APPLICATIONS OF DATA MINING IN INTEGRATED CIRCUITS MANUFACTURING

Sidda Reddy Kurakula[1], Lokesh Kulkarni[1], Madhu Dasari[1], Helen Armer[2]

[1]Applied Materials India (P) Ltd, Bangalore, India,
Sidda_Reddy_Kurakula@amat.com, Lokesh_Kulkarni@amat.com,
Madhu_Dasari@amat.com
[2]Applied Materials, Inc., Santa Clara, California, USA
Helen_Armer@amat.com

ABSTRACT

Integrated circuits (a.k.a chips or IC's) are some of the most complex devices manufactured. Making chips is a complex process requiring hundreds of precisely controlled steps such as film deposition, etching and patterning of various materials until the final device structure is realized. Also, each chip goes through a huge number of complicated tests and inspection steps to ensure quality. In IC manufacturing, yield is defined as the percentage of chips in a finished wafer that pass all tests and function properly. Yield improvement translates directly into increased revenues. A humongous amount of data (Terabytes per day) is logged from the equipment in the fab. This paper describes some applications of advanced data mining techniques used by chip makers and equipment suppliers in order to improve yield, match equipment, increase equipment output and also to predict the change in equipment performance before and after maintenance activities.

KEYWORDS

Integrated Circuit (IC), Yield, Equipment, Wafer, Models

1. INTRODUCTION

The process of creating integrated circuits (IC's) is called wafer fabrication. It is a sequence of chemical and photographic steps (like lithography, etching, deposition, oxidation and diffusion) in which the circuits are constructed on a semiconductor material typically called a wafer. In order to perform Automatic Process Control and offline data mining, a large amount of data is collected, stored and retrieved from the equipment in which the said processes are being carried out. In IC manufacturing, yield is defined as the percentage of chips in a finished wafer that pass all tests and function properly. Due to the large number of processing steps and the complex interactions between steps, yield is a complex function to analyse. Data mining methods for yield analysis are now starting to be developed and deployed. Table 1 depicts the yield values for 2 different process technologies with a total number of steps (N) equal to 200 and 450. As can be seen, total perfection at each process step is absolutely necessary for achieving higher yields.

Table 1. Yield for two different process technologies.

x	Y for N = 450	Y for N = 200
95%	9.45×10^{-9} %	3.51×10^{-3} %
99%	1.09 %	13.4 %
99.9%	63.7 %	81.9 %

Where

x = Success rate of each process step and
Y= Yield of working devices (i.e. $(x/100)^N$ %)

2. DATA MINING METHODOLOGY

In the proposed data mining methodology a holistic approach is being followed by the FabVantage™ group at Applied Materials. This methodology goes beyond traditional statistical process control methods which primarily emphasize process monitoring for change-point detection; instead it focuses more on yield driven control limits and strategies.

2.1. Data sources and typical volumes

Different types of data that are generated and used for various purposes in an IC fabrication unit are event logs data, unit processes data, integration data, inspection & review data, metrology data and parametric and final yield data. The size of each type of data varies from a few gigabytes/day to a few terabytes/day depending on production capacities. "Unit processes" constitute 30 to 40% of total process steps involved in making an IC. All the data mining techniques described in later sections were mostly used to analyse this "Unit Processes" data. The Equipment Data Acquisition is typically performed by one or more factory data gathering or analysis software applications (clients) using different standards like SECS (Semiconductor Equipment Communications Standard), GEM (Generic Equipment Model) or Interface A. One example of this type of client is the Applied Materials E3™. E3 is the only equipment engineering system solution that combines statistical process control (SPC), fault detection and classification (FDC), equipment performance tracking (EPT), advanced data mining (ADM), run-to-run control (R2R) and tool automation on a unified platform. In addition to this most of the Equipment Controller Software provides the option to export/store the data in various file formats. The generic input and output model in data mining for IC fabs is shown in Figure 1.
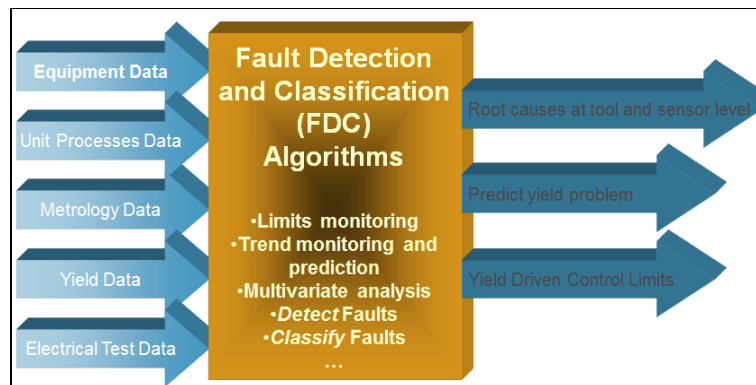

