

# Integrated Yield-Mining Solution with Enhanced Electrical Test Data Correlation

Chih-Min Fan, Ruey-Shan Guo, Argon Chen

Graduate Institute of Industrial Engineering,  
National Taiwan University, Taiwan, R.O.C.  
rsguo@ntu.edu.tw

Amos Hon

Minefact Technologies  
Taipei, Taiwan, R.O.C.

John Wei, Mingchu King

Taiwan Semiconductor  
Manufacturing Corporation  
mcking@tsmc.com.tw

*Abstract – This paper describes the yield-mining solutions in semiconductor manufacturing from the theoretical as well as application points of view. In specific, an innovative methodology is developed to enhance the correlation between CP yield data and electrical test (ET) data so that the engineers can diagnose problems and improve yield more effectively. Also, an interactive yield-mining system is developed to embed engineers' knowledge in the data mining process. To cope with the single-tool problem and the small sample size feature in a yield ramp-up or in a high product-mix, low volume production environment, a project-based multi-product analysis methodology is developed to enhance the diagnosis capability. The proposed solutions have been fine tuned and validated in a local foundry fab.*

## INTRODUCTION

Recently "data mining" has emerged as a new tool to support the rapid yield improvement efforts and engineering data analysis. Although proven to be useful in some cases, most data mining techniques still lack engineers' full trusts due to the following issues:

- i) the underlying models and algorithms do not address the special problems only seen in semiconductor manufacturing,
- ii) most mining tools provide only problem-specific solutions instead of an "integrated" solution for yield enhancement,
- iii) engineering knowledge is usually not included in the data mining process, and
- iv) the sample size is usually too small during a yield ramp-up or in a high product-mix, low volume production environment.

Chen et al. [1] pointed out shortcomings of commonly used data mining methods for yield-loss factor identification, and proposed improved solutions to address the first issue mentioned above. In this paper, an integrated mining tool is developed to address the other three issues. In specific, an innovative methodology is developed to enhance the correlation between CP yield data and electrical test (ET) data so that the engineers can diagnose problems and improve the yield more effectively with their engineering knowledge.

To cope with the small sample size issue, a project-based multi-product analysis methodology is also developed to enhance the diagnosis capability. The goals of this paper are then to:

- i) explain the confounded correlation between yield and ET,
- ii) propose an integrated yield-mining framework with enhanced ET correlation,
- iii) design an interactive yield-mining system with engineering knowledge embedded, and
- iv) validate the proposed solutions using a multi-product analysis example.

## CONFOUNDED CORRELATION

To identify problems that cause the low yield, integration engineers usually conduct three types of correlation analysis:

- i) Correlation between yield/bin data and process tools.
- ii) Correlation between electrical test data and process tools.
- iii) Correlation between yield/bin data and electrical test data.

As pointed out in [1], the first two types of analyses have two major deficiencies in commonly used methods: first, a high false-positive rate due to one-factor-at-a-time analysis; and second, a high false-negative rate due to the lack of time segmentation [1].

As for type iii analysis, the indistinguishable correlations between yield and ET make it difficult for engineers to exercise their engineering knowledge to diagnose the yield-loss problems. The indistinguishable correlations are mainly caused by confounded factor effects. For example, wafer lots were processed by four ion implant machines. Fig. 1 shows the scatter plot of the CP yield vs.  $Rs\_N+PO1$ . The dark data points represent data collected from all lots regardless of the implant machine used. It is difficult to observe a significant correlation between the yield and  $Rs\_N+PO1$ . Based on this plot, one may quickly disregard  $Rs\_N+PO1$  as an impact ET parameter. However, if only those lots processed by a certain ion machine are examined (light data points in Fig. 1), an apparent correlation can be

established between the yield and Rs\_N+PO1. Here, the ion implant machine is the confounding factor to make the correlation indistinguishable. There are other types of factors, such as a specific time period in which the lots are processed, that will cause the similar effects. How to identify the underlying confounding factors and isolate the effects of ET parameters on the CP yield becomes critical in an integrated yield-mining tool.

#### INTEGRATED YIELD-MINING SOLUTION FRAMEWORK WITH ENHANCED ET CORRELATION

An integrated yield-mining solution is proposed and is shown in Fig. 2. As can be seen, the input data include yield/bin, electrical test, and lot history data. All the possible yield-loss factors  $\{X_i\}$  that cause the variations in yield/bin and electrical test data are identified first. Also, the effects of these factors on yield/bin and electrical test data are quantified. Data mining techniques discussed in [1] are adopted here; that is, both time-segmentation and multivariate techniques are used to enhance the factor-identification capability. Based on the identified factors and their corresponding effects, an innovative yield and electrical test cross-reference table will be generated for easy incorporation of engineering knowledge. Engineers can review the table together with their insights and knowledge of ET parameters and process factors. If a specific symptom under certain conditions (combination of factor, yield, and electrical parameter) is diagnosed to be a probable cause, engineers can further drill down to the related in-line process/tool data for validation. With this solution framework, engineers will be able to quickly and effectively conduct data analyses and fault diagnosis.

#### INTERACTIVE YIELD-MINING SYSTEM WITH ENGINEERING KNOWLEDGE EMBEDDED

Based on the data mining techniques discussed in [1] and the integrated yield-mining solution framework in Fig. 2, an interactive yield-mining system, Yield Miner, is developed. In addition to the theoretical features discussed earlier, there are four major features in Yield Miner. We'll explain the four features from the viewpoint of industry applications.

##### Project-Based Design for Multi-Product Analysis

To enhance the statistical significance of analysis results, different product data under the same technology can be grouped and normalized for analysis in yield-mining project. Also, all the analysis results and knowledge can be maintained through the project management function. The multi-product analysis capability is especially useful for a high product-mix, low volume production environment, yield ramp-up phase with limited sample size, and the single-tool problem in specific back-end process steps where a tool commonality analysis does not work because there is only one machine.

##### Object-Oriented Design for Flexible Mining Rule Extension

How to integrate the latest innovation on engineers' product, process, and data mining knowledge is one of the most challenging issues in developing a yield-mining system. We have adopted the multi-level object-oriented hierarchical design for better system flexibility and reusability. In specific, all the mining rules and algorithms are modeled and implemented as an object, and can be easily expanded or reused.

##### Memory-Based Computing for Fast Calculation

In semiconductor manufacturing, there are many attributes, such as step ID, machine tool ID, recipe, machine tool data, in-line metrology data, end-of-line ET data, CP yield data, etc., to analyze. Memory-Based computing mechanism enables fast yield-mining calculation and allows engineers to exercise their knowledge with fast what-if analysis.

##### Interactive Process for What-if Analysis

All the mining rules have their underlying assumptions and cannot fully capture the real world problem. Also, engineers usually have a lot of "dynamic" knowledge about the data, which cannot be considered during the system design phase. Our system provides an interactive what-if analysis process to solve this issue, which is shown in Fig. 3. As can be seen, there are five major steps:

- i) Input Data Screening : to filter out particular factors, such as experimental lots, metrology process steps, and/or metrology tools, which are unlikely to result in yield loss.
- ii) ET-Step Relation Constraining: to define the process line segments and their relations to ET parameters as constraints for ET-Process-Step analysis. Each process line segment consists of an ID, name, start process step, and end process step. Each ET parameter can then be defined to associate with certain defined process-line segments.
- iii) Root Cause Mining: to conduct the pre-defined queue-time, recipe, ill/golden machine, and tool re-entrant analysis on the pre-screening data
- iv) Root Cause Screening: to select and sort various factor types, to conduct corresponding Excel chart analysis for each candidate factor, to manually adjust time segmentation, and to screen out factors that are unlikely to have impact on yield, etc.
- v) Root Cause Selection: to conduct multi-factor analysis on pre-screened data and factors and view the final analysis results.

In the interactive yield-mining process, engineers can screen and change the input data as well as the system-identified factors in any one of the five steps, and force Yield Miner to recalculate the results for a fast what-if analysis. Furthermore, all the data changes and knowledge learned during the what-if analysis can be managed in the

corresponding project for knowledge management purpose and future applications.

### VALIDATION

The developed system along with the proposed solution framework have been fine-tuned and validated in a local foundry fab. Specific yield loss cases such as multi-chamber and single-tool problems were successfully identified. In this paper we describe one case study of a single-tool problem using multi-product analysis.

At the beginning of the validation case, the engineer used Yield Miner to conduct the single-product analysis and found two suspected machine tools that might cause the yield loss: EQP\_104 and EQP\_268. However, as shown in Fig. 4, both machine tools have similar shift patterns and it is difficult to judge which one is the root cause. To enhance the resolution of the shift patterns, the engineer combined three other products under the same technology into a multi-product analysis. Fig. 5 shows the results, in which important information such as the identified factor types, start-time, end-time, effect on the bin data, and statistical significance, are provided for further analyses. It can be seen that EQP\_104 has the most significant impact on the yield loss and could be the major abnormal machine. Fig. 6 shows the multi-product bin-data trend charts of EQP\_104 and EQP\_268. With the aid of time segmentation and the trend pattern resolution, it can be concluded that EQP\_104 is the major abnormal machine because EQP\_268 still has normal data points in the shifted region.

Furthermore, to identify the root cause, the engineer wants to know the impact of each identified factors on yield and ET parameters. The cross-reference table of the bin-data, ET parameters, and yield-loss factors is shown in Fig. 7 in which each cell's value represents the effect significance of each yield-loss factor on the yield or on an ET parameter. The effect significance can be sorted and classified in different colors for each yield-loss factor. For example, Factor 1 has significant impacts, in dark colors, on ET79, ET81, and ET78, etc. A process integration engineer can easily diagnose the true cause from this table based on his/her process and device knowledge. This table is more effective than the traditional empirical Key Node Table [2] by providing an up-to-date enhanced cross-reference among the CP .bin/yield, process factors and ET parameters. Moreover, the table enables engineers to drill down further to analyze the raw data with box plots, trend charts, scatter plots, etc. As shown in Fig. 8, a scatter plot of yield vs. ET\_79 under Factor 1 provides a better resolution in identifying the yield-loss root causes.

### CONCLUSIONS

The confounded correlations between yield and ET data always make it difficult for engineers to exercise their engineering knowledge to diagnose the yield-loss problems. By isolating the impact of individual yield-loss factors on the bin/yield data and on the ET parameters, the correlations between bin/yield and ET data are greatly enhanced.

To further include the engineering knowledge in the data mining process, an interactive yield-mining system is developed for a fast what-if analysis, in which a project-based multi-product analysis approach is proposed to cope with the issues of single-tool problem and the small sample-size problem during yield ramp-up or in a high product-mix environment. The solutions proposed in this paper are all aiming to integrate the engineering knowledge and data mining for rapid yield improvement.

### REFERENCES

- [1] A. Chen, R. Guo, C. Fan, A. Hon, O. Ho, L. Hu, J. Wae, and M. King, "Data Mining Techniques for Engineering Data Analysis: Issues and Solutions," ISSM 2002
- [2] R. Guo, C. Tsai, J. Lee and S. Chang, "Intelligent Process Diagnosis Based on End-of-line Electrical Test Data," IEMT 1996.

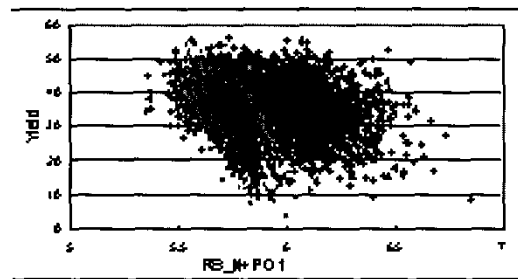


Fig.1 Yield vs. electrical test data under a specific yield-loss factor

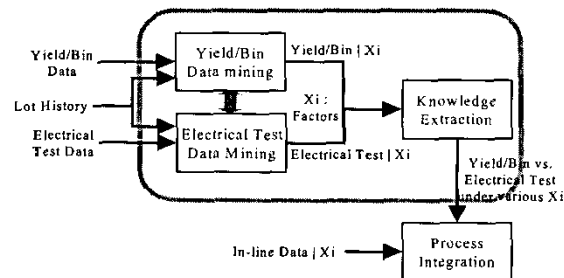


Fig. 2 Yield-mining and improvement framework

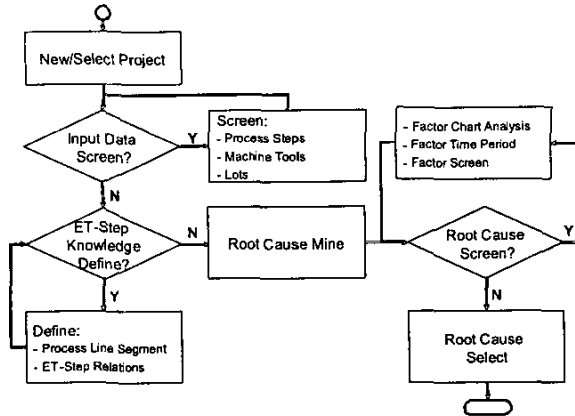


Fig. 3 Interactive yield-mining process for fast what-if analysis

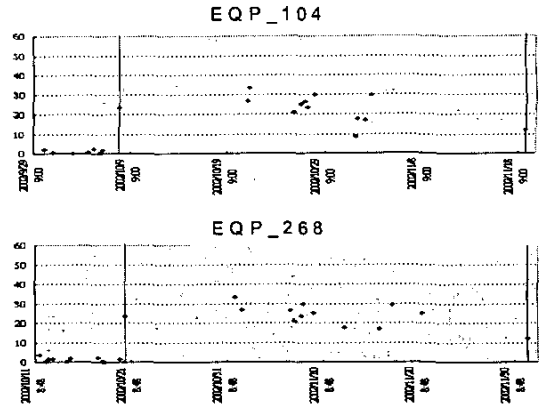


Fig. 4 Single-tool problem through single-product analysis

Variable Selection Set	Effect	Type	STEP-EQP	Period Start Time	Period End Time	Selection Significant Value
Factor 1	-7.979	Machine Time Analysis : Negative	EQP104@slot2	2001/1/1/5	2001/1/1/15	3.29E-06
Factor 2	-8.7358	Machine Time Analysis : Negative	EQP19@slot28	2001/1/1/2	2001/1/1/3	4.06E-13
Factor 3	-9.5423	Step Machine Time Analysis	HS_43-EQP127@slot4	2001/1/0/20	2001/1/1/3	0.000291242
Factor 4	-25.026	Step Machine Time Analysis	HS_132-EQP322@slot1	2001/1/0/15	2001/1/0/25	0.002313326
Factor 5	-1.5771	Machine Time Analysis : Negative	EQP256@slot1	2001/1/0/6	2001/1/1/26	6.39E-05
Factor 6	-6.8055	Step Machine Time Analysis	HS_394-EQP112@slot3	2001/1/1/16	2001/1/1/20	0.001486945
Factor 7	4.2332	Step Machine Time Analysis	HS_95-EQP261@slot2	2001/1/0/23	2001/1/1/1	2.10E-06

Fig.5 Yield-loss factor list

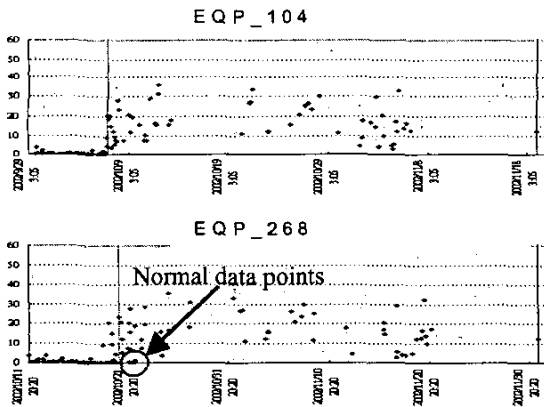


Fig. 6 Single-tool problem through multi-product analysis

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
CP YIELD					
ET_79	0.316488344	0.247075298	0.927920361	0.000000000	0.000000000
ET_91	0.348707506	0.231523219	0.890860223	0.000000000	0.000000000
ET_78	0.480	0.311	0.977093125	0.010961995	0.127
ET_86	0.363863972	0.223789319	0.834134412	0.000000000	0.137
ET_176	0.262	0.222	0.836	0.000000000	0.127
ET_169	0.279722466	0.213	0.930	0.000000000	0.137
ET_82	0.485	0.171	0.468	0.239	0.239
ET_3	0.458	0.26160217	0.60837461	0.000000000	0.239
ET_194	0.282	0.255	0.931	0.000000000	0.239
ET_84	0.301845342	0.201220124	0.51157647	0.39519202	0.239
ET_191	0.536	0.196349076	0.347913278	0.349796006	0.239
ET_45	0.324460465	0.312787089	0.752966214	0.188018448	0.239
ET_190	0.316	0.291274856	0.977182952	0.179931741	0.239
ET_187	0.371023953	0.266507143	0.649914365	0.112836728	0.239
ET_20	0.350349232	0.349129936	0.031981584	0.749019938	0.239
ET_4	0.005	0.682	0.128	0.290	0.712
ET_176	0.005862225	0.344220459	0.28158418	0.837759405	0.514657042
ET_90	0.011176865	0.710	0.112012945	0.313172529	0.835234493
ET_48	0.016	0.784880976	0.112012945	0.194270631	0.942216703

Fig.7 Cross-reference table for yield, electrical test item, and possible factors

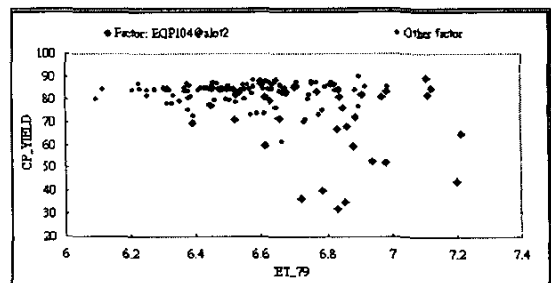


Fig.8 Yield vs. ET\_79 under Factor 1