A Robust Metric for Screening Outliers from Analogue Product Manufacturing Tests Responses

Shaji Krishnan
Analytical Research Department
TNO
Zeist, The Netherlands
Email: shaji.krishnan@tno.nl

Hans G. Kerkhoff
Testable Design & Testing of Integrated Systems
University of Twente, CTIT
Enschede, The Netherlands
Email: h.g.kerkhoff@utwente.nl

***Abstract*—Mahalanobis distance is one of the commonly used multivariate metrics for finely segregating defective devices from non-defective ones. An associated problem with this approach is the estimation of a robust mean and a covariance matrix. In the absence of such robust estimates, especially in the presence of outliers to test-response measurements, and only a sub-sample from the population is available, the distance metric becomes unreliable. To circumvent this problem, multiple Mahalanobis distances are calculated from selected sets of test-response measurements. They are then suitably formulated to derive a metric that has a reduced variance and robust to shifts or deviations in measurements. In this paper, such a formulation is proposed to qualitatively screen product outliers and quantitatively measure the reliability of the non-defective ones. The application of method is exemplified over a test set of an industrial automobile product.**

I. Introduction

Stringent quality requirements on the finished electronic products are continuously forcing the semiconductor industries, especially the automobile, to insert additional reliability tests in their production flow. For this purpose they subject their packaged devices to time consuming burn-in procedures and purchase expensive automatic test equipments. In order to simultaneously benefit from supplying reliable products to customers, stay competitive in the market and profit from their business, these industries are constantly searching for cheaper manufacturing and product reliability test methodologies that reduce the overall cost in their production flow, leaving no compromise to the quality of their products.

One of the most commonly applied solution to this problem is identification and elimination of latent devices as early in the production flow. Several statistical methods are applied to identify latent devices at wafer-level. Parts Average Testing (*PAT*) is one such first procedure introduced by the Automobile Electronic Council to identify abnormal parts from a population [1]. Later in time, methodologies like die-level predictive models, capable of predicting the reliability of the device, based on the failure rate of the neighboring dies, were introduced [2]. Some of the other techniques that have been proposed to identify unreliable devices at wafer sort are based on test-response measurements either from parametric tests like IDDQ/DeltaIDDQ [3], [4], supply ramp [5] or functional tests [6]. Test-response measurements from these wafer-level tests are statistically post-processed to screen outliers.

Outliers are identified by handling the data either in a univariate or multivariate space. In some situations, an inlier in the univariate space can be an outlier in the multivariate space. Outlier detection methods follow either a supervised or a unsupervised approach [7]. Supervised approaches are learning methods that construct the model from training data [8], [9]. Arbitrary data is termed outlier if it does not fit the model. Unsupervised approaches on the other hand are unaided methods that usually discriminate a data as an outlier based on distance to its neighbors [10], [11]. The approaches vary in the manner on how the distance is calculated. Multivariate unsupervised outlier identification methods like linear regressions models for outlier detection, capitalize on known relationships among variables to calculate the error in predicting the dependent variable [3], [4], [6], [12]. The distance is then the deviation of error from the population mean. The problem with linear regressions models for outlier identification is that they do not generally account for the manufacturing variability and test measurement shifts, commonly occurring in analogue and RF devices. Since the test measurements have less predictability, the population mean of the measurements also has lesser certainty. This effects the distance metric and hence the marginal outliers have a higher chance of being undetected. Another approach is to transform the test measurements to a principal component space and then apply the outlier identification methods on the transformed data [13], [14]. One of the underlying problems in principal component (*PC*) related methods is their high sensitivity to scaling. Since test-response measurements exhibit spatial and temporal shifts, the distance metric is non-stationary and hence affects the results of the PC based outlier identification methods.