Figure 1. Inputs and outputs of FDC algorithms in IC fabs.

2.2. Data quality check

The unit process data (mentioned above) typically called equipment (tool) sensor data consists of all the physical variables like temperature, flow, pressure, acceleration, torque, angle and the like. This data is recorded during each wafer run. A data quality check algorithm verifies this data for any missing sensors from the defined data collection plan, missing data or stagnant data and also detects any non-optimal sampling rate in any sensor.

2.3. Knowledge Base and sensor priority

Knowledge base is the repository of methodologies, tool documentation, Best Known Methods (BKMs), lessons learned, sensor data collection plans, previously used models and the like. A key part of the knowledge base is sensor prioritization. Each sensor collected from a tool is assigned a priority of P1 through P4 to denote its impact on yield. P1 sensors are known to impact yield if they go out of range, while P4 sensors are known to have no yield impact if they go out of range. P2 and P3 sensors are suspected to impact or not impact yield, respectively, if they go out of range. Sensor priorities are used to reduce the burden of analysing 1000's of sensors in the initial stages of data mining. Table 2 shows an example of sensor priorities.

Table 2. Example Priority Sensor list for one specific tool/process type.

Sensor Name	Sensor Units	Sensor Priority	Sampling Frequency(Hz)
Temperature	degC	P1	1
H_2 Gas Flow	sccm	P1	1
RF Forward Power	Watts	P1	5
Chamber Process Pressure	mtorr	P1	1
Foreline Pressure	mtorr	P2	1
E-Chuck Voltage	V	P2	1
Gas line pressure	psi	P3	0.5
Wafer Counter	None	P4	1 sample per wafer run

Where

P1: Confirmed/known to have caused a yield problem,
P2: Science suggests there will be a problem but no experience from data,
P3: No knowledge whether or not there will be a problem,
P4: Known to be a non-issue.

2.4. Tools and modelling methods

A combination of data mining software including Applied Materials E3, R [1], JMP [2], and UNIX scripting were used to perform the statistical modelling, associated data preparation and reporting. A variety of modelling techniques are used. These include Rules Ensemble [3], Random Forest [4], Support Vector Machines (SVM), Partial Least Squares (PLS) Analysis and Discriminant Analysis in supervised learning and clustering analysis or Principal Component Analysis in unsupervised modelling.

3. ADVANCED DATA MINING APPROACH

3.1. Data visualization

Data visualization is the first and foremost analysis task that helps to identify obvious abnormalities while performing wafer to wafer comparison, lot to lot comparison or recipe to recipe comparison in the tool sensors data. It also helps to rebuild the recipe to quickly compare the recipe under investigation with the BKM recipe. The differences can be summarized and can be used at a later stage while running any classification/regression models as sometimes the differences are deliberately set. Table 3 shows one example recipe wherein the differences from BKM recipe are highlighted in *italics*.

Table 3. Comparison of a recipe with BKM.

Attribute/Step Number	1	2	4	5
Step Name	Stabilization	Deposition	Purge	Pump
Step Time (Sec)	2	30	5	10
Temperature (deg C)	200	*250* *(BKM is 270)*	190	180
Pressure (Torr)	5	5	2	1
Flow (sccm)	200	*200* *(BKM is 250)*	50	10

Customized plotting like shown in Figure 2 was performed in R in order to define the Univariate Analysis (UVA) models using appropriate statistics in different process regions like maximum/slope in transient regions and similarly mean and standard deviation in stable regions.
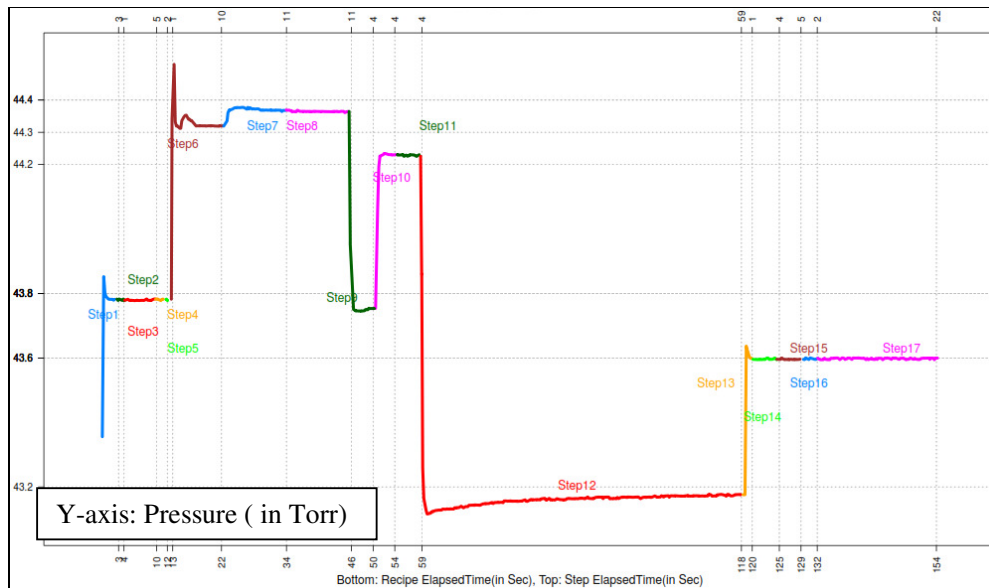

Figure 2. Trace plot for pressure sensor from a single wafer run.

3.2. Modelling

Once appropriate UVA models are identified and summary statistics calculated, the first pass of modelling results are studied. This typically involves filtering out redundant variables to reduce the dimensions to a more manageable subset. Depending on the problem being analysed, a variety of supervised and unsupervised learning algorithms are available for use. Supervised algorithms, like some regression or classification techniques, help establish relationships between a dependent variable (e.g. metal film thickness) and a set of independent variables (e.g. gas flows); unsupervised methods like Principle Components Analysis (PCA) or clustering help highlight interaction between different variables (sensors).

The first pass modelling results are followed by a number of iterations to successively and systematically remove noise and unwanted variables to improve model quality. Next, a model quality report is generated with the top ranked variables and their respective contributions (Figure 3). A glance at the plot of predicted values vs. actual values (Figure 4) provides a good measure of the model quality.

More often, a combination of two algorithms is significantly more effective than using any one algorithm. For example, consider fitting a prediction model for transistor current (Idrive) after an etching process. It was found that Rules Ensemble alone performed very poorly in terms of the predictive ability of the model. Likewise, an algorithm like Random Forest, which builds decision trees based on splitting sensor values, is prone to overfitting. However, a combination of the two methods proved to be significantly more powerful – a Rules Ensemble algorithm was used to reduce the number of variables (sensors), while Random Forest was then used to build a model with high predictive power and good generalizability.
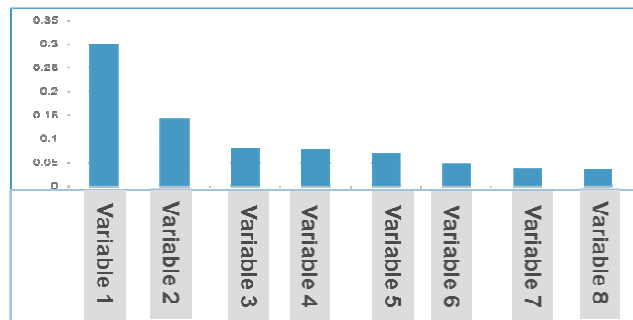

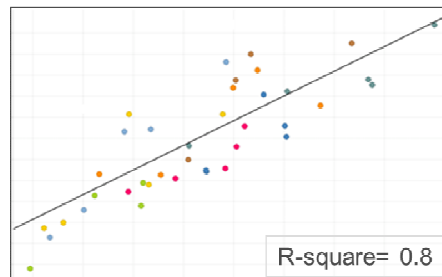Figure 3: Contribution Pareto of top-ranked variables.

Figure 4: Plot of predicted value of dependent variable (x) vs. actual value of dependent variable (y).

Yield driven control limits were set at the end of the data mining phase wherein the independent variable is allowed to vary in a specific range based on its correlation with the yield numbers as

depicted in Figure 5. Idrive is the dependent variable and the FDC (Fault Detection and Classification) sensor is the independent variable like gas flow that was found to be highly correlated with the dependent variable. The vertical lines on the right side of the chart indicate the independent variable control limits needed to meet the dependent variable specification limits as shown in the left side of the chart.
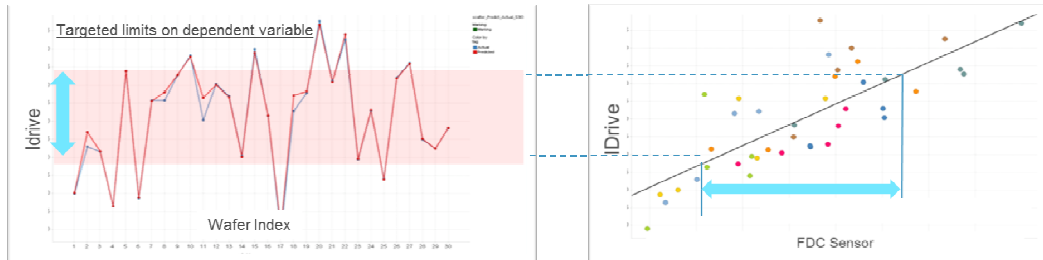

Figure 5. Left: Actual (blue) vs. predicted (red) dependent variable and Right: Scatter plot of independent variable vs. dependent variable.

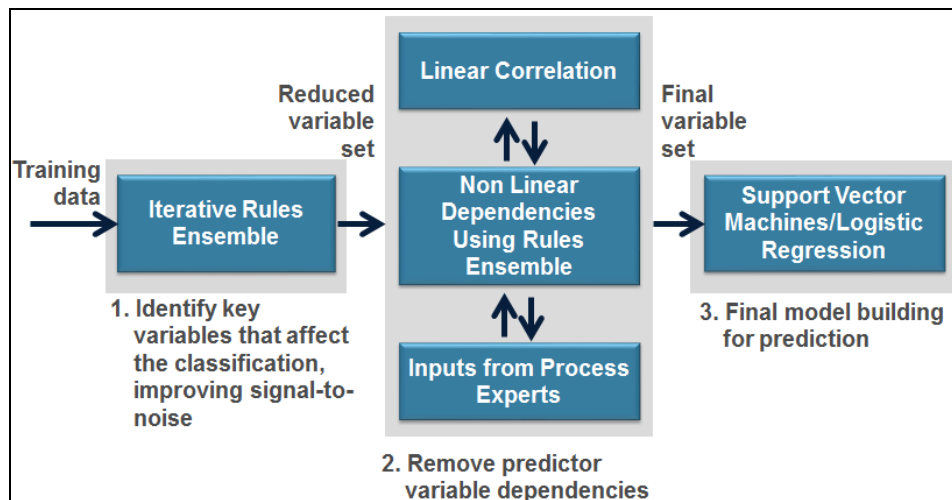Figure 6 is the visual representation of the said data mining methodology.

Fig 6. Advanced data mining approach.

4. RESULTS AND DISCUSSION

Some of the high level data mining problem statements that we address are listed below:

- Transistor drive current not matched post-etch process.
- Thickness uniformity of hard mask layers are not in specification limits of <1.5%.
- Improving the temperature matching on Epi chambers from ±10ºC to ≤5ºC.
- Identifying key sensors controlling transistor drive current from Epi process.
- Determining the root cause of arc and deep scratches at Copper Chemical and Mechanical Planarization process and reducing them.
- Reducing particle count on various key device layers.

Multivariate regression was performed using Random Forest and Rules Ensemble models separately in R platform in order to find the key sensors controlling the transistor drive current variation across multiple chambers from an etch process. The model prediction power (based on R-square value) was 0.95. From the drill down and physical verification of the model the root cause for transistor drive current variation was identified. Subsequent recipe optimization resulted in reducing the standard deviation of the process by 30% as depicted in Figure 7.
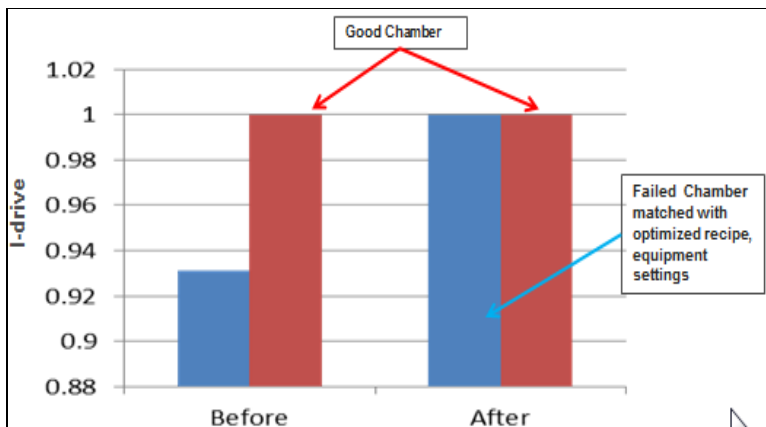

Figure 7. Chamber matching performance matching after changes implemented.

In another case multivariate regression analysis found the root cause sensors affecting the uniformity mismatch of dielectric layers wherein the slope of liquid gas flow sensor across 6 chambers were found to be the variable of highest correlation to the dependent variable (i.e., uniformity). Figure 8 shows the improvement in non-uniformity values after the changes were implemented based on the data mining results on data from 300 wafers.
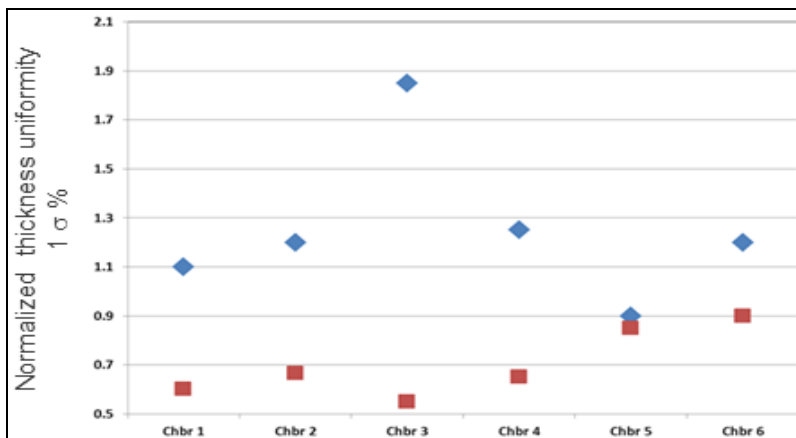

Figure 8. Dielectric film non-uniformity improvement in 6 chambers.

The chart below (Figure 9) shows the pre- and post-implementation results obtained after implementing changes based on the FabVantage data mining analysis. Analysis of gate critical dimension (CD) data over a span of six months revealed that a bad RF generator was causing a bias impedance mismatch (resistance and reactance), which was in turn driving the variation in gate CD.
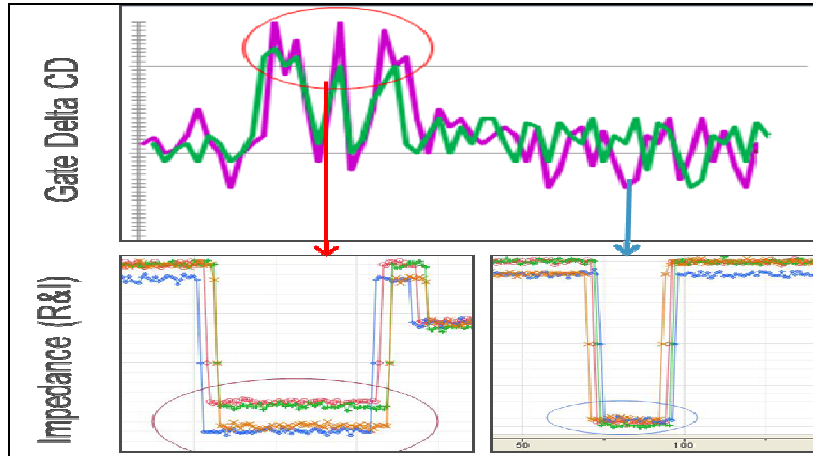
Figure 9. Identification of root cause driving variation in gate CD across 2 etch chambers.

5. CONCLUSIONS

Advanced data mining techniques are deployed in order to achieve higher yields in IC manufacturing units. Different statistical modelling techniques were used to study the impact of independent physical variables on the chamber matching, fault detection and to set the new control limits in order to maintain the yield at desired levels. Similar methodology can be used in any semiconductor, electronic, and photovoltaic manufacturing.

ACKNOWLEDGEMENTS

REFERENCES

[1]   R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org/.

[2]   JMP®, Version 11. SAS Institute Inc., Cary, NC, 1989-2007.

[3]   J.H. Friedman, and B. E. Popescu, "Predictive Learning via Rules Ensemble," Stanford University, Department of Statistics, Technical Report, 2005.

[4]   A. Liaw and M. Wiener (2002). Classification and Regression by randomForest. R News 2(3), 18--22.

AUTHORS

Sidda Reddy Kurakula obtained a Master's Degree in Semiconductor Physics in 2007 from Indian Institute of Science, India. He joined Applied Materials in 2008, where he has worked as a process engineer and data analyst. Mr. Kurakula specializes in data mining and analysis of sensor and event data from process equipment. He has performed data analysis for nearly all of Applied Materials' semiconductor and solar processing equipment. His work has helped to identify root causes of tool excursions and mis-processing on many types of equipment, including CVD, Etch, CMP, and implant. Mr. Kurakula has co-authored 2 papers and has contributed to many internal white papers and project reports.

Lokesh Kulkarni received both his Bachelor's and Master's degrees in industrial engineering from Texas A&M University at College Station, Texas (USA). In 2013, Lokesh joined Applied Materials in the capacity of an Engineer – Technology, with much of his work focused on leveraging data mining techniques in the context of semiconductor device manufacturing.

Madhu Babu Dasari obtained a Bachelor's Degree in Computer science from Jawaharlal Nehru University, India. He joined Applied Materials in 2005, where he started working as a software engineer – developer on various modules related to Applied Materials E3 software.

Helen Armer received a Bachelor's Degree in Chemical Engineering from Tulane University in Louisiana (USA), and a Ph.D. in Chemical Engineering from The University of Houston in Texas (USA). She joined Applied Materials in 1999. She has held various roles, including engineering, product management, and knowledge base management. She is currently responsible for the knowledge base for Applied's consulting and advanced services businesses. She holds 26 U.S. patents and has published >25 papers.