



A SLvT Adaptive Test Method for Integrated Circuit Test Parameter Sets without Yield Loss

Qiong Wu¹ · kaiming Hao¹ · Wenfa Zhan²

Received: 9 June 2024 / Accepted: 31 October 2024 / Published online: 9 December 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

With the development of semiconductor technology, the fabrication process of integrated circuits is complicated and expensive, and the testing of integrated circuits has become increasingly difficult. The reduction of testing costs has become a very important issue, and how to reduce the yield loss caused by testing has become increasingly important. Therefore, an adapted SLvT (Simplified Loss recovery Test) methodology is proposed, leveraging historical wafer data. Initially, the approach employs RFECV (Recursive Feature Elimination with Cross-Validation) and Pearson correlation analysis to sequentially select pivotal parameters., and XGBoost is used to achieve accurate prediction. For chips predicted as faulty, additional parameter items are added using the MI (Mutual Information) method improved by normal distribution. Prediction is then achieved through the retrained XGBoost., which greatly reduces the test yield loss. Finally, the chips that are predicted to be bad are fully tested to achieve zero yield loss. In addition, experimental validation underscores the efficacy of the method, demonstrating a substantial decrease in test resource occupancy, only 32.5%, at the expense of only 0.09% of test escape. At the same time, compared with other adaptive methods, the performance of zero yield loss is better than 83.5%

Keywords XGBoost · SPK-SMOTE · Parameter data · Segmented testing

1 Introduction

Scientific and technological advancements have significantly contributed to the prosperity of the semiconductor industry [1]. Quality control plays a pivotal role in semiconductor manufacturing, facilitating cost savings and ensuring timely delivery [2]. As integrated circuit sizes decrease and processes become more complex, the CP (Circuit Probing) process has emerged as pivotal in determining wafer yield, albeit requiring substantial time and specialized equipment [3, 4]. Consequently, engineers leverage WAT (Wafer Acceptance Test)-based forecasting to adjust CP processes, thereby reducing manufacturing duration and costs. In the

competitive semiconductor landscape, cost management has become paramount for corporate profitability. The term "wafer acceptance test" is normally used to refer to limited probe during and just at the end of wafer fabrication where only sample sites are tested, typically 9 or 20.

At the onset of this year, an earthquake in Ishikawa Prefecture, Japan, necessitated shutdown inspections for semiconductor silicon wafer manufacturers such as Shin-Etsu and Toshiba, significantly disrupting their production activities. Similarly, an earthquake in Yilan County, Taiwan, in August impacted Taiwan Semiconductor Manufacturing, Liandian, and other wafer foundry enterprises, adversely affecting their production equipment. These unforeseen events pose risks of damaging chip manufacturing equipment, halting production lines, and resulting in production cuts, potentially disrupting global supply chains.

Furthermore, technological advancements serve as the pivotal force propelling incessant product innovation within the chip industry. Consequently, a myriad of chip manufacturers have intensified their focus on technical investments. This heightened emphasis underscores the paramount significance of swift product introductions for cost recuperation and profit optimization. In light of these dynamics,

Responsible Editor: R. A. Parekhji.

✉ Wenfa Zhan
aqsfyjanjiusheng@163.com

¹ School of Mathematics and Phycsis, Anqing Normal University, Anqing 246133, China

² School of Electronic Engineering and Intelligent Manufacturing, Anqing Normal University, Anqing 246133, China

harnessing historical wafer data to devise sophisticated adaptive testing methodologies that meticulously minimize test items is poised to revolutionize the sector. This innovative approach not only conserves testing resources but also upholds stringent quality control standards, thereby streamlining production processes and augmenting overall efficiency amidst the industry's challenges. These methodologies are designed to enhance product profitability by maximizing the availability of chips within the constraints of limited production capabilities.

Traditional testing entails subjecting all chips to the same test set, where test limits, content, and flow remain fixed. This approach utilizes the full test set for both defective products and chips with a high yield, leading to significant wastage of test resources.

Adaptive testing methods refer to the use of test data for analysis on the basis of standard tests to make test selections and judgments.

In the realm of adaptive testing methods, both academia and industry have undertaken extensive explorations. Within the Wafer Acceptance Test (WAT) process, the number of parameters requiring testing progressively increases, leading to a simultaneous rise in time consumption and testing costs [5]. Given the substantial quantity of WAT parameters, the complexity of relationships between parameters becomes evident, acquiring critical parameters proves challenging [6]. Researchers at Arizona State University, as documented in literature [7], have employed test items to assess the efficiency of fault-chip testing and the correlation among test items. They utilized kernel density estimation to calculate the joint probability density function between test items, aiding in the sorting and reduction of test items. Chien et al. [8], through expert experience, selected 12 highly correlated WAT factors and subsequently developed an improved analysis method based on Modified Partial Least Squares (mPLS). Ultimately, they used these highly correlated factors as input parameters for modeling, achieving superior predictive performance compared to the traditional Partial Least Squares (PLS) method. Literature [9] proposes a method for parameter selection in the presence of missing and unlabeled data. This method involves using TSR (Total Sum of Squares Regression) interpolation for missing value data and selecting important parameters through specific predictive models, such as PLS and LASSO (Least Absolute Shrinkage and Selection Operator) regression. Ahmadi et al. [10] utilized Principal Component Analysis (PCA) to reduce the dimensionality and features of high-dimensional data in testing parameter datasets, transforming parameter set reduction into Integer Linear Programming (ILP). Their proposed testing selection method requires minimal testing infrastructure support, yet achieves a remarkably low test escape rate. Additionally, Chen et al. [11] employed clustering techniques to cluster high-quality process quality

parameters, followed by the Kruskal–Wallis test to examine significant differences between data sets. They utilized decision trees to infer parameter clustering results, thereby constructing an analytical framework for key parameters in wafer manufacturing yield analysis. Wang and colleagues [12] devised a method for selecting critical parameters based on information entropy, aiming to identify those parameters that impact production cycle fluctuations. Reference [13], leveraging offline statistical information obtained from circuits and real-time statistical information from each tested chip, selects testing parameters, effectively reducing testing time while maintaining low levels of test escape rates. Reference [14] introduces a machine learning-based framework encompassing various techniques for feature selection and handling imbalanced classes. Utilizing the chosen parameters, machine learning models are employed to predict failures in given units. In summary, methods for identifying key wafer parameters include traditional statistical analysis methods, filtering methods, heuristic wrapper methods [15–17], and machine learning methods. However, as the scale of wafer parameters expands and constraint factors increase, the performance of these methods becomes limited, making it challenging to automatically and efficiently identify critical WAT parameters for wafers.

In summary, conducting tests based on identifying critical parameters from the test set can enhance testing efficiency and reduce testing costs. However, these methods exhibit certain drawbacks. Firstly, some methods merely utilize the impact of individual parameter terms on testing without considering the nonlinear relationships among testing items. Consequently, they are not very efficient in reducing parameter terms, and their advantage in lowering testing costs is not very evident. Secondly, when facing test escapes, these methods do not achieve a very low level. Lastly, although these methods can reduce certain testing costs, they result in significant waste in yield, thus not effectively reducing the overall testing costs of chips.

To address the issues of high test escapes, high yield wastage, and long testing times in the aforementioned methods, a SLvT testing approach is proposed. This method involves analyzing data imbalances to achieve the goal of low test escapes and reduces the utilization of testing resources through a two-stage parameter selection process. Finally, conducting full tests on products predicted as faulty achieves zero yield loss. As the availability of chips increases, the chip shortage issue can be partially alleviated. The contributions of this study are outlined as follows:

1. Use different parameter selection strategies to achieve a low level of test resource occupancy.
2. The current self-adaption method lacks a reduction in the loss of test yield, and this method achieves zero yield loss.

3. SPK-SMOTE is introduced as a method for synthesizing data that is tailored specifically for chip testing and exhibits remarkable performance. The generated data is optimized to meet the unique requirements of chip testing, ensuring its suitability and efficacy in this application domain.

4. This method is completely based on software, does not occupy hardware resources, and has good application conditions in the actual test process.

The remainder of this paper is structured as follows: Section II presents the research foundation, Section III elaborates on the methodology, Section IV presents the experimental findings, and Section V concludes the study.

2 Research Basis

2.1 Parameter Selection

Integrated circuit (IC) testing represents a complex and meticulous process aimed at verifying whether a chip adheres to established product specifications and determining its quality grade. Throughout the testing procedure, the selection of various parameter items plays a pivotal role in ensuring product quality while also impacting testing efficiency and cost. The optimization of these test parameters, namely the reduction of unnecessary test items while maintaining test result accuracy, stands as a crucial factor in enhancing the efficiency of integrated circuit testing.

Traditional methods of parameter selection, such as mathematical and statistical approaches, though effective in certain scenarios, may not fully exploit the potential connections between parameter items due to the complex nonlinear relationships among test parameters. Consequently, there is a need for more advanced algorithms capable of identifying redundant test parameters and discerning essential ones.

The proposed SLvT algorithm represents further innovation built upon existing technology. This algorithm not only focuses on singular relationships between parameters but also comprehensively considers a variety of factors, enabling a deeper analysis of parameter interactions. Through this approach, the algorithm effectively identifies parameters with minimal impact on test results, thereby optimizing the testing process. This optimization not only reduces testing time, but also lowers testing costs without compromising product quality.

