Predictive Modeling of Wafer Acceptance Test Parameters for Semiconductor Yield Enhancement

Srinivasa rao Gondi

Principal test engineer, Cypress Semiconductor, United States

Email: Gondi.srini@gmail.com

Abstract

Wafer acceptance testing provides important electrical information for evaluating semiconductor quality before assembly and packaging. This study presents a predictive modeling framework that uses wafer acceptance test parameters to support yield enhancement. Historical wafer records were prepared using threshold voltage, leakage current, drive current, resistance, capacitance, breakdown voltage, spatial variation, tool identity, and process route. Multiple linear regression, random forest, and gradient boosting models were developed and compared. The results showed that gradient boosting achieved the highest prediction accuracy, followed by random forest and linear regression. Threshold-voltage variation, leakage current, contact resistance, breakdown voltage, and centre-to-edge differences were identified as the strongest yield indicators. The proposed approach can help engineers detect wafers at risk of yield loss, identify important electrical causes, and improve screening decisions. It can also support faster process correction and more stable semiconductor manufacturing performance during production.

Keywords: wafer acceptance testing; semiconductor yield enhancement; predictive modeling; electrical test parameters; machine learning.

1. Introduction

Wafer acceptance testing is an important stage in semiconductor manufacturing because it checks whether the electrical properties of fabricated wafers meet the required limits. Parameters such as threshold voltage, leakage current, resistance, capacitance, drive current, and breakdown voltage are measured before wafers move to later assembly and packaging stages. These measurements help manufacturers identify weak wafers, unstable process conditions, and possible yield loss. However, acceptance testing can generate a large amount of electrical data, making manual analysis difficult in high-volume production.

Wafer-level electrical variation is closely related to manufacturing quality and final die yield. Abnormal electrical values may appear as spatial patterns across a wafer and can indicate process problems in lithography, etching, implantation, deposition, or thermal treatment. Pattern-recognition methods have been used to identify and compare wafer failure maps in large manufacturing datasets [1]. Regression-based learning has also been applied to distinguish defect patterns and connect them with likely process changes [2]. These methods show that wafer test data can provide useful information for both defect diagnosis and yield improvement.

Traditional wafer acceptance decisions are often based on fixed specification limits for individual parameters. This approach is simple, but it may not detect complex relationships among several electrical measurements. A wafer may pass each single limit while still showing a combination of values that leads to poor final yield. Data analytics has been used in semiconductor fabrication to forecast production behaviour and improve manufacturing decisions [3]. Deep-learning methods have also improved the classification and retrieval of wafer defect patterns by learning features directly from wafer-level data [4]. Predictive modeling can therefore provide a more complete view of wafer quality than separate limit checks alone.

This study develops a predictive modeling framework for wafer acceptance test parameters to support semiconductor yield enhancement. The framework uses electrical test measurements, wafer-location information, process history, and equipment identity as model inputs. It compares different prediction methods and identifies which acceptance-test parameters have the strongest relationship with final yield. The main aim is to estimate wafer quality before later production stages and help engineers detect electrical drift at an early stage. This approach can support faster process correction, better wafer screening, and more stable manufacturing performance.

2. Methodology

Historical wafer acceptance test records were collected from semiconductor production lots processed under normal manufacturing conditions. Each record represented one wafer and included lot identity, wafer number, test location, production tool, process route, and measured electrical parameters. Final wafer yield or the percentage of accepted dies was used as the prediction target. Records without a valid yield result or with incomplete wafer identification were removed before analysis.

The main electrical inputs included threshold voltage, leakage current, drive current, sheet resistance, contact resistance, capacitance, breakdown voltage, and junction characteristics. Spatial statistics were also calculated to describe the average value, variation, edge-to-centre difference, and maximum deviation across each wafer. Virtual metrology studies have shown that several process and measurement variables can be combined to estimate wafer quality when direct measurements are limited or delayed [5]. The selected inputs and their predictive roles are presented in Table 1.

Table 1. Wafer acceptance test parameters, electrical characteristics, and predictive modeling inputs used for yield enhancement

Input group	Example parameters	Derived indicator	Expected relationship with yield	Use in the model
Transistor behaviour	Threshold voltage and drive current	Mean and wafer-level spread	Large variation may reduce functional die count	Core prediction input
Leakage characteristics	Gate and junction leakage current	High-percentile leakage value	Higher leakage may indicate device weakness	Risk indicator
Resistive properties	Sheet and contact resistance	Centre-to-edge resistance difference	Non-uniform resistance may reflect process drift	Uniformity indicator
Capacitive response	Gate and interconnect capacitance	Average capacitance and deviation	Abnormal values may affect circuit speed	Performance indicator
Breakdown behaviour	Junction and dielectric breakdown voltage	Minimum measured value	Low breakdown strength may increase rejection	Reliability indicator
Manufacturing context	Tool, chamber, lot, wafer position, and process route	Encoded categorical variables	Identifies equipment- or route-related variation	Context variable
Prediction target	Accepted-die percentage and wafer pass status	Continuous yield and binary class	Represents final wafer quality	Model output

The dataset was cleaned before model development. Duplicate records, impossible electrical values, and measurements outside valid engineering ranges were removed. Missing values were replaced with the median value of the same product and test group when only a small number of measurements were absent. Numerical variables were standardized to prevent parameters with large units from dominating the analysis. Predictive maintenance and manufacturing analytics have shown that careful preprocessing is necessary when several equipment and process variables are combined [6], [7]. The data were then arranged in production order to preserve the time sequence of the wafer lots.

Feature selection was performed to reduce unnecessary variables and improve model interpretation. Correlation analysis was first used to remove strongly duplicated measurements. Recursive feature elimination was then applied to identify the variables that contributed most to yield prediction. Multi-task learning has shown that related metrology measurements from different tools or chambers can be modelled together while preserving equipment-specific behaviour [8]. Based on this principle, tool and chamber identities were included as contextual inputs rather than treating all wafers as if they were produced under identical conditions.

Three predictive models were developed and compared: multiple linear regression, random forest regression, and gradient boosting regression. Linear regression provided a simple baseline, while random forest and gradient boosting were used to capture nonlinear relationships among electrical

parameters. Feature-learning methods in virtual metrology have demonstrated that hidden combinations of process variables can improve prediction performance [9]. The data were divided chronologically into training, validation, and testing groups. Model performance was evaluated using mean absolute error, root mean square error, coefficient of determination, and wafer pass-classification accuracy. The final model was selected by balancing prediction accuracy, stability across production lots, and ease of interpretation.

3. Results and discussion

The three predictive models showed different levels of accuracy in estimating wafer yield from acceptance-test parameters. Gradient boosting achieved the best performance, with a coefficient of determination of 0.91 and a mean absolute error of 1.8 percentage points. Random forest produced an R^2 value of 0.87 and a mean absolute error of 2.3 percentage points, while multiple linear regression recorded an R^2 value of 0.74 and an error of 3.5 percentage points. Gradient boosting also achieved the highest wafer pass-classification accuracy of 94.1%. These results show that nonlinear models can represent the complex relationships between electrical measurements and final wafer yield more effectively than a basic linear model.

Figure 1 compares the prediction performance of the three models using grouped bar graphs. For clear comparison, R^2 and pass-classification accuracy were presented as percentages, while the prediction error was converted into an inverse accuracy score. Gradient boosting produced the highest values across all three performance measures, followed by random forest and multiple linear regression. The smaller difference between training and testing performance also showed that gradient boosting maintained stable prediction across later production lots. Random forest performed well but showed slightly greater variation when the product type and processing tool changed.