To circumvent the problems associated to calculating a robust distance metric, the Mahalanobis distance [15] is employed as the distance parameter in the unsupervised outlier identification methods [16]. Mahalanobis distance is invariant to measurement shifts and most importantly accounts for the relationships among the data variables (covariance) in calculating the metric. However, since the Mahalanobis distance metric is sensitive to outlier data, robustly estimated population mean and covariance are used in identifying outliers [17]–[19]. To further reduce the variance in the distribution of

1530-1877/11 $26.00 © 2011 IEEE
DOI 10.1109/ETS.2011.31

the Mahalanobis distance metric, this paper proposes a novel way of using multiple sets of Mahalanobis distance metric for reliability analysis of analogue electronic products. One of the main advantages of the proposed technique is the capability of the model to include sets of correlating and uncorrelating test variables into a single structure and conduct a qualitative and quantitative reliability analysis of a functionally qualified product.

II. Multivariate Reliability Classifier Function and Reliability Analysis Flow

The reliability classifier function to be formulated and used as a model for reliability analysis is multivariate in nature. The associated analysis is called multivariate analysis. Multivariate analysis comprises a set of techniques dedicated to the analysis of data sets with more than one variable [20]. For electronic circuits, the responses from several tests can be combined in a multivariate sense to describe the response behavior of the system in terms of product reliability. The following subsections will firstly discuss a procedural way of formulating the multivariate classifier reliability function as a model for qualitative analysis of product reliability. The qualitative analysis is performed to remove the outliers to the multivariate reliability classifier functional model. This is followed by a description on a way of utilizing the model to assign the probabilistic value for reliability. Finally the overall reliability analysis flow for simultaneously qualifying and quantifying product reliability will be discussed.

A. Multivariate Reliability Classifier Function

The multivariate reliability classifier function as a model for reliability analysis is constructed from multiple test-sets. All tests within a set are uncorrelated, while the test-sets are correlated among each other. Each test-set contains a minimal number of tests and hence corresponds to the most significant tests. The significance of these tests is based on their ability to construct a reliability classifier model with parameters sensitive to changes that influence the final results. Once such a selection of tests is made, a suitable metric is developed that is invariant to measurement shifts. Invariant to measurement shifts prevents recomputing the parameters of the model for every instance of change in the measurement environment. In addition, every classifier model attempts to avoid as many number of false positives or the *type I* error and false negatives or *type II* error that are possible errors occurring in a statistical decision process. One of the ways to lessen this erroneous decision process is to reduce the variance of distribution. Reducing the variance in distribution reduces the overlap if any, among the distributions and hence the false positives and false negatives. In the following sections, the descriptions of various steps that ultimately leads to a reliability classifier function is discussed.

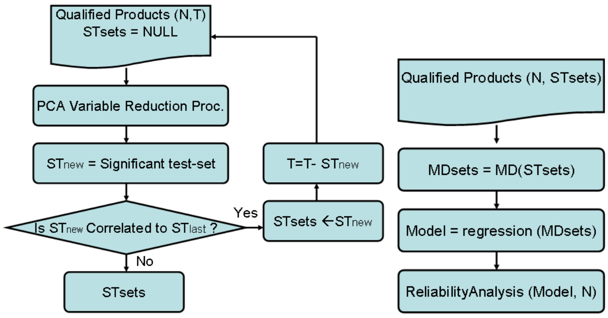
1) PCA Variable Reduction Procedure: The Principal Component Analysis (*PCA*) variable reduction procedure is an approach to select the most significant variables that are sufficient to describe the observed characteristic behavior of the system. Usually in an electronic circuit test environment, only n out of p number of test measurements ($n << p$) are significant and sufficient in the multivariate sense, to describe a certain response behavior of the device. An efficient and well-known approach to determine those n number of significant tests, is to subject the p-tuple measurement set to the *PCA*, wherein the first *m* principal components are determined using *Scree* plot [21].