2.2 Quality Prediction Model

Quality prediction based on machine learning is an important idea in adaptability testing, and some existing schemes have confirmed that this idea can effectively reduce testing

costs [18–20]. First, some chips are sampled on the wafer to be tested for standard testing, and the prediction model is trained using the test results of these chips. The chips that have not been sampled are used as the chips to be tested, and the test results are given by the quality prediction model. Next, as shown in Fig. 1, when establishing the quality prediction model, the characteristic test items need to be screened first. The screened test items will be used as features, and the test results (Result) will be used as labels. The test items screened for the first time are the features trained by Model A, and the parameter items of Model B are a collection of two filters. Secondly, the data of the screened test items should be balanced, and finally, the prediction result of the chip is given to the model training through the test item data. For the chip to be tested, only the characteristic test items need to be tested when predicting through the quality model, and finally the trained model A makes predictions about the quality of the chip (good / bad). The chip predicted to be bad needs to test the parameters of the second screening, and the chip is predicted by the quality model B.

3 SLvT Adaptation Method Without Yield Loss

This section will describe in detail the proposed low-escape adaptability test method for SLvT without yield loss. As shown in Fig. 2, the method is divided into two stages: (1) parameter item selection; (2) quality model training.

In Part (1), the parameter set is meticulously screened based on the historical data of wafers, with the aim of selecting an efficient parameter set that preserves comprehensive information concerning both parameter items themselves and their relationships with test results. This selection procedure is conducted in two distinct phases. Notably, the parameter set chosen in the second phase encompasses the testing content from the previous phase. Subsequently, based on the outcomes of Part (1), the XGBoost algorithm is utilized for model training in both phases, enabling it to thoroughly capture the differences among various features and achieve accurate classification and prediction for the chips under test.

3.1 Parameter Set Selection

3.1.1 Data Preprocessing

Sometimes the data contains irrelevant observations that affect the performance of the prediction model. Data preprocessing is crucial to improve the data quality and model performance [21]. In the research, when a certain parameter item detects a faulty chip during the ATE test wafer, the subsequent test content of the chip will not be carried out. To address this issue, the researchers supplemented the

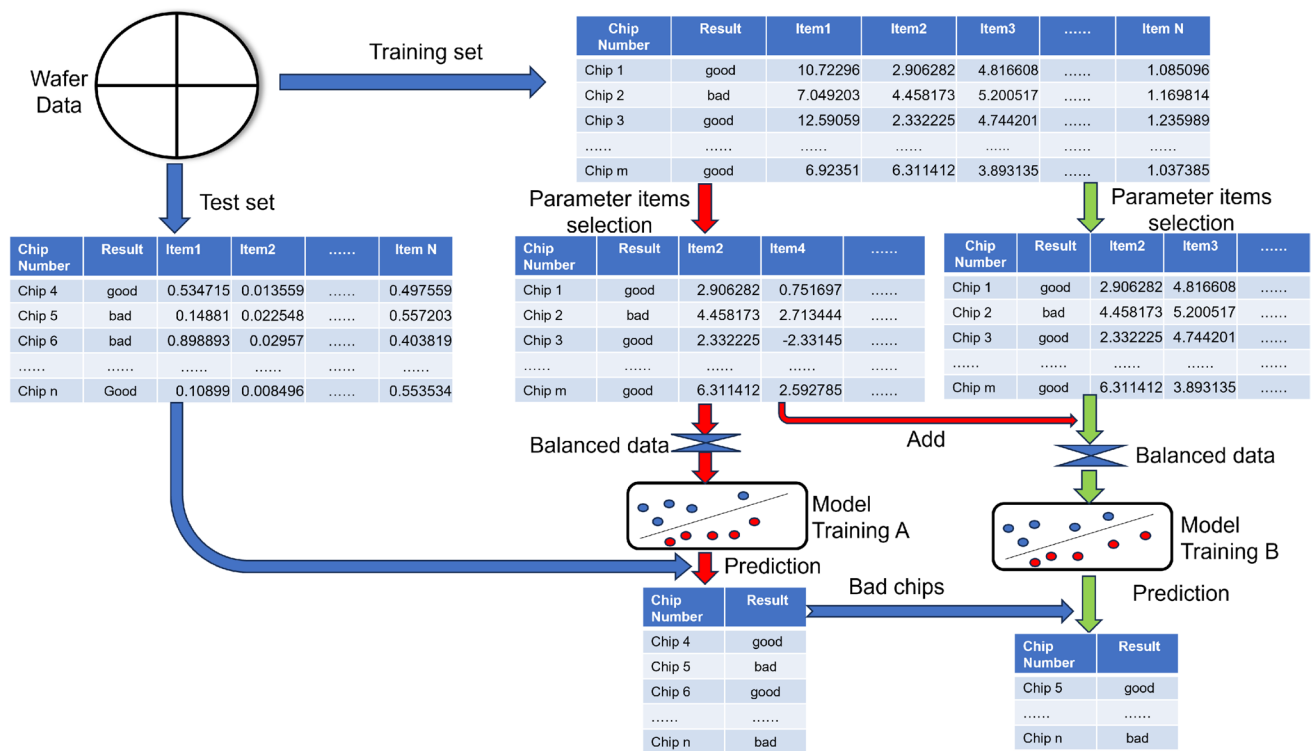


Fig. 1 Quality prediction model establishment

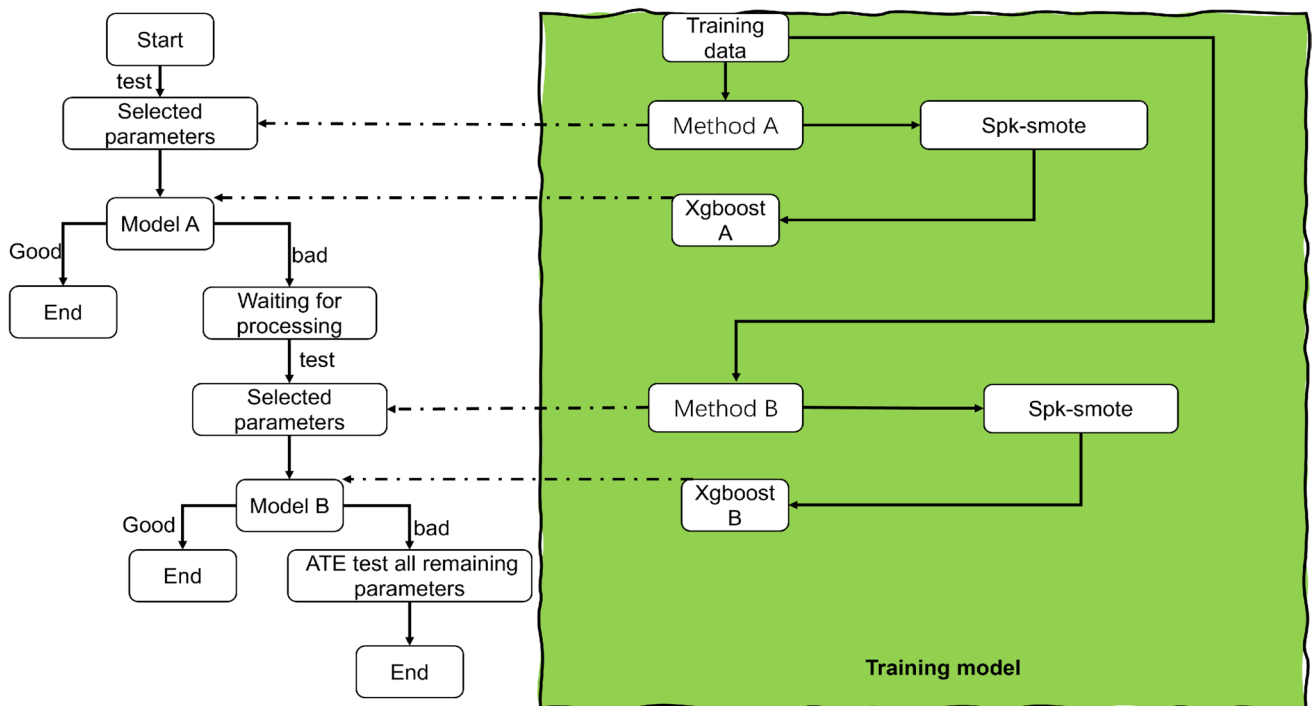


Fig. 2 SLvT adaption test flow

data vacancy value. Research findings indicate that in the wafer test, whether a chip fails is related to the failure of the surrounding chips. At the wafer failure level, the defect source often causes the faulty chips on the wafer to gather together[22].

In the research method employed, the knn filling method is adopted. This approach involves using adjacent chips to fill the missing parts of the data, thereby making the filling value close to the original value.

3.1.2 Parameter Set Selection

(1) Selection method A The purpose of parameter set screening is to obtain a test set that meets low test escape requirements while also being cost-effective to test. Concurrently, parameter selection, which is essentially variable selection in statistics and machine learning, is a vital aspect of modeling. It [23] significantly impacts the model's prediction ability, generalization, computational efficiency, and interpretability.

In the study depicted in Fig. 3, two distinct parameter selection methods were employed for parameter items: RFECV and Pearson. These methods were utilized to identify important parameter items of the wafer during the training phase. The RFECV method maximizes the utilization of the relationship between AUC and parameter items, utilizing XGboost to facilitate parameter selection.

RFECV represents a wrapper feature selection method aimed at identifying the most pertinent features for a given test parameter set. To ensure its robustness, RFECV synergizes recursive feature elimination [24] and cross-validation [25] to determine the optimal number of features that maximize model performance.

In RFECV, a classification machine learning model is employed to score each feature and iteration, subsequently discarding features that fail to improve classification accuracy. The feature search process operates via backward selection, commencing with the complete feature set and progressively eliminating features that do not contribute

Fig. 3 Filtering Method A Flow

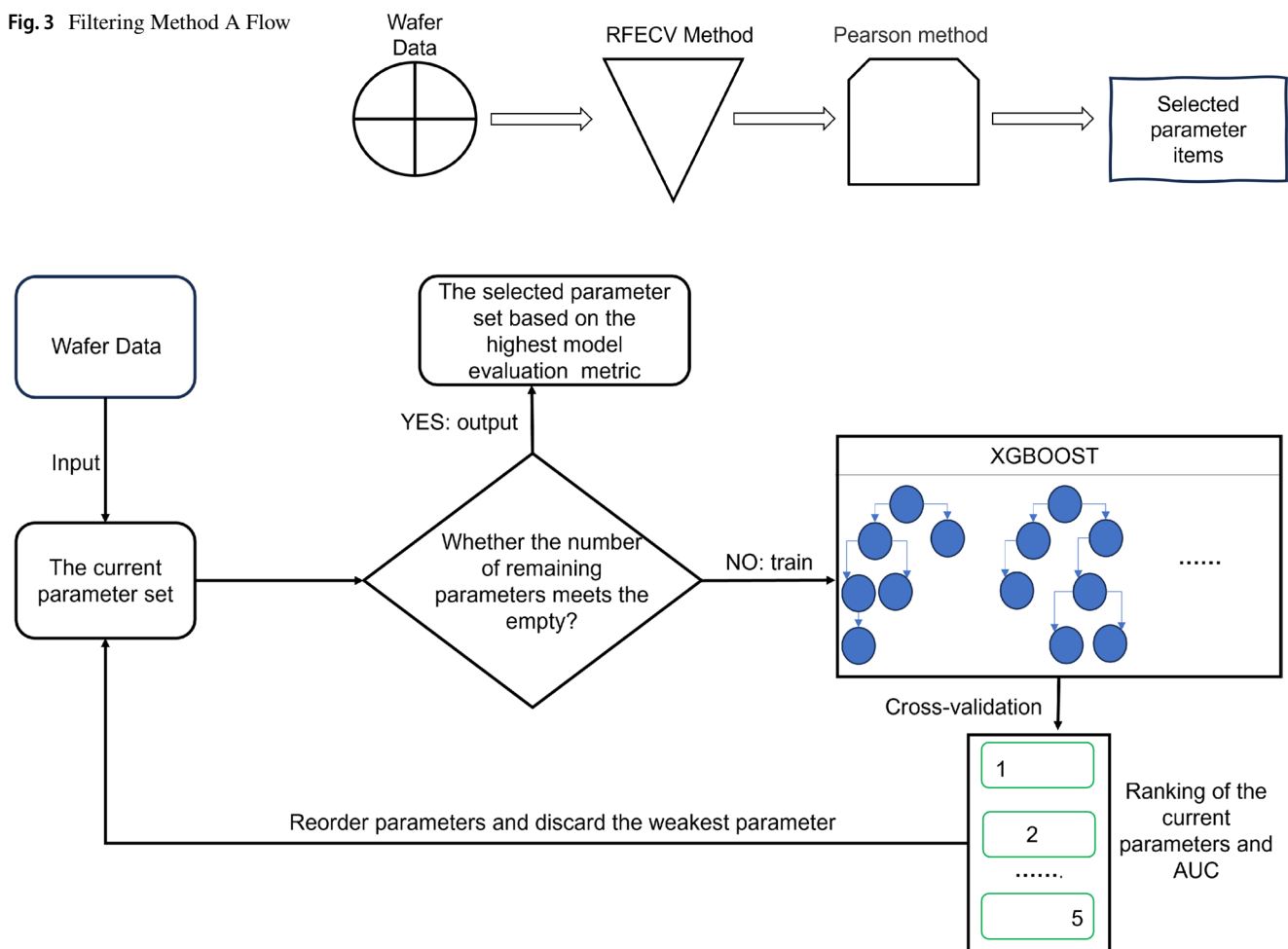


Fig. 4 RFECV method parameter set selection flow

significantly to classification accuracy. This iterative process culminates in the identification of the most effective feature subset.

In the research depicted in Fig. 4, RFECV is implemented with XGBoost serving as the classifier. The cross-validation fold (k) is set to 5, utilizing StratifiedKFold as the splitting strategy to maintain a consistent sample percentage for each class. By employing fivefold cross-validation, the dataset is divided into five equally sized folds, thereby enabling a robust evaluation and selection of features.

Pearson correlation coefficient is a measure of linear relationship between two random variables. Historically, it is the first formal measure of correlation. It is still one of the most widely used relationship measures [26].

Pearson correlation coefficients X and Y of two variables are defined as the product of the covariance of two variables divided by their standard deviation, which can be expressed as

$$r_{xy} = \frac{\sum (x_i - \bar{x}) \sum (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} \quad (1)$$

where $\bar{x} = \frac{1}{n} \sum_{i=1}^N x_i$ is the average of X , which $\bar{y} = \frac{1}{n} \sum_{i=1}^N y_i$ is the average of Y . The r_{xy} range is -1 to 1.

If $r_{xy} > 0$, it means that the two variables are positively correlated, that is, the larger the value of one variable, the larger the value of the other variable; if $r_{xy} < 0$, it means that the two variables are negatively correlated; when $r_{xy} = 0$, it means that x and y are not correlated. The larger the absolute value of the correlation coefficient, the stronger the correlation; the closer the absolute value of the correlation coefficient is to 0, the weaker the correlation. Measuring the correlation between features and categories can eliminate irrelevant features.

(1) Selection method B Figure 5 shows the process of selection method B in determining the parameter items to

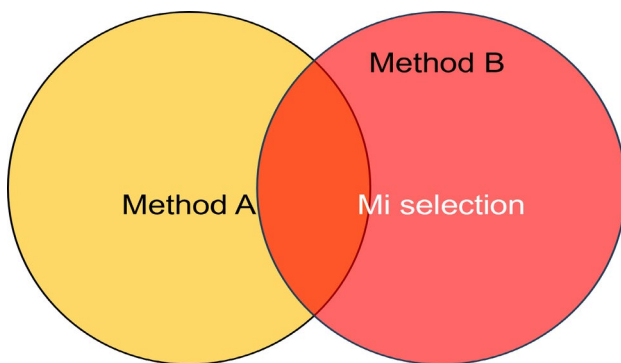


Fig. 5 Filtering Method B Flow

increase the test. First, the increased parameter items are determined through MI, and then the parameter set outside the intersection is selected with the parameter set in method A. MI is a symmetric measure of the amount of information contained in each other by two variables X and Y [27]. One of the advantages of MI for parameter item selection is that it can detect the nonlinearity relationship between variables. This method focuses on the joint correlation and redundancy between parameter items.

MI X and Y of two random variables are defined as

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) = H(Y) - H(Y|X) \\ &= H(X) + H(Y) - H(X;Y) \end{aligned} \quad (2)$$

Here $H()$ is entropy, $H(X|Y)$ and $H(Y|X)$ are conditional entropy, where $H(X;Y)$ is joint entropy.

$$H(X) = - \int_x p_X(x) \log p_X(x) dx \quad (3)$$

$$H(Y) = - \int_y p_Y(y) \log p_Y(y) dy \quad (4)$$

$$H(X;Y) = - \int_x \int_y p_{X,Y}(x,y) \log p_{X,Y}(x,y) dx dy \quad (5)$$

Here $p_{X,Y}(x,y)$ is the joint probability density function, $p_X(x)$ and $p_Y(y)$ are the marginal density functions of X and Y , and the marginal density function is

$$p_X(x) = \int_y p_{X,Y}(x,y) dy \quad (6)$$

$$p_Y(y) = \int_x p_{X,Y}(x,y) dx \quad (7)$$

Simultaneous upper formula, MI equation is

$$I(X;Y) = \int_x \int_y p_{X,Y}(x,y) \log \frac{p_{X,Y}(x,y)}{p_X(x)p_Y(y)} dx dy \quad (8)$$

In parameter selection, mutual information is a measure of the degree of interdependence between parameter items and test results. From the previous definition, it can be seen that the greater the mutual information value, the greater the degree of interdependence between parameter items and test results. That is, the greater the mutual information between a certain parameter item and test results, the more non-determinism of test results is reduced after a certain feature is known, indicating that the "stronger the correlation" between the two is. If the mutual information between the two is 0, the two are independent and have little relationship.

Here, the data is first converted into percentiles to convert the distribution of parameter items into a nearly uniform

distribution. This conversion can also be regarded as a data standardization method, making the comparison and subsequent processing between different parameter items easier. Percent conversion processing can make the distribution of features more standardized and provide a more robust and reliable basis for subsequent data analysis and modeling.

The proposed method standardizes the calculated mutual information, and the formula is as follows

$$scorex_i = \sum_y I(x_i, Y) \quad (9)$$

$$Nscorex_i = \frac{scorex_i - \mu}{\sigma} \quad (10)$$

where $Nscorex_i$ is the standardized mutual information score of the i -th parameter item, the original mutual information score of the i -th parameter item, denoted as $scorex_i$, encapsulates the overall relevance of the parameter (x_i) to the test result (Y) by condensing the mutual information (MI) into a single, comprehensive value. This score serves as a metric for assessing the significance of the parameter in relation to the outcome under consideration, μ is the average of mutual information scores for all parameter items, σ is the standard deviation of the mutual information score of all parameter items.

This process converts the mutual information score into a distribution with zero mean (mean 0) and unit variance (standard deviation 1), allowing for fairer comparisons between the mutual information scores of different features. In this way, the normalized score visually shows which features have a higher or lower amount of information shared relative to the average, helping to identify the features that are most useful for predicting the target variable.

At the same time, the mentioned method only takes parameters other than 2σ and focuses on those parameters that may have more information, so that the model can perform better in prediction.

3.1.3 Description of Selection

Regarding the performance of MI (mutual information) selected parameters in RFECV (Recursive Feature Elimination Cross Verification) and Pearson Correlation Analysis, and why these parameters were not all selected in RFECV and Pearson, we can discuss the following aspects:

1)The working principle of the three selection methods

RFECV is a feature selection method that combines Recursive Feature Elimination (RFE) and Cross Validation (CV). It evaluates the importance of features by repeatedly fitting the model and weeding out the least important

features until a specified number of features or an ideal level of model performance is reached.

Pearson correlation measures the linear relationship between two variables. Features highly correlated with the target variable might be selected based on this criterion, but it does not capture non-linear relationships or interactions between features.

MI quantifies the shared information between features and the target variable, emphasizing the relevance of each feature. However, a high MI value does not necessarily translate to a significant improvement in model performance, as it may be influenced by data distribution, noise, or non-linear relationships.

2)The focus of different analytical methods

In RFECV, the importance of features is usually evaluated by the prediction performance of the model, such as auc, mean square error, etc. This means that even if a feature shows a high mutual information value in MI analysis, it may be eliminated in RFECV if it has little effect on improving the prediction performance of the model.

MI is focused on measuring the amount of information shared between variables, MI-selected parameters may be more suitable for nonlinear models or complex data relationship analysis.

The Pearson correlation coefficient focuses on measuring the linear relationship between variables.

Furthermore, our rationale for selecting these three methods stems from the following considerations:

1)Given that our prediction model is XGBoost, its predictive accuracy inherently relies on the quality and nature of the data. To optimize model performance, we opted for Recursive Feature Elimination with Cross-Validation (RFECV), integrating XGBoost itself within this framework to meticulously identify the most suitable parameters. This approach ensures that the model's architecture is tailored to the specific nuances of our dataset.

2)Recognizing that the initial set of screened parameters did not effectively diminish the parameter count, particularly in light of the substantial sample size encountered during preliminary predictions, we deemed it essential to narrow down to parameters exhibiting simpler, more manageable relationships. Consequently, we employed Pearson correlation analysis to filter out parameters with strong linear dependencies, thereby mitigating potential biases or overfitting issues and refining the model's focus.

3)Finally, with the remaining feature set already enriched with robust linear relationships, we shifted our attention to capturing nonlinear intricacies. To this end, we utilized Mutual Information (MI) as a selection criterion. MI excels at identifying nonlinear associations, enabling us to select parameters that not only demonstrate a strong

link to the target variable but also enrich our model with nonlinear insights, thereby enhancing its predictive prowess and versatility.

Understanding these fundamental differences is vital when interpreting the performance of features selected by each method. Each technique offers unique insights into the data and its predictive potential, necessitating a holistic approach to feature selection that considers multiple criteria and methodologies.

3.2 Quality Model Training

3.2.1 Data Balancing

Min–Max Normalization The existence of high values in different features will affect the learning process of machine learning classifiers, and training high dimensional data sets requires a lot of computational resources. In order to solve these problems, we choose to use various normalization methods [28], such as Min–Max normalization, Z-fractional normalization, decimal scaling or maximum normalization. The choice of methods usually depends on the application scenario. In this step, use the following formula to apply Min–Max to the data set.

$$X_{i-norm} = \frac{X_i - X_{i-min}}{X_{i-max} - X_{i-min}} \quad (11)$$

where X_i is the data of the original parameter item i , X_{i-min} and X_{i-max} are the minimum and maximum values of the parameter item i in the dataset, respectively, and X_{i-norm} is the normalized data.

This scaling technique maps raw values to a range between 0 and 1, preserving the relative relationship between data points.

It is important to note that normalization is not applied in the feature selection phase, as the machine learning model used in this phase is not sensitive to high numerical values of features. However, the training and test datasets used to train the machine learning classifier are normalized to ensure accurate and reliable model performance.

Data Balance Processing SMOTE (Synthetic Minority Over-sampling Technique) algorithm uses interpolation strategy to synthesize minority samples and increase the probability of occurrence of minority data [29]. This algorithm can overcome the imbalance of the sample set and is suitable for processing wafer data. For a minority sample set $X = \{x_i | i = 1, 2, \dots, M\}$, for a certain sample x , the Euclidean distance is used as the standard to calculate its distance

to all samples in the set X to obtain the k -nearest neighbors of each sample. When the magnification of up-sampling is N , N nearest neighbors $\{\hat{x}_i | j = 1, 2, \dots, N\}$ are randomly selected from its K nearest neighbors; Random linear interpolation between the set X and $\{\hat{x}_i | i = 1, 2, \dots, N\}$ is used to synthesize new samples x_{new}

$$x_{new} = x_i + rand \cdot (\hat{x}_i - x_i) \quad (12)$$

In the formula, $rand$ represents a random number between 0 and 1. Finally, the new sample is combined with the original minority sample set to obtain a new data set. SMOTE is a technique for dealing with data imbalances, primarily used in machine learning and data science. However, CPK is an indicator to measure the capacity of the production process and is often used in manufacturing and quality control.

$$CPK = \min\left(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right) \quad (13)$$

In the formula, μ is the process average value, σ is the process standard deviation, USL is the upper specification limit, and LSL is the lower specification limit. By combining CPK (Complex Process Capability Index) and SMOTE, it can ensure that the generated synthetic samples are more in line with common sense in terms of quality control and avoid introducing unnecessary deviations. However, CPK and SMOTE are the contents of different fields, which may lead to over-synthesis of samples, thus making the synthesized minority samples too concentrated in certain specific areas, resulting in the model overfitting in this area, thereby reducing the generalization ability of the model. For this reason, the method here is proposed as

$$SPK = \max\left(\frac{DUP - \mu}{6\sigma}, \frac{\mu - DLO}{6\sigma}\right) \quad (14)$$

where μ is the process average, σ is the process standard deviation, DUP is the maximum value of the data parameter item, and DLO is the minimum value of the data parameter item. SPK (Side Process Capability Index) focuses on the distribution characteristics of each parameter itself, which helps to generate more realistic and realistic synthetic samples. Figure 6 shows the control range of CPK and SPK in data distribution. In the CPK diagram, it is evident that data points are distributed even beyond the designated USL and LSL. It is crucial to note that CPK is utilized not merely for independently assessing the stability and capability of the production process, but also for safeguarding data integrity. Regarding SPK, given its definition rooted in the actual upper and lower limits of the data, it is capable of encompassing all sample data comprehensively, which holds particular significance in scenarios involving rare samples, such as defective chips,

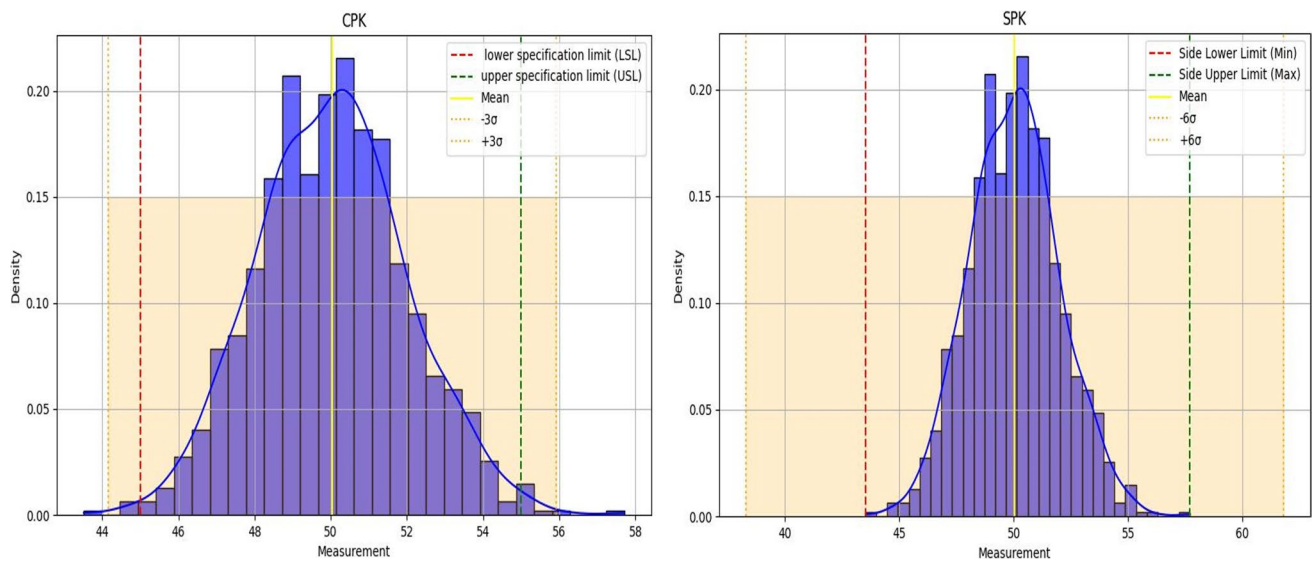


Fig. 6 CPK and SPK data control range

and uneven defect distributions stemming from various factors. Conversely, 6σ represents a management strategy and methodology aimed at substantially enhancing product quality and drastically lowering defect rates. Its fundamental objective is to attain a near-zero defect production state through a series of systematic process improvements, typically defined as no more than 3.4 defects per million opportunities. In comparison, 3σ standards fall short of comprehensively addressing all potential quality fluctuations, making 6σ a more rigorous and comprehensive quality control framework.

