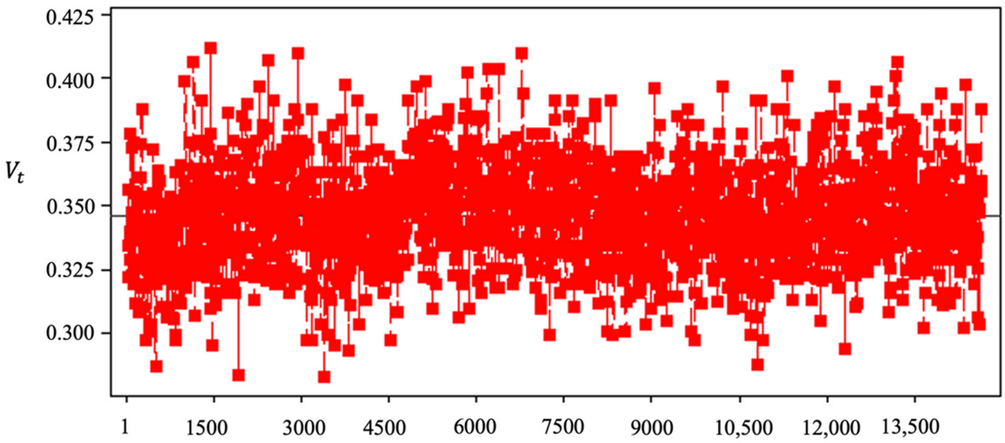

Figure 2. The run chart result for Vt values.

According to the statistical analysis, the *p*-values for the trend and the cluster non-randomness were 0.0012 and 0.0002, respectively. This meant that the trend and the cluster non-randomness had no effect on the Vt values because both were smaller than the assumed significance level of 0.05.

3.1. Association-Rule Analysis Using Individual Control-Boundary Clustering

An individual control-boundary was used to calculate the Vt values; for this, the upper and lower bounds were set as 0.36921 and 0.2999, respectively. The Apriori algorithm was then applied to the three clusters of data (corresponding to data above the upper bound, between the upper and lower bounds, and below the lower bound). We set values of support > 80%, 75% < support < 80%, 70% < support < 75%, and confidence > 85% to generate the rules. The results are shown in Table 2.

From Table 2, it can be seen that there are some cases where the Vt value exceeds the upper limit: for example, machine DI303 at station 049050 and machine DC393 at station 060000 should be checked, calibrated, and serviced first. If the support and confidence requirements are low, machines EH464, EH308, DD249, DI303, and DC393 at the five stations 02100, 046000, 047500, 049050, and 060000 should be inspected, calibrated, and serviced first. Since each station code corresponds to a particular machining sequence, the interactions between each station listed in Table 2 can be seen in Figure 3 for cases where the support is between 70% and 75% and the confidence level is greater than 85%.

Table 2. Results of association rules using individual control-boundary clustering.

Support	Confidence	beyond the Upper Bound	between the Upper and Lower Bounds	beyond the Lower Bound
>80%	>85%	049050 (DI303) 060000 (DC393)		
75~80%	>85%		049050 (DI303) 060000 (DC393)	
70~75%	>85%	021700 (EH464) 047500 (DD249) 049050 (DI303) 046000 (EH308) 060000 (DC393)		

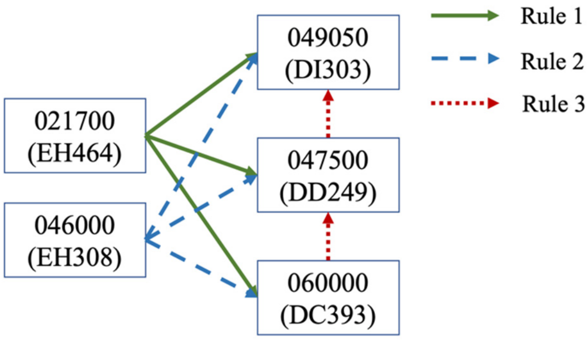

Figure 3. Rules for station–equipment combinations with support rates between 70% and 75% and confidence levels greater than 85%.