2) Mahalanobis Distance Metric: The Mahalanobis distance is a well-known distance measure that takes into account the covariance matrix [15]. For a p-dimensional multivariate sample $x_i (i = 1, 2, \ldots, k)$ the Mahalanobis distance is given as in Equation 1. This distance metric differs from the Euclidean distance in that the amount of correlations in the data set is taken into consideration. One of the most striking features of this metric is that it is scale-invariant.

$$\mathrm{MD_i} = ((x_\mathrm{i} - \mathrm{t})^\mathrm{T} \mathrm{C}^{-1} (x_\mathrm{i} - \mathrm{t}))^{1/2} \tag{1}$$

where t is the estimated multivariate location and C the estimated covariance matrix. Usually, t is the multivariate arithmetic mean or centroid, and C is the sample covariance matrix. For multivariate normally distributed data, the squared values ($\mathrm{MD_i^2}$) are approximately chi-square (χ^2) distributed with p degrees of freedom [18].

3) Construction of PCA-MD based Reliability Multivariate Classifier Function: The constructional flow of the Principal Component-Mahalanobis Distance (*PCA-MD*) based multivariate reliability classifier function is depicted in Figure 1. This functional model to be constructed and capable of qualitatively analyzing the reliability of products is parameterized by k number of significant test-sets, $STsets$. While the significant test-sets are correlated among themselves, the tests within each significant test-set are uncorrelated. Figure 1a depicts the flow for determining significant test-sets. The PCA variable reduction procedure is used iteratively to build k number of significant test-sets. As shown in Figure 1a, the procedure begins with the functional test list T and corresponding response measurements from N number of functionally qualified products. At each PCA variable reduction iteration, all previously determined tests and corresponding measurements are removed from the overall test list and measurements, and a new set of tests is determined. The iteration stops at k, if there is no sufficient correlation (less that 0.7, for instance) among the respective test pairs of the significant test-sets. In the following step, as shown in Figure 1b, the Mahalanobis distance metric $MD_l \in MDsets,\ l = 1, 2, \ldots, k$, one for every significant test-set, is computed for all the products. Since any two significant test-sets are correlated (respective test pair correlation), the corresponding Mahalanobis distance vectors (vector defined from the centroid of the measurement to the specific measurement) of the products are collinear and since the Mahalanobis distance vectors are collinear, one of them can be expressed as a regression function of the others as shown in the Equation 2. The regression fitting coefficients can be estimated from the Mahalanobis distance metric calculated for all products. The regression function constructed in this way, can be considered as the multivariate reliability classifier function. This functional model utilizes the information on distribution of error ϵ, in the regression equation to reduce the variance in measurement [22] and qualitatively classify the reliability of the product.



(a) Significant test sets Determination (b) Reliability Analysis

Fig. 1. PCA-MD based Multivariate Reliability Classifier Flow

$$\mathrm{MD_{i,l}} = \sum_{\mathrm{m=1,m\neq l}}^{\mathrm{k}} \mathrm{a_m MD_{i,m}} + \epsilon \tag{2}$$

B. Quantification

Quantification is the process of assigning a reliability metric to the qualitative results of the multivariate reliability classifier. The process of quantification begins once the outliers are qualitatively identified and removed from the test response data. Hence the goal of quantification is to assign a probabilistic value that quantifies the reliability of all functionally qualified products that has not been identified as an outlier in the qualitatively analysis.

As stated earlier, the multivariate reliability classifier functional model is parameterized by sets of Mahalanobis distance metrics that are collinear. The distribution of errors in the regression fit is qualitatively analyzed to identify the outliers. Rewriting the multivariate reliability classifier functional model equation from Equation 2 to Equation 3, converts the error ϵ, from normal distribution to a chi-square (χ^2) distribution with one degree of freedom [23].

$$\epsilon^2 = (\mathrm{MD_{i,l}} - \sum_{\mathrm{m=1,m\neq l}}^{\mathrm{k}} \mathrm{a_m MD_{i,m}})^2 \tag{3}$$

The error ϵ, is usually not normally distributed due to small mismatches in measurements. These mismatches, setting aside the measurement variability that are conditioned by the model, may indicate the reliability risk to the functionally qualified product. Since the errors have slightly skewed distributions, the square of these deviations are much more expressed in the chi-square (χ^2) distribution. The deviation of the squared error distribution from a Chi-square (χ^2) distribution is quantified to assign probabilistic values that reflect the reliability of the qualified product within the scope of the available test-response measurements.