In order to give full play to the advantages of SPK, we combine it with SOMTE and have obvious advantages in data generation. Table 1 below is the spk-smote algorithm.

3.2.2 Model Training

The XGBoost algorithm proposed by Chen and Guestrin [30] in 2016 attracted attention for its success in machine learning competitions. XGBoost is a hoisting decision tree integration machine learning algorithm that uses layer boosting. XGBoost trains a new decision tree at each iteration to improve the performance of the current decision tree.

Here, given a wafer data, the dataset is expressed as Φ . Where $\Phi = \{(x_i, y_i) : x_i \in R^m, y_i \in R\}$ There are n chips and m parameter items, x_i represents the parameter item of the i -th chip, and y_i represents the test result of the i -th chip.

The chip test predictions using the regression tree (CART) set are as follows:

Table 1 Algorithm SPK-SMOTE

Algorithm SPK- SMOTE

1. Given the set of a minority class, S ; the number of nearest neighbors, k ;
2. Specify a random sample $x_i \in S$;
3. Seek k nearest neighbors of x_i
4. Randomly choose one of k nearest neighbors, \hat{x}_i ;
5. Calculate the SPK for each parameter
6. Create the synthetic sample by $x_{\text{new}} = x_i + (\hat{x}_i - x_i) \cdot \delta + \partial * z * SPK$, where $\delta \in [0,1], \partial \in \text{constant}, z \in \text{random matrix}$

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i) f_k \in F \quad (15)$$

In the equation, each function represents an independent regression tree, and F represents a set of CARTs. $f_k(x_i)$ is the predicted value of the i -th chip of the data on the k -th CART, and \hat{y}_i represents the final prediction result. Unlike the original layer lifting algorithm, XGBoost aims to minimize the regularization objective function defined as follows

$$Ob_j = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (16)$$

where c represents the test result of the i -th chip, and s represents the loss function. F is the regularization term, which helps to prevent the model from overfitting.

The loss function is used to define the degree of inconsistency between the predicted value and the true value of the model and determine the effect of the training model. The composition of the loss function is selected according to the application characteristics of the integrated circuit.

$$loss = \alpha * TER + (1 - \alpha)TYL \quad (17)$$

where TER is the test escape rate, which means that N_E is the proportion of chips that are predicted to be good for bad chips to all chips that are s , and the actual fault chip is obtained from the full test. The lower the test escape rate, the higher the test quality. The calculation formula is as follows

$$TER = \frac{N_E}{s} * 100\% \quad (18)$$

TYL is the loss of test yield, which refers to the proportion of N_L as the number of qualified chips predicted to be bad chips to all chips as s . The calculation formula is as follows

$$TYL = \frac{N_L}{s} * 100\% \quad (19)$$

It is worth noting that the XGBoost algorithm has built-in a series of objective functions, which are the sum of loss functions and regularization terms, and does not support custom objective functions. The purpose of model training is to obtain hyperparameters in the best state of the model prediction effect when the defined loss function Minimum.

For model A, we choose $\alpha = 1$ to greatly reduce test escape, and for model B, we choose $\alpha = 0$ to greatly reduce test yield loss.

The parameters obtained in the process of model training are called model parameters, and the parameters that specify their range or value before modeling are called hyperparameters, which serve as guidance for model training. For example, the learning rate and depth of the tree in decision tree model training are both hyperparameters. The setting of

hyperparameters will affect the performance of the model, and more accurate prediction results can be obtained through adjustment.

When the XGBoost model is actually used in the experiment, the following parameters are adjusted to make the model perform best:

1. $n_evaluations$: $n_evaluations$ is the number of iterations in training. Too large an $n_estimate$ is not appropriate because it will lead to overfitting, and too small an $n_estimate$ will lead to insufficient fitting, so that the model cannot give full play to its learning ability.
2. max_depth : It is the maximum depth of the tree. The deeper the tree, the more complex the tree model and the stronger the fitting ability, but at the same time, the model is also easier to overfitting.
3. $Learning_rate$: learning rate is a very important parameter in most algorithms and needs to be adjusted, as is the case in XGBoost. It greatly affects the performance of the model, and choosing the appropriate learning rate will make the model more robust.

To apply the method described in this article more efficiently, we introduce a grid search strategy. Grid search is a technique for finding optimal parameter settings by systematically traversing a combination of multiple parameters. This approach is particularly common and effective in hyperparameter tuning. Next, we will elaborate on the suggested search range to ensure that the potentially optimal parameter combinations can be explored comprehensively and in detail.

$n_estimators$: [100, 200, 300], max_depth : [3–5], $learning_rate$: [0.01, 0.1, 0.2].

4 Experimental Results

4.1 Experimental Platform Setup and Data Description

The experimental platform utilizes an AMD Ryzen 5 5600H with Radeon Graphics running at 3.30 GHz, coupled with 16 GB of DDR4 3200 MHz SDRAM high-bandwidth memory. For the simulation platform, Python 3.9 serves as the primary environment. The key libraries employed include NumPy, Pandas, and Scikit-learn, as depicted in Fig. 6.

The test data of integrated circuits usually involves high-level trade secrets. The data used in this experiment comes from the ICND2263 chip of Chizhou HISEMI-electronics. The types of test parameters include power supply voltage, current, output signal voltage amplitude, offset, etc. The dataset includes 10,911 chips, and each chip has 151 test parameters. Among them, there are 10,875 good products,

Fig. 7 Experimental environment

that is, chips that pass the test; the rest are non-conforming products, with a yield of about 99.6%.

4.2 Results of Parameter Selection

4.2.1 Method A

Figure 7 displays the results obtained from the method employing RFECV and Pearson correlation. Using the RFECV method, a parameter set with the highest AUC value, comprising 59 parameter items, was identified.

Among these items, coefficients greater than 0.03 indicate a discernible influence of a feature on the target variable. Even features with relatively weak correlations can impact the target variable to some extent, aiding in data understanding and predictive model construction. Filtering out irrelevant features: By selecting coefficients greater than 0.03, features with weaker associations with the target variable are excluded, thereby simplifying the model and mitigating overfitting risks. This enhances the model's generalization ability, focusing on pivotal data features. Overall, a Pearson correlation coefficient exceeding 0.03 suggests a linear

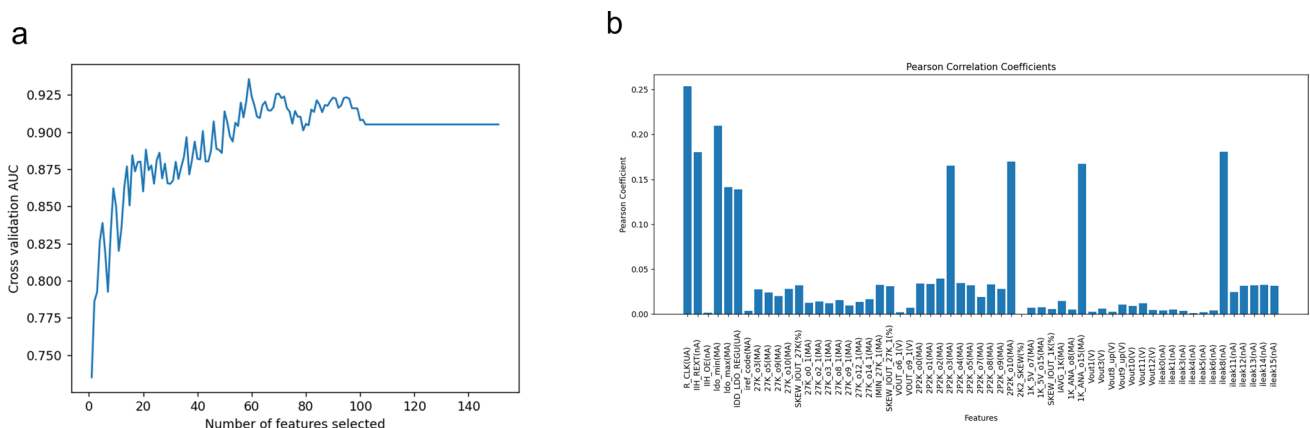
**Fig. 8** Parameter set selection results, a. RFECV selection process, b. Pearson selection process

Table 2 Data set

Data Set	majority	minority	parameters
Data A	98.5%	1.5%	22
Data B	98.5%	1.5%	25

relationship between the feature and the target variable, contributing predictive power and warranting consideration in feature selection. Pearson identified 23 parameters, which were utilized as predictors in constructing a machine learning model for chip quality prediction.

4.2.2 Method B

As depicted in Fig. 8, the MI method was employed to select data outside of 2σ , resulting in the addition of three parameter items. Building upon the selection from method 1, both the parameter items chosen by method A and those selected by method B are utilized as predictors. This approach enhances the accuracy of the machine learning model from the dataset.

4.3 Data Balance Results

Using the chip data in this article as an illustrative example, here are the specific details pertaining to the data. The majority class comprises 98.5% of the instances, whereas the minority class accounts for merely 1.5%. The distinguishing factor between these two classes lies in the number of parameters they possess; the majority class has 22 parameters, whereas the minority class possesses 25 parameters. Table 2 below is Specific information about the data set.

To more effectively demonstrate the merits of SPK-SMOTE, we employ three key metrics: Intra-Cluster Distance, Average Inter-Cluster Index, and DB Index. Here's a refined description of Intra-Cluster Distance:

4.3.1 Intra-Cluster Distance

This metric quantifies the average distance between samples within the same cluster, serving as an indicator of the compactness or tightness within a cluster. A smaller Intra-Cluster Distance signifies that the samples within the cluster are closer together, indicating a higher degree of cohesion and a more distinct cluster formation. The formula for calculating Intra-Cluster Distance is as follows:

$$S_i = \frac{1}{m_i} \sum_{j=1}^{m_i} d(x_{i,j}, c_i) \quad (20)$$

where $x_{i,j}$ represents the j -th sample point in the i -th cluster, and $d(x_{i,j}, c_i)$ represents the distance from the sample point

$x_{i,j}$ to the cluster center c_i , where d can be Euclidean distance, Manhattan distance, etc. The i -th cluster contains m_i sample points, m_i represents the number of samples in the i -th cluster.

4.3.2 Inter-Cluster Distance

Measure the average distance between different clusters and use it to evaluate the degree of separation between clusters. The larger the distance, the better the degree of separation between clusters. The formula is as follows:

$$D_{ij} = \frac{1}{m_i \times m_j} \sum_i \sum_j d(x_i, x_j) \quad (21)$$

$m_i \times m_j$ represents the number of sample point pairs of the i -th cluster and the j -th cluster, and $d(x_i, x_j)$ represents the distance between all sample point pairs of the i -th cluster and the j -th cluster.

4.3.3 DB Index (Davies-Bouldin Index)

The synthesis effect is evaluated by calculating the ratio of the degree of dispersion of samples in the cluster to the degree of separation of the nearest cluster between clusters. The smaller the DB index, the better the synthesis effect. The formula is as follows:

$$DB = \frac{1}{n} * \sum_{i=1}^n \frac{S_i}{\min_{j \neq i} D_{ij}} \quad (22)$$

where n is the number of clusters.

Upon examining the index of average intra-cluster distance, it becomes evident that SPK-SMOTE outperforms CPK-SMOTE in both Dataset A and B. Notably, even as the complexity of the dataset escalates, SPK-SMOTE maintains a lower index of 0.13 compared to CPK-SMOTE, highlighting the tighter clustering of specific classes within the data synthesized by SPK-SMOTE.

Furthermore, the index measuring the average inter-cluster distance underscores SPK-SMOTE's superiority over CPK-SMOTE by a margin exceeding 0.1 in both datasets.

Table 3 Calculation results of indicators

Indicators	Data Set	SPK-SMOTE	CPK-SMOTE
Intra	Data A	0.6063	0.8066
Cluster Distance	Data B	0.7141	0.8472
Inter	Data A	0.1944	0.0779
Cluster Distance	Data B	0.1841	0.0791
DB Index	Data A	6.2380	20.7044
	Data B	7.7578	21.4110

This signifies that SPK-SMOTE effectively ensures a healthy separation between a subset of samples and the majority, fostering favorable conditions for machine learning modeling.

A comparative analysis of the DB index reveals a striking disparity, with CPK-SMOTE's value exceeding that of SPK-SMOTE by more than twofold. This disparity underscores the inferior quality of data synthesized by CPK-SMOTE, characterized by insufficient proximity among same-label data and inadequate distinction between data of different labels. Such shortcomings hinder effective data modeling and analysis, as further elaborated in subsequent sections of the article. Table 3 below is the Results of comparison.

4.4 Model Training Results

With effective parameter selection, exceptional results can be achieved during model training. In order to illustrate the advantages of SPK-SMOTE, in the methods mentioned, they are compared with CPK-SMOTE, SMOTE, and No treatment respectively.

TRO (test resource occupancy rate): The ratio of the product of the number of tests s_i and the number of test items p_i to the product of the total number of test items p and the total number of tests s . The smaller the ratio, the smaller the test resource occupancy, and the lower the test cost.

$$TRO = \frac{\sum p_i * s_i}{p * s} * 100\% \quad (23)$$

In the comparison presented in Fig. 9, it is evident that SPK-SMOTE occupies test resources relative to CPK-SMOTE has decreased by 2.3%, accounting for only 32.5%,

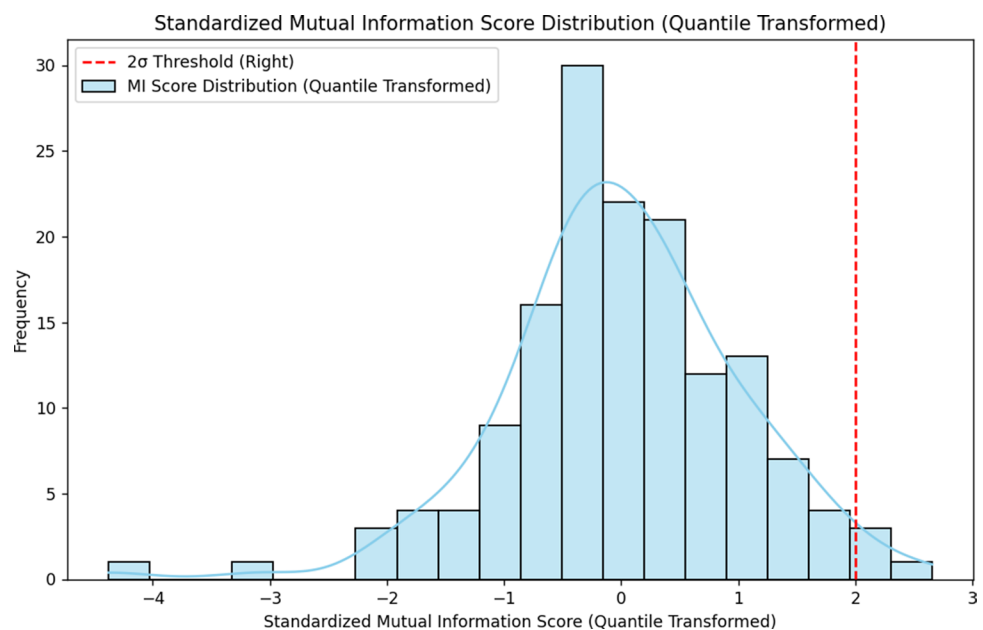
while the occupation rate of test resources by SMOTE is notably higher at 51.5%. Despite the low occupation of untreated test resources, it demonstrates poor performance on test escape. This is apparent from Fig. 10, where its test escape rate is as high as 0.32%. In contrast, the use of SPK-SMOTE significantly reduces test escape, surpassing CPK-SMOTE with only 0.09% test escape. This indicates that the method, inclusive of SPK-SMOTE as a performance scheme, has achieved a 67.5% reduction in test resources compared to traditional methods, with just a 0.09% sacrifice in test escape, thereby substantially lowering test costs.

Simultaneously, compared with the self-adaption algorithm mentioned in Document [14], Fig. 11 illustrates that the mentioned method reduces test resources by 0.6%, slightly higher than the method described in the literature in terms of test escape. However, in terms of test yield loss, the Mentioned method achieves zero yield loss, which significantly surpasses the 83.5% yield loss observed in the comparison method. A 0.6% reduction in test resources might indeed seem modest in terms of an immediate return on investment (ROI) for some users. In the context of cost–benefit analysis, such a small improvement might not initially seem compelling enough to justify the effort required to implement the approach.

However, it's important to consider several factors that could shift the balance in favor of pursuing this approach:

Cumulative Effect: While 0.6% may seem insignificant in isolation, the impact can be substantial when applied across a large scale or over an extended period. For example, in a manufacturing environment with high-volume production, even small percentage reductions in test resources can translate into significant cost savings over time.

Fig. 9 MI method selection process



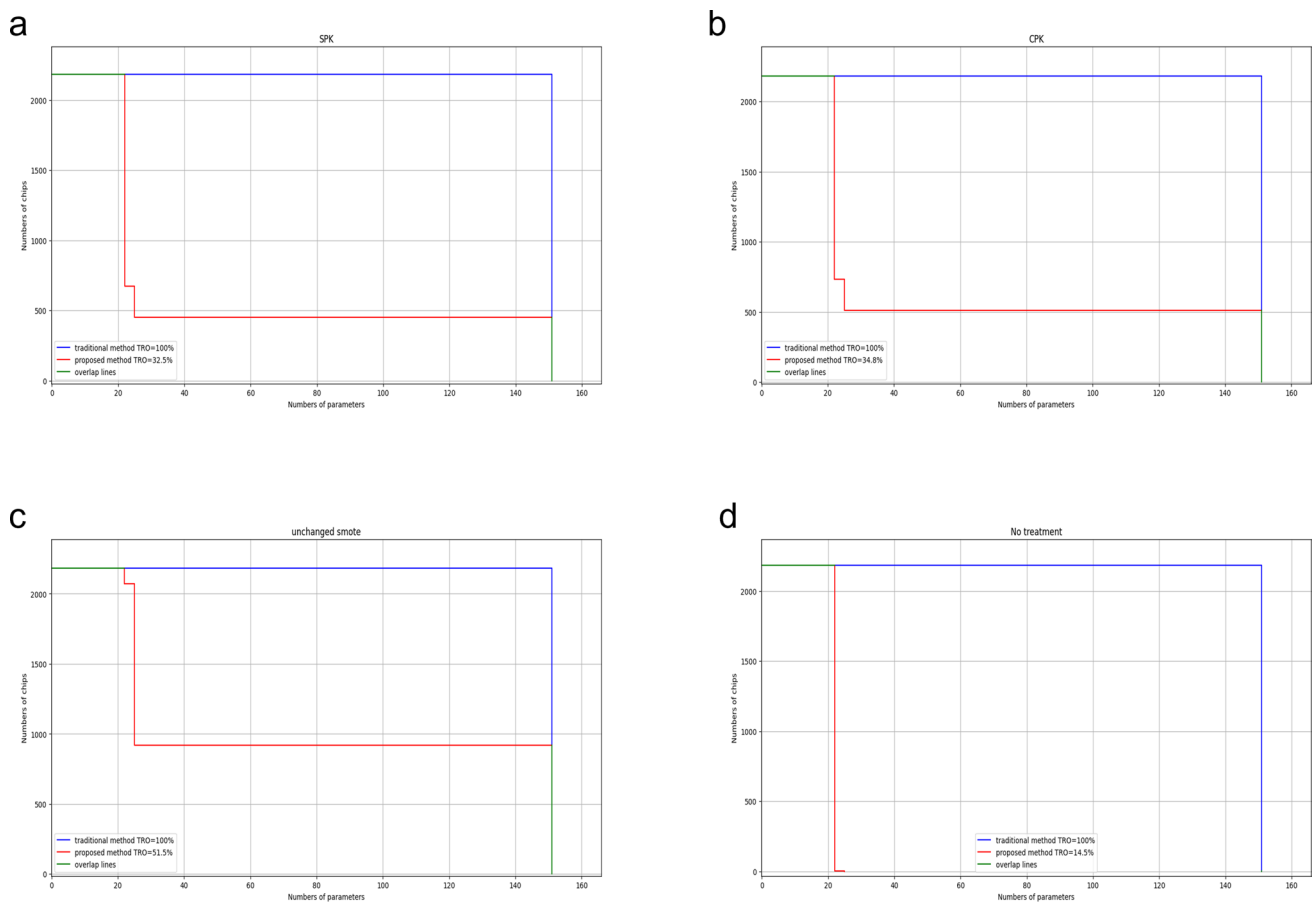
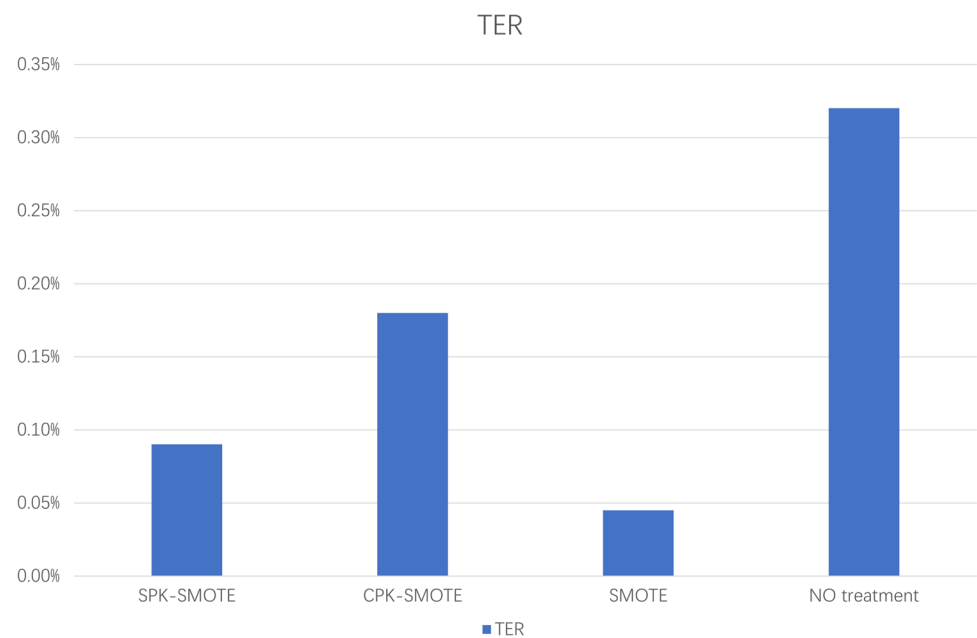


Fig. 10 Test resource occupancy rate results, a. SPK-SMOTE test resource occupancy rate performance; b. CPK-SMOTE test resource occupancy rate performance; c. SMOTE test resource occupancy rate performance; d. No processing test resource occupancy rate performance

Fig. 11 Test escape rate performance



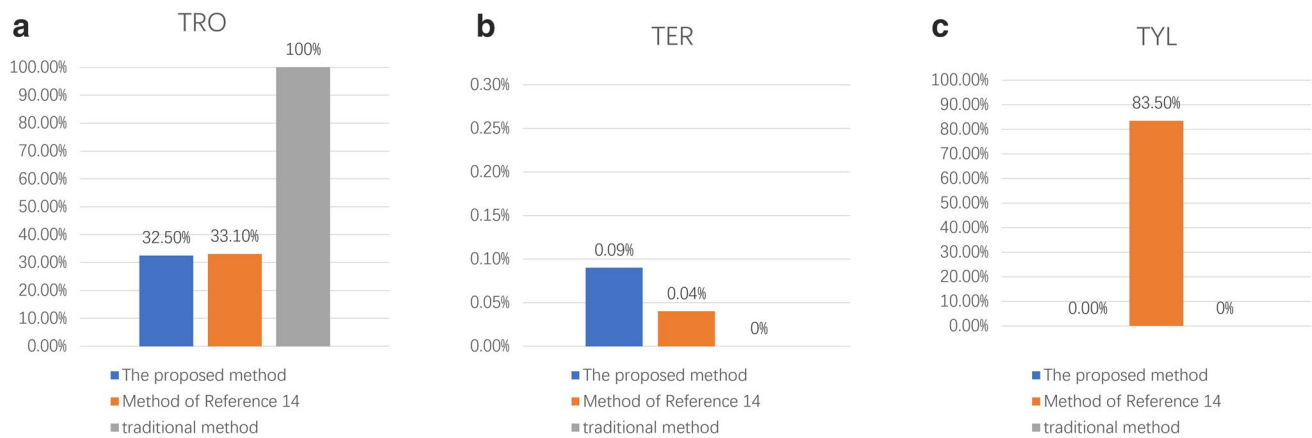


Fig. 12 Comparison results of literature 14, a. test resource occupancy performance; b. test escape performance; c. test yield loss performance

Qualitative Benefits: Besides the direct cost savings, there might be qualitative benefits that are harder to quantify but equally important. For instance, reducing test time can speed up the overall production cycle, improving responsiveness to market demands and potentially enabling faster product iterations.

Risk Mitigation: Overkill can lead to unnecessary yield loss and increased costs. By optimizing test resources, users can potentially minimize this risk while still ensuring product quality.

Future-Proofing: Adopting an approach that prioritizes efficiency and resource optimization can position a company to better adapt to future challenges, such as increased competition, stricter regulations, or shifts in consumer preferences.

In summary, while a 0.6% reduction in test resources may seem modest at first glance, its true value depends on the specific context and long-term implications. By considering the cumulative effect, qualitative benefits, and risk mitigation potential, users can make informed decisions about whether to pursue this approach or explore alternatives that offer a more significant ROI Fig. 12.

5 Conclusion

Aiming at the parameter test in wafer testing, this paper proposes a SLvT method for the integrated circuit test parameter set. For the parameter test items, a variety of selection methods are used to find the appropriate parameter set, so as to obtain the parameter items that need to be tested, and SPK-SMOTE is used to achieve data balance, and XGBoost is used to achieve classification prediction. Compared with the original set of test items, the set of

parameter test items that need to be tested is smaller, thus reducing the test cost, and the test escape rate has been kept at a low level. At the same time, it has achieved the same excellent performance as the traditional test method in the loss of test yield. Indeed, the current performance in terms of Defective Parts Per Million (DPPM) is below expectations. Enhancing this metric and minimizing test escape rates are pivotal to ongoing efforts. Within the context of this discussion, the main focus is on optimizing test resources and achieving efficient test recovery. Maintaining a manageable level of test escape while pursuing these goals is crucial, recognizing that achieving low DPPM is a multifaceted challenge intricately tied to both the testing process and the testing objective.

Author Contribution Qiong Wu: Conceptualization, Methodology, Investigation, Kaiming Hao: Software, Validation, Formal analysis, Wenfa Zhan: Conceptualization, Methodology, Writing – review & editing, Project administration.

Funding This work was supported by the National Natural Science Foundation of China (62474002, 61306046, 61640421), the Yicheng Elite Project (202371), the Open Project of National Local Joint Engineering Laboratory of RF Integration and Micro-assembly Technology (KFJJ20230101), the National Key Laboratory of Integrated Chips and Systems Project (SLICS-K202316), the Anhui University Research Project (2023AH050481), and the Research on Testing Methods and Accuracy of High Frequency Signal Chips (2023AH050500) and the Graduate Education Quality Engineering Project of Anqing Normal University(2023xcysj137).

Data Availability The data that has been used is confidential.

Declarations

Competing Interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Zheng Y, Ling D, Wang YW, Jang SS, Tao B (2016) Model quality evaluation in semiconductor manufacturing process with EWMA run-to-run control. *IEEE Trans Semicond Manuf* 30(1):8–16
- Ruth J, Berndt R (2016) Quality control for ultrafiltration of ultrapure water production for high end semiconductor manufacturing. In: *Proc. of 2016 27th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)*. IEEE, pp 16–22. <https://doi.org/10.1109/ASMC.2016.7491077>
- Tan F, Pan T, Li Z, Chen S (2015) Survey on run-to-run control algorithms in high-mix semiconductor manufacturing processes. *IEEE Trans Industr Inf* 11(6):1435–1444
- Moyne J, Samantaray J, Armacost M (2016) Big data capabilities applied to semiconductor manufacturing advanced process control. *IEEE Trans Semicond Manuf* 29(4):283–291
- Zhu Q, Wu N, Qiao Y, Zhou M (2016) Optimal scheduling of complex multi-cluster tools based on timed resource-oriented Petri nets. *IEEE Access* 4:2096–2109
- Yang F, Wu N, Qiao Y, Zhou M, Su R, Qu T (2017) Petri net-based efficient determination of optimal schedules for transport-dominant single-arm multi-cluster tools. *IEEE Access* 6:355–365. <https://doi.org/10.1109/TVLSI.2017.2205027>
- Yilmaz E, Ozev S, Butler KM (2012) Per-device adaptive test for analog/RF circuits using entropy-based process monitoring. *IEEE Trans Very Large Scale Integr (VLSI) Systems* 21(6):1116–1128. <https://doi.org/10.1109/TVLSI.2012.2205027>
- Chien CF, Lee PC, Dou R, Chen YJ, Chen CC (2017). Modeling collinear WATs for parametric yield enhancement in semiconductor manufacturing. In: *Proc. of 2017 13th IEEE Conference on Automation Science and Engineering (CASE)*. IEEE, pp 739–743. <https://doi.org/10.1109/COASE.2017.8256192>
- Kim KJ, Kim KJ, Jun CH, Chong IG, Song GY (2018) Variable selection under missing values and unlabeled data in semiconductor processes. *IEEE Trans Semicond Manuf* 32(1):121–128
- Ahmadi A, Nahar A, Orr B, Past M, Makris Y (2016) Wafer-level process variation-driven probe-test flow selection for test cost reduction in analog/RF ICs. In: *Proc. of 2016 IEEE 34th VLSI Test Symposium (VTS)*. IEEE, pp 1–6. <https://doi.org/10.1109/VTS.2016.7477263>
- Chien CF, Wang WC, Cheng JC (2007) Data mining for yield enhancement in semiconductor manufacturing and an empirical study. *Expert Syst Appl* 33(1):192–198
- Wang J, Zhang J, Wang X (2018) A data driven cycle time prediction with feature selection in a semiconductor wafer fabrication system. *IEEE Trans Semicond Manuf* 31(1):173–182
- Yilmaz E, Ozev S (2008) Dynamic test scheduling for analog circuits for improved test quality. In: *Proc. of 2008 IEEE International Conference on Computer Design*. IEEE, pp 227–233. <https://doi.org/10.1109/ICCD.2008.4751866>
- Chen X, Zhao Y, Lü H, Shao X, Chen C, Huang Y (2021) A machine learning-based approach for failure prediction at cell level based on wafer acceptance test parameters. In: *Proc. of 2021 IEEE Microelectronics Design & Test Symposium (MDTS)*. IEEE, pp 1–5. <https://doi.org/10.1109/MDTS52103.2021.9476151>
- Li TS, Huang CL, Wu ZY (2006) Data mining using genetic programming for construction of a semiconductor manufacturing yield rate prediction system. *J Intell Manuf* 17:355–361
- Chen K, Chang PY, Yeh CH (2010) Wafer die yield prediction by heuristic methods. In: *Proc. of The 40th International Conference on Computers & Industrial Engineering*. IEEE, pp 1–4. <https://doi.org/10.1109/ICCIIE.2010.5668273>
- Huang CL, Wang CJ (2006) A GA-based feature selection and parameters optimization for support vector machines. *Expert Syst Appl* 31(2):231–240
- Katragadda V, Muthee M, Gasasira A, Seelmann F, Liao JH (2018) Algorithm based adaptive parametric testing for outlier detection and test time reduction. In: *Proc. of 2018 IEEE International Conference on Microelectronic Test Structures (ICMTS)*. IEEE, pp 142–146. <https://doi.org/10.1109/ICMTS.2018.8383784>
- Kuo YT, Lin WC, Chen C, Hsieh CH, Li JCM, Fang EJW, Hsueh SSY (2021) Minimum operating voltage prediction in production test using accumulative learning. In: *Proc. of 2021 IEEE International Test Conference (ITC)*. IEEE, pp 47–52. <https://doi.org/10.1109/ITC50571.2021.00012>
- Stratigopoulos HG (2018) Machine learning applications in IC testing. In: *Proc. of 2018 IEEE 23rd European Test Symposium (ETS)*. IEEE, pp 1–10. <https://doi.org/10.1109/ETS.2018.8400701>
- Fan C, Chen M, Wang X, Wang J, Huang B (2021) A review on data preprocessing techniques toward efficient and reliable knowledge discovery from building operational data. *Front Energy Res* 9:652801
- Ooi MPL, Sok HK, Kuang YC, Demidenko S, Chan C (2013) Defect cluster recognition system for fabricated semiconductor wafers. *Eng Appl Artif Intell* 26(3):1029–1043
- Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. *J Mach Learn Res* 3:1157–1182
- Jeon H, Oh S (2020) Hybrid-recursive feature elimination for efficient feature selection. *Appl Sci* 10(9):3211
- Browne MW (2000) Cross-validation methods. *J Math Psychol* 44(1):108–132
- Cohen I, Huang Y, Chen J, Benesty J, Benesty J, Chen J, ... Cohen I (2009) Pearson correlation coefficient. Noise reduction in speech processing, pp 1–4. https://doi.org/10.1007/978-3-642-00296-0_5
- Kraskov A, Stögbauer H, Grassberger P (2004) Estimating mutual information. *Phys Rev E* 69(6):066138
- Raju VG, Lakshmi KP, Jain VM, Kalidindi A, Padma V (2020) Study the influence of normalization/transformation process on the accuracy of supervised classification. In: *Proc. of 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, pp 729–735. <https://doi.org/10.1109/ICSSIT48917.2020.9214160>
- Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 16:321–357
- Chen T, Guestrin C (2016) Xgboost: A scalable tree boosting system. In: *Proc. of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp 785–794. <https://doi.org/10.1145/2939672.2939785>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Qiong Wu received the BTech degree in mathematics from Anqing Normal University, the M.E. degree in computer application from Hefei University of Technology, and the Ph.D degrees in electromagnetic field and microwave technology from Anhui University in 1989, 2005 and 2007 respectively. She is now the vice president of Anqing Normal University, a leading technology talent in Anhui Province, the fifth inspector of Anhui Province, a professional certification expert from the Ministry of Education, and a director of

the Anhui Electronics Society. She presided over the completion of one project of the National Natural Science Foundation of China and participated in more than 10 provincial and departmental scientific research projects such as the Anhui Natural Science Foundation and the Anhui Key Scientific Research Plan. He has published more than 10 academic papers in important journals at home and abroad, authorized one invention patent (the first inventor), won one third prize for science and technology in Anhui Province, one award for innovation and promotion of industry-university-research cooperation in China (industry-university-research cooperation innovation achievement award), and two third prizes for outstanding academic papers in natural science in Anhui Province. She is mainly engaged in higher education management, fault-tolerant calculation, integrated circuit testing methods, electromagnetic field theory and applied research.

Kaiming Hao received the BTech degree in mathematics from Yanbian University and is currently studying for the M.E. degree in mathematics from Anqing Normal University in 2019 and 2022 respectively. His

research interests are fault-tolerant computing, integrated circuit testing methods, clustering and classification methods.

Zhan Wenfa received The M.E. degree, from the School of electrical engineering and automation of Hefei University of Technology. He received The PhD degree from the School of Computer and Information, Hefei University of Technology in 2004 and 2009 respectively. He is now the vice president of the School of electronic engineering and intelligent manufacturing of Anqing Normal University. His main research directions are: test data compression, testability design, integrated circuit testing and network teaching. He presides over 2 items of the National Natural Science Foundation, has published more than 60 papers and won more than 20 national invention patents. In 2013, he was awarded the strategic development technology leader in Anhui Province, in 2015, he was awarded the reserve candidate for the academic technology leader in Anhui Province, and in 2019, he was awarded the Anqing Science and Technology Excellence Award.