Figure 3, Rule1, Rule2, and Rule3 show the machine interactions at each station. Station 049050 is affected by machine EH464 at station 02100 and machine EH308 at station 046000, resulting in abnormal final Vt values. Machine DD249 at station 047500 will be affected by machine EH464 at station 021700 and machine EH308 at station 046000. The yield will also be affected by the abnormal yield of machines DI303 and DC393 at sites 049050 and 060000.

Based on the knowledge that the production of semiconductors requires a high degree of precision, the association rules for highly correlated station-specific equipment groups were devised. According to these rules, if an abnormality occurs at a particular station, it must have been caused by one of the earlier stations. The station–equipment association rules generated from the Vt anomalies shown in Table 2 and Figure 3 were organized into the following four groups.

1. Group 1 (049050/DI303 and 060000/DC393). Whenever an abnormality occurs in machine DC393 at station 060000, an abnormality must first have occurred in machine DI303 at station 049050.
2. Group 2 (02100/EH464, 047500/DD249, 049050/DI303, and 060000/DC393). When the tolerance range of the association rule is extended, machine EH464 at station 021700 will affect machines DD249, DI303, and DC393 at later stations. Station–equipment combinations generated by this rule are susceptible to anomalies where the final Vt value exceeds the control limit.
3. Group 3 (046000/EH308, 047500/DD249, 049050/DI303, and 06000/DC393). The Vt values of machine DD249 at station 047500, machine DI303 at station 049050, and machine DC393 at station 0600 are affected by machine EH308 at station 046000.
4. Group 4 (047500/DD249, 049050/DI303, and 060000/DC393). Machine DI303 at station 049050 will also affect the Vt values between machine DD249 at station 047500 and machine DC393 at station 060000.

Using the individual control-boundary clustering method, it is possible to quickly determine which stations and equipment have the greatest influence on each other and to map the various relationships between the groups of equipment and stations. From this and the association rules, engineers can also quickly find out which pieces of equipment with abnormal Vt values need to be calibrated, thus saving time that would otherwise have been wasted on unnecessary error detection.

The processed data were discussed with the senior engineers at the factory to coordinate the capture of the data fields that were required for verification. The Vt values from the EDA database were retrieved from the results of the wafer acceptance test for DRAM production yield. The senior engineers reviewed the results and then removed the stations that they considered had less effect on the results: 39 stations and 15,628 data remained for analysis. Some of the data are shown below.

3.2. Association-Rule Analysis Using k-Times Standard Deviation Clustering

The Vt values that had been found previously were grouped into six clusters: (1) Vt > $\overline{x}_{V_t} - 2S_{V_t}$; (2) $\overline{x}_{V_t} - 2S_{V_t} <$ Vt $< \overline{x}_{V_t} - 1S_{V_t}$; (3) $\overline{x}_{V_t} - 1S_{V_t} <$ Vt $< \overline{x}_{V_t}$; (4) $\overline{x}_{V_t} <$ Vt $< \overline{x}_{V_t} + 1S_{V_t}$; (5) $\overline{x}_{V_t} + 1S_{V_t} <$ Vt $< \overline{x}_{V_t} + 2S_{V_t}$; and (6) Vt > $\overline{x}_{V_t} + 2S_{V_t}$. Next, ANOVA was used to determine whether there was a difference between the six clusters that had been defined. The results gave a p-value of 0.002 (<0.05) and an R^2 of 92.28%, which meant that the six clusters were well able to explain the variation in the Vt values and constituted the most suitable subgroups. The Apriori algorithm was then applied to the six data clusters. The levels of the support and confidence were set as support > 80%, 75% < support < 80, 70% < support < 75%, and confidence > 85% to generate the association rules. The results are shown in Table 3.

Table 3. Results of association rules using k-times standard deviation clustering.