A well-known technique to test the goodness-of-fit of a given set of observations to any hypothesized distribution is the *Chi-square*(χ^2) test [24]. Any discrepancy between the observed and the expected values is expressed in terms of a χ^2 statistic as shown in the Equation 4, where d and e are the observed and expected values, while k is the number of independent variables or the degrees of freedom. From the computed χ^2 value for a set of observations, the probability that these set of observations belong to the hypothesized distribution can be determined from a χ^2 table [25].

$$\chi^2 = \sum_k (d_k - e_k)^2 / e_k \tag{4}$$

For the squared error distribution to be tested for the goodness-of-fit to the χ^2 distribution, the sample squared error distribution space is divided into an equal number of classes. From the number of observations for each class and the expected number of observations determined from the quantiles of the χ^2 distribution, the χ^2 is computed for a predefined ($n > 1$) number of classes, depending on the granularity requirements of the results. From the evaluated χ^2 statistic for the set of classes and the qualified products belonging to that class, the probability of these products belonging to the χ^2 distribution and hence the measure of reliability, is determined.

III. Experiment and Results

To demonstrate the capability of the multivariate reliability classifier to qualify and quantify the reliability of electronic circuits, functionally qualified samples from an industrial product have been chosen. In the following section, a brief description of the product, chosen samples for reliability analysis will be first described, followed by the multivariate reliability classifier analysis flow and discussion of results.

A. Product and Sample Information

The product is a single IC implementing a car radio tuner for AM and FM intended for microcontroller tuning with the I2C-bus. The number of functional tests at wafer level is 175 that corresponds to several levels of DC tests, modes of AC and RF tests. Those products that pass these functional tests are referred to as *qualified* products at wafer level. The samples chosen for demonstrating the multivariate reliability classifier are all qualified car radio tuner products instances chosen from a single wafer picked up from the manufacturing lot. The number of qualified products or known good dies (KGDs) are 379 dies, while the number of known defective dies (KDDs) are 99 dies.

B. Multivariate Reliability Classifier Flow Results and Discussions

The following sub-sections will first describe this PCA-MD based reliability classifier function construction procedure in detail, followed by the description of the qualitative and quantitative reliability analysis of the samples and discussion of the results.

1) PCA Variable Reduction: Functional test-data from 175 tests and 379 qualified products were iteratively subjected to the PCA variable reduction procedure. During each iteration, a set of significant tests that belong to the significant test-set are selected based on the *Scree* plot. Figure 2 shows the Scree plot of the test data for two iterations. The number of significant tests selected for each significant test-set *T1* and *T2*, are 9 tests. These test variables were selected based on the coefficients of the respective principle components lying above the elbow in the Scree plot. Since the test variables within a test-set are chosen from different principle components (first 9 PC's explaining the variance in data), the variables selected from these principle components are uncorrelated or heterogeneous. Again, since the test variables in the next iteration (2^{nd}) are selected from the same set of 9 principle components (given the small margin of change in the associated Eigen vector after removal of the test variable selected in the previous iteration), the significant test-sets are equivalent. Combining these two arguments it can be concluded that the test variables within the test-sets *T1* and *T2* are uncorrelated or heterogeneous, while the test-sets are equivalent or pair-wise correlated.