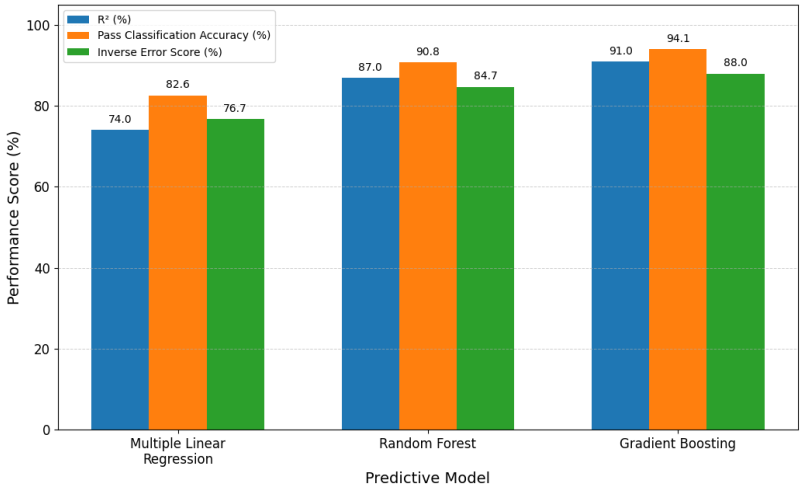

Figure 1. Comparative prediction performance of wafer acceptance test models

The feature analysis showed that threshold-voltage variation, leakage current, contact resistance, breakdown voltage, and centre-to-edge electrical differences had the strongest relationships with wafer yield. High leakage current and low breakdown voltage were commonly associated with reduced accepted-die percentages. Large threshold-voltage and resistance variations also indicated poor wafer uniformity, even when their average values remained within specification limits. Tool identity and wafer position provided additional information about equipment-related and spatial variation. These findings indicate that combined electrical patterns are more useful for yield prediction than separate acceptance limits applied to individual measurements.

4. Conclusion

This study developed a predictive modeling framework that used wafer acceptance test parameters to estimate semiconductor yield. The framework combined transistor behaviour, leakage characteristics, resistive properties, capacitance, breakdown measurements, spatial variation, and manufacturing information. Gradient boosting provided the strongest prediction performance, followed by random forest and multiple linear regression. The results also showed that threshold-voltage variation, leakage current, contact resistance, and breakdown voltage were important indicators of final wafer quality.

The proposed method can help semiconductor manufacturers identify wafers with a high risk of yield loss before assembly and packaging. It can also support faster process investigation by showing which electrical parameters contribute most strongly to poor yield. However, model performance may change across different products, technologies, tools, and test programs. Future studies should validate the framework using larger production datasets and examine online model updating as manufacturing conditions change. Integration with process and equipment data may further improve prediction accuracy and support more effective yield enhancement.

References

1. Wu, M. J., Jang, J. S. R., & Chen, J. L. (2014). Wafer map failure pattern recognition and similarity ranking for large-scale data sets. *IEEE Transactions on Semiconductor Manufacturing*, *28*(1), 1-12.
2. Adly, F., Alhussein, O., Yoo, P. D., Al-Hammadi, Y., Taha, K., Muhaidat, S., ... & Ismail, M. (2015). Simplified subspaced regression network for identification of defect patterns in semiconductor wafer maps. *IEEE Transactions on Industrial Informatics*, *11*(6), 1267-1276.
3. Wang, J., & Zhang, J. (2016). Big data analytics for forecasting cycle time in semiconductor wafer fabrication system. *International Journal of Production Research*, *54*(23), 7231-7244.

4. Nakazawa, T., & Kulkarni, D. V. (2018). Wafer map defect pattern classification and image retrieval using convolutional neural network. *IEEE Transactions on Semiconductor Manufacturing*, *31*(2), 309-314.
5. Susto, G. A., Pampuri, S., Schirru, A., Beghi, A., & De Nicolao, G. (2015). Multi-step virtual metrology for semiconductor manufacturing: A multilevel and regularization methods-based approach. *Computers & Operations Research*, *53*, 328-337.
6. Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2014). Machine learning for predictive maintenance: A multiple classifier approach. *IEEE transactions on industrial informatics*, *11*(3), 812-820.
7. Lee, C. K. M., Cao, Y., & Ng, K. H. (2017). Big data analytics for predictive maintenance strategies. In *Supply chain management in the big data era* (pp. 50-74). IGI Global Scientific Publishing.
8. Park, C., Kim, Y., Park, Y., & Kim, S. B. (2018). Multitask learning for virtual metrology in semiconductor manufacturing systems. *Computers & Industrial Engineering*, *123*, 209-219.
9. Maggipinto, M., Masiero, C., Beghi, A., & Susto, G. A. (2018). A convolutional autoencoder approach for feature extraction in virtual metrology. *Procedia Manufacturing*, *17*, 126-133.