Support	Confidence	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
>80%	>85%				049050 (DI303) 060000 (DC393)	049050 (DI303) 060000 (DC393)	049050 (DI303) 060000 (DC393)
75~80%	>85%				049050 (DI303) 060000 (DC393)		047500 (DD249) 049050 (DI303) 060000 (DC393)
70~75%	>85%			049050 (DI303) 060000 (DC393)		047500 (DD249) 049050 (DI303) 060000 (DC393)	047500 (DD249) 049050 (DI303) 060000 (DC393)

The support and confidence levels were then relaxed to 70% < support < 75% and confidence > 85%. An association-rule group composed of three stations (machines) 047500 (DD249), 049050 (DI303), and 060000 (DC393) was generated, as shown in Figure 4.

Figure 4 shows that machine DI303 at station 049050 is affected by both machine EH308 at station 046000 and machine DD249 at station 047500 and that it indirectly affects machine DC393 at the later station 060000. This means that if the Vt value is abnormal at machine DI303 at station 049050, this must have been caused by machine EH308 at station 046000 and machine DD249 at station 047500. These latter two stations will also indirectly affect machine DC393 at station 060000.

From the results of the individual control-boundary clustering and the k-times standard deviation clustering, it was concluded that the higher the level of support and confidence, the smaller the number of association-rule groups and the more significant the association between station groups. However, as the level of support and confidence set by the association rules decreased, the application of the rules caused new groups of stations and equipment to be slowly added. The association rules with more significant effects then gradually moved outside the control limits.

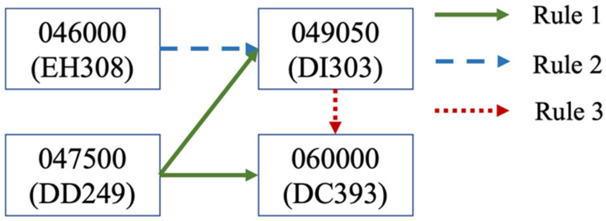

Figure 4. Rules for station–equipment combinations with 70% < support < 75% and confidence > 85% using k-times standard deviation clustering.

3.3. Validation of the Results

The procedures described in Sections 3.1 and 3.2 were then applied to the data again. First, the Vt run chart was examined for the presence of trends and clusters in the sample data. The results of a statistical analysis showed that the *p*-values for the trend and cluster non-randomness were 0.0001 and 0.00003, respectively. This meant that the trend and cluster non-randomness had no effect on the Vt values. Association-rule analysis was then applied to the new data, and the results were compared with those obtained previously. The new results are shown in Table 4.

Table 4. Comparison of association rules obtained using new data with those obtained previously.

	Individual Control Boundary Cluster Method		**K Times Standard Difference Cluster Method**	
	First Data Set	**New Data Set**	**First Data Set**	**New Data Set**
Support	70~75%	75~80%	70~75%	75~80%
Station and equipment combinations	021700 (EH464) 047500 (DD249) 049050 (DI303) 046000 (EH308) 060000 (DC393)	018100 (EH463) 046000 (EH308) 047800 (DI206) 049200 (DW005)	047500 (DD249) 049050 (DI303) 060000 (DC393)	018100 (EH463) 046000 (EH308) 047800 (DI206) 049200 (DW005)

From an analysis of these results, two main conclusions can be drawn.

(1) In the results previously obtained using the individual value control-boundary clustering method, the groups of stations that had Vt values beyond the control limit were mostly concentrated in the latter part of the production process. Additionally, most of the generated association rules were related to a small number of equipment and stations: machine DI303 at station 049050, machine DC393 at station 060000, and machine DW005 at station 046000. The results of the new data analysis show machine EH308 at station 046000 and machine DW005 at station 049200.

(2) The clustering results obtained using the individual value control-boundary clustering method are better than the association rules generated by the k-times standard deviation clustering method. This suggests that it is better to implement the association rules after clustering has been performed using individual control-boundary clustering.

4. Discussion

According to the initial settings of this study, the support level was between 70% and 80%, and the confidence level was greater than 85%. Using the I-MR control limit method before improving the process target to group station-specific equipment was more significant than using the k-fold standard deviation method. As the equipment's calibration and the target value improve, the final association rule groups tend to be the same for both the I-MR control boundary grouping method and the k-fold standard difference grouping method. Cross-evaluation of the two actual process data obtained at different time points revealed that the results of the I-MR control boundary clustering and k-fold standard deviation clustering were more significant than the results of the other two clusters in the selected range of 75% to 80% for the support and 85% for the confidence of the associated rule level.

From the 39 stations provided by the engineers, this study proposes a method that allows us to quickly select the 13 stations and equipment with the greatest impact on the process. Figure 5 shows the association rules arranged in the order of the stations. The station groups are indicated in yellow; the equipment at individual stations that affect each other are indicated in red. The correlation rules generated in this study identified four stations—046000, 049200, 049050, and 060000—as being the most likely to affect the yield of the DRAM semiconductor manufacturing plant. In the plant, more than one piece of equipment was located at each station. The results of the association-rule analysis showed that the most significant machines at these four stations were EH308, DW005, DI303, and DC393, respectively.

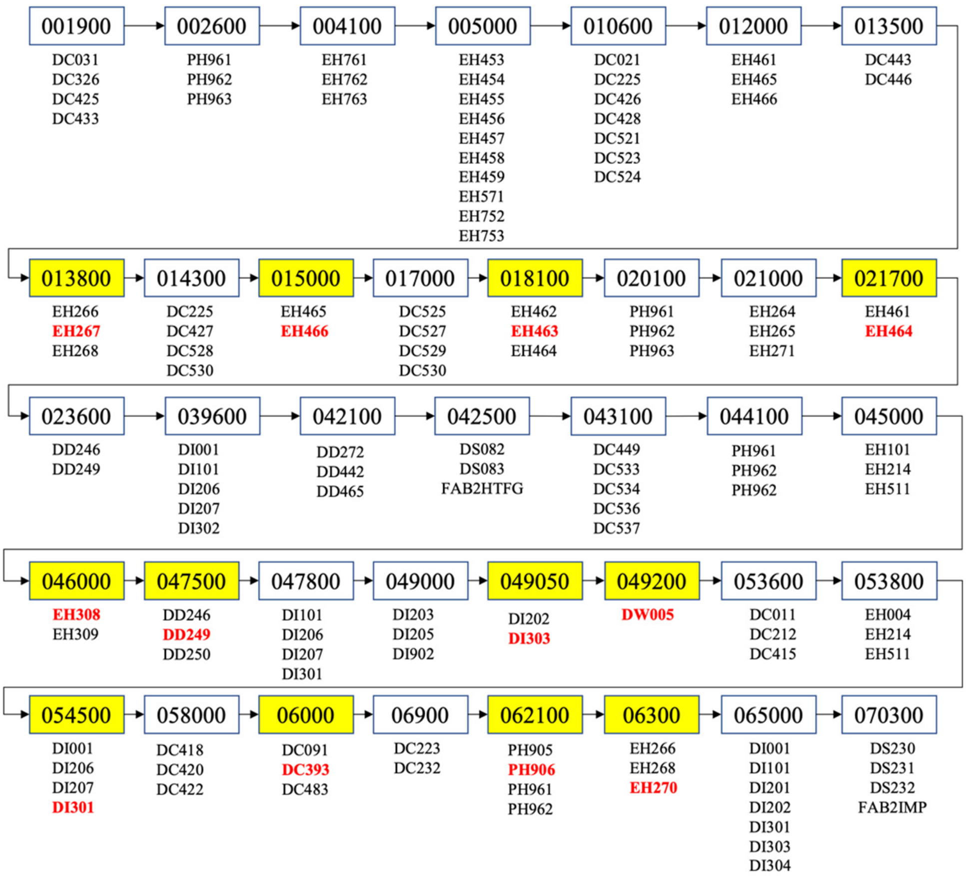

Figure 5. The results of the association rules are arranged in order of the stations.

The results shown in Figure 5 were confirmed by consultation with senior engineers. The results for the most significant equipment that had been extracted from the results for the 39 stations were consistent with the actual conditions at the plant, thus confirming that the combinations of equipment and stations that were identified do in fact impact the plant's yield. The two clustering methods proposed in this study are currently only available to the analytical engineers in a semiconductor manufacturing plant for their judgment. Through this study, engineers at a DRAM semiconductor plant can use the groupings established in this study to quickly identify associated groups of equipment from an extensive database when tracing process anomalies, thus improving the time spent in the traditional process of using various statistical charts to determine defect rates and tracing the affected stations from scratch. Furthermore, the support and confidence in the correlation rule analysis are not fixed values. In performing the correlation rule analysis, the selection criteria must consider the degree of correlation of the final rule results, i.e., the degree of influence of the sequence of occurrence of individual equipment. In this study, it is suggested that, in general, when using correlation rules for analysis, one can set support to be between 70% and 80% and confidence > 85%, and then adjust the judgement level of support and confidence according to the correlation rule results after clustering to obtain the best correlation rule clusters.

5. Conclusions

This study focused on the station–equipment combinations used in the production of DRAM devices and the final, measured Vt values. This contrasts with previous studies, where station–equipment combinations were considered a random factor and their influence was ignored. The results of an empirical analysis showed that the best association rules could be obtained by setting the support at between 70% and 80% and the confidence at over 85% for the individual control-boundary clustering, and then using the association rules.

The association-rule analysis could quickly identify the most influential stations among the major stations and determine the correlation and association rules between the equipment. This means that, using this technique, the time that might otherwise have been spent applying various statistical techniques and control charts to determine which of the stations at a DRAM semiconductor manufacturing plant are abnormal can be saved. In addition, the association rules for the abnormal stations can be used to detect errors in the manufacturing process. These rules can also be used as a reference by process engineers and in production scheduling to reduce the occurrence of errors.

Author Contributions: Conceptualization, C.-C.W. and Y.-Y.Y.; methodology, C.-C.W. and Y.-Y.Y.; validation, C.-C.W. and Y.-Y.Y.; formal analysis, C.-C.W. and Y.-Y.Y.; data curation, Y.-Y.Y.; writing—original draft preparation, C.-C.W. writing—review and editing, C.-C.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, J.; Wang, J.; Wu, J.; Qi, H.; Wang, C.; Fang, X.; Cheng, C.; Yang, W. Rapid detection of ultra-trace nanoparticles based on ACEK enrichment for semiconductor manufacturing quality control. *Microfluid. Nanofluidics* **2019**, *23*, 2. [CrossRef]
2. Senoner, J.; Netland, T.; Feuerriegel, S. Using explainable artificial intelligence to improve process quality: Evidence from semiconductor manufacturing. *Manag. Sci.* **2022**, *68*, 5704–5723. [CrossRef]
3. Fan, S.K.S.; Lin, W.K.; Jen, C.H. Data-driven optimization of accessory combinations for final testing processes in semiconductor manufacturing. *J. Manuf. Syst.* **2022**, *63*, 275–287. [CrossRef]
4. May, G.S.; Spanos, C.J. *Fundamentals of Semiconductor Manufacturing and Process Control*; John Wiley & Sons: Hoboken, NJ, USA, 2006.
5. Espadinha-Cruz, P.; Godina, R.; Rodrigues, E.M. A review of data mining applications in semiconductor manufacturing. *Processes* **2021**, *9*, 305. [CrossRef]
6. e Oliveira, E.; Miguéis, V.L.; Borges, J.L. Automatic root cause analysis in manufacturing: An overview & conceptualization. *J. Intell. Manuf.* **2022**, 1–18. [CrossRef]
7. Chien, C.F.; Hsu, C.Y.; Morrison, J.R.; Dou, R. Semiconductor manufacturing intelligence and automation. *Comput. Ind. Eng.* **2016**, *99*, 315–317. [CrossRef]
8. Yoon, H.J.; Chae, J. Simulation Study for Semiconductor Manufacturing System: Dispatching Policies for a Wafer Test Facility. *Sustainability* **2019**, *11*, 1119. [CrossRef]
9. Lin, J.; Li, Y.Y.; Song, H.B. Semiconductor final testing scheduling using Q-learning based hyper-heuristic. *Expert Syst. Appl.* **2022**, *187*, 115978. [CrossRef]

10. Thakur, R.P.S.; DeBoer, S.J.; Ping, E.X.; Jesse, C.A. Process simplification in DRAM manufacturing. *IEEE Trans. Electron Devices* **1998**, *45*, 609–619. [CrossRef]
11. Qian, Z.; Wei, J.; Xiang, Y.; Xiao, C. A performance evaluation of DRAM access for in-memory databases. *IEEE Access* **2021**, *9*, 146454–146470. [CrossRef]
12. Wang, C.C.; Lo, K.L. Practical study on optimizing the quality of dynamic random-access memory for improving leakage current. *J. Qual.* **2019**, *26*, 353–365.
13. Lu, C.H.; Chang, W.F. Optimizing the Process Window of Bond Line Thickness for Printable Die Attach Adhesive in DDR DRAM Packaging. In Proceedings of the 2010 11th International Conference on Electronic Packaging Technology & High Density Packaging, Xi'an China, 16–19 August 2010; pp. 909–915.
14. Kim, Y.K.; Min, K.K.; Park, B.-G. Trap-induced data-retention-time degradation of DRAM and improvement using dual work-function metal gate. *IEEE Electron Device Lett.* **2020**, *42*, 38–41. [CrossRef]
15. Yousaf, J.; Faisal, M.; Youn, J.; Nah, W. Design of experiment (doe) analysis of system level esd noise coupling to high-speed memory modules. *Electronics* **2019**, *8*, 210. [CrossRef]
16. Leu, Y.; Lin, C.-M.; Yang, W.-N. Reducing Thickness Deviation of W-Shaped Structures in Manufacturing DRAM Products Using RSM and ANN_GA. *IEEE Trans. Compon. Packag. Manuf. Technol.* **2021**, *11*, 899–910. [CrossRef]
17. Sparsh, M.; Inukonda, M.S. A survey of techniques for improving error-resilience of DRAM. *J. Syst. Archit.* **2018**, *91*, 11–40.
18. Qin, R.; Qiu, S.; Xia, Y.; Hu, S.; Chang, J.; Zhang, J.; Zhang, W.; Wang, P.; Zhou, X.; Bitincka, E.; et al. Fast In-Device Overlay Metrology on DRAM Storage Node Contact and Its Applications in Process Control. In Proceedings of the 2021 International Workshop on Advanced Patterning Solutions (IWAPS), Foshan, China, 12–13 December 2021.
19. Fan, S.K.S.; Lin, S.C.; Tsai, P.F. Wafer fault detection and key step identification for semiconductor manufacturing using principal component analysis, AdaBoost and decision tree. *J. Ind. Prod. Eng.* **2016**, *33*, 151–168. [CrossRef]
20. Wang, T.; Qiao, M.; Zhang, M.; Yang, Y.; Snoussi, H. Data-driven prognostic method based on self-supervised learning approaches for fault detection. *J. Intell. Manuf.* **2018**, *31*, 1611–1619. [CrossRef]
21. Kim, E.; Cho, S.; Lee, B.; Cho, M. Fault detection and diagnosis using self-attentive convolutional neural networks for variable-length sensor data in semiconductor manufacturing. *IEEE Trans. Semicond. Manuf.* **2019**, *32*, 302–309. [CrossRef]
22. Hsu, C.Y.; Liu, W.C. Multiple time-series convolutional neural network for fault detection and diagnosis and empirical study in semiconductor manufacturing. *J. Intell. Manuf.* **2021**, *32*, 823–836. [CrossRef]
23. Xu, H.; Zhang, J.; Lv, Y.; Zheng, P. Hybrid feature selection for wafer acceptance test parameters in semiconductor manufacturing. *IEEE Access* **2020**, *8*, 17320–17330. [CrossRef]
24. Fan, S.K.S.; Cheng, C.W.; Tsai, D.M. Fault diagnosis of wafer acceptance test and chip probing between front-end-of-line and back-end-of-line processes. *IEEE Trans. Autom. Sci. Eng.* **2021**, *19*, 3068–3082. [CrossRef]
25. Suh, Y.J.; Choi, J.Y. Efficient Fab facility layout with spine structure using genetic algorithm under various material-handling considerations. *Int. J. Prod. Res.* **2022**, *60*, 2816–2829. [CrossRef]
26. Ghasemi, A.; Azzouz, R.; Laipple, G.; Kabak, K.E.; Heavey, C. Optimizing capacity allocation in semiconductor manufacturing photolithography area–Case study: Robert Bosch. *J. Manuf. Syst.* **2020**, *54*, 123–137. [CrossRef]
27. Uzsoy, R.; Lee, C.Y.; Martin-Vega, L.A. A review of production planning and scheduling models in the semiconductor industry part I: System characteristics, performance evaluation and production planning. *IIE Trans.* **1992**, *24*, 47–60. [CrossRef]
28. Zhang, A.; He, Y.; Han, X.; Li, Y.; Yang, X.; Zhang, Z. Modeling Product Manufacturing Reliability with Quality Variations Centered on the Multilayered Coupling Operational Characteristics of Intelligent Manufacturing Systems. *Sensors* **2020**, *20*, 5677. [CrossRef]
29. Zhou, D.; Xu, K.; Lv, Z.; Yang, J.; Li, M.; He, F.; Xu, G. Intelligent Manufacturing Technology in the Steel Industry of China: A Review. *Sensors* **2022**, *22*, 8194. [CrossRef]
30. Wolfe, H.A.; Taylor, A.; Subramanyam, R. Statistics in quality improvement: Measurement and statistical process control. *Pediatr. Anesth.* **2021**, *31*, 539–547.
31. Ryan, T.P. *Statistical Methods for Quality Improvement*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
32. Montgomery, D.C. *Introduction to Statistical Quality control*; John Wiley & Sons: Hoboken, NJ, USA, 2020.
33. Verma, A.; Dhalmahapatra, K.; Maiti, J. Forecasting occupational safety performance and mining text-based association rules for incident occurrences. *Saf. Sci.* **2023**, *159*, 106014.
34. Shafiq, D.A.; Marjani, M.; Habeeb, R.A.A.; Asirvatham, D. Student Retention using Educational Data Mining and Predictive Analytics: A Systematic Literature Review. *IEEE Access* **2022**, *10*, 72480–72503. [CrossRef]
35. Telikani, A.; Gandomi, A.H.; Shahbahrami, A. A survey of evolutionary computation for association rule mining. *Inf. Sci.* **2020**, *524*, 318–352. [CrossRef]
36. Chen, W.C.; Tseng, S.S.; Wang, C.Y. A novel manufacturing defect detection method using association rule mining techniques. *Expert Syst. Appl.* **2005**, *29*, 807–815. [CrossRef]
37. Agrawal, R.; Mannila, H.; Srikant, R.; Toivonen, H.; Verkamo, A.I. Fast discovery of association rules. *Adv. Knowl. Discov. Data Min.* **1996**, *12*, 307–328.
38. Hu, X.; Wei, Z. Internet Public Opinion Analysis Based on Apriori Association Rule Mining. In Proceedings of the 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Chongqing, China, 11–13 October 2019; pp. 1855–1858.

39. Guo, Y.; Zhou, J.; Qin, Q.; Wei, Y.; Zhang, W. An Improved Algorithm and Implementation of Data Mining for Intelligent Manufacturing Association Rules Based on Pattern Recognition. *IEEE Consum. Electron. Mag.* **2022**, *12*, 94–99. [CrossRef]
40. Al-Bana, M.R.; Farhan, M.S.; Othman, N.A. An Efficient Spark-Based Hybrid Frequent Itemset Mining Algorithm for Big Data. *Data* **2022**, *7*, 11. [CrossRef]