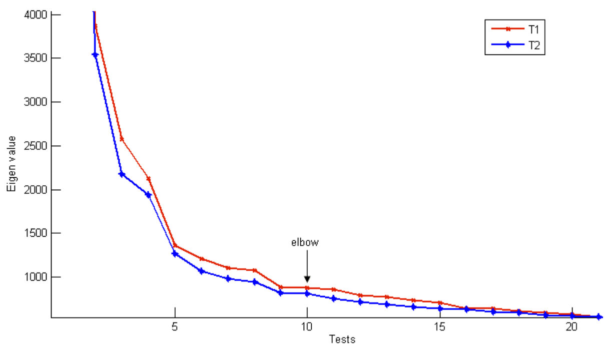


Fig. 2. Scree Plot of Test-data from Qualified Products

2) PCA-MD based reliability classifier: The *PCA-MD* based reliability classifier is constructed from the sets of significant test-sets that are correlated among the test-sets on hand and uncorrelated within the test-sets. For each significant test-set *T1* and *T2*, a Mahalanobis distance metric *MD1* and *MD2*, is calculated for all products, using the test response data corresponding to the tests in respective significant test set. Since the respective Mahalanobis distance metric for all products are collinear to each other, one of them can be linearly expressed as the function of the other. Hence a linear regression model as shown in Equation 5, is fit with the *MD1* as the independent variable and *MD2* as the dependent variable. This linear regression model is the *PCA-MD* based reliability classifier function. The error in the regression fit of the reliability classifier function is further analyzed to qualitatively and quantitatively classify the reliability of the *qualified* products.

$$\mathrm{MD2} = 0.97 \times \mathrm{MD1} + 0.04267 \tag{5}$$

3) Qualitative Analysis Results: The goal of qualitative reliability analysis is to determine statistical outliers to the model. Statistical outliers to a model can be determined by iteratively constructing the model and estimating the goodness-of-fit after every iteration. However, during each iteration, a set of data that do not seem to fit the model are removed and the model parameters are re-estimated. The set of data that are removed at each iteration until no further improvement is observed in the fit, are the outliers to the model.

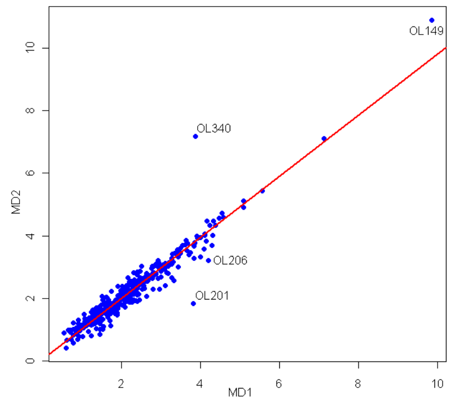


Fig. 3. Regression Plot (MD1,MD2) and 4 Outliers

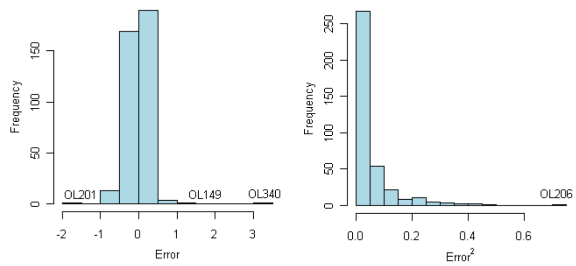


Fig. 4. Error Distribution and Squared Error Distribution of *Qualified* Products and Outliers

Following this criterion for the reliability classifier functional model described in Equation 5, the empirical distribution errors are utilized to qualitative determine the outliers of the dataset. Since the linear regression fit follows a least square fit, the error distribution of the fit will closely follow a normal distribution and the square of the error distribution will follow a χ_2 distribution. The distribution of the errors for all *qualified* products along with 4 outliers, are depicted in Figure 3, while Figure 4 shows the distribution of the errors and squared errors along with the outliers to the respective distributions. The statistical outliers determined in this procedure are qualitatively classified as products that are potentially unreliable.

4) Invariance to Response Measurement Shifts: One of the major advantages of the reliability classifier model is its invariance to response measurement shifts. This is because the model employs the Mahalanobis distance as parameter which is robust and invariant to scaling [15]. In addition to this, our model utilises the error distribution of multiple Mahalanobis distances that are collinear to each other. This further reduces the naturally occurring variance in test measurements. Altogether, the reliability classifier model is robust to measurement variations and shifts.

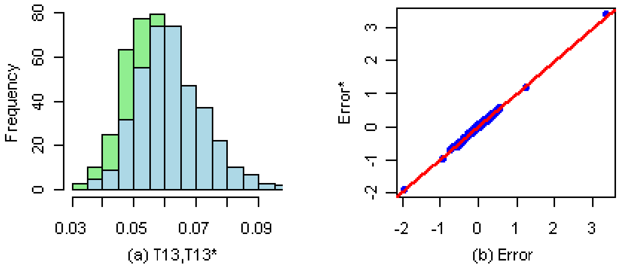


Fig. 5. Histogram T13 and T13*, Scatter-plot of Error and Error*

In order to validate the invariance of the reliability classifier model to measurement shifts, the test data of all the qualified products were shifted (increased by 10% to its measured value) and subjected to the reliability classifier for qualitative analysis. Figure 5a, shows the overlapped histogram of the measured (T13) and shifted test data (T13*) related to test variable $T13$ from all qualified products. Figure 5b is the regression plot of the error distribution Error and Error*, of the reliability classifier function model contructed before and after shift of the test-response data. The regression line lying at 45° and passing through the origin, indicates that the error distributions are equal to one another, asserting the invariance of reliability classifier model to measurement shifts.

5) Comparative Results to Other Outlier Detection Methods: Table I shows the comparative results of the PC-MD based reliability classifier (PC-MD-RC) model to other multivariate outlier detection models like the linear regression (LR), Principal Components (PC) and a univariate method like PAT. The labels denote the outliers identified by each of the methods. Although the results from other outlier methods (LR, and PC), put together, produced comparable results, they do not identify the outliers exclusively. This is in contrast to our PC-MD-RC model.

TABLE I
COMPARATIVE RESULTS

PC-MD-RC	LR	PC	PAT
OL201	OL340	OL340	-
OL340	OL149	OL201	
OL149	-	-	
OL206			

6) Quantitative Analysis Results: Figure 6 shows the density function of squared errors derived from the reliability classifier functional model for all qualified products, after removing the outliers identified in the qualitative reliability analysis procedure discussed earlier. A corresponding χ^2 density function with 1 degree of freedom is overlayed to display the deviation from the density function of squared errors.

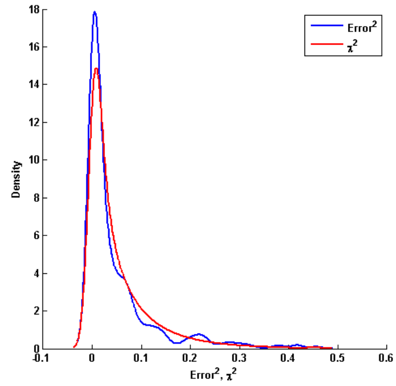


Fig. 6. Overlayed Density Function of Squared Errors and χ^2

To determine the goodness-of-fit of the squared error with respect to the χ^2 distribution, a χ^2 test is conducted. To facilitate the χ^2 test, the distributions are evenly sub-divided into classes. For each classes lying between the 2^{nd} and 4^{th} quartile of the χ^2 distribution, the number of observed and the expected number of samples are determined from the squared error distribution and the χ^2 distribution respectively. The d^2/e statistic is then computed for each of these classes. In order to accommodate the variations between adjacent classes, two neighboring classes along with the current class are subjected to χ^2 test. In other words, to determine the goodness-of-fit of a class from the squared error distribution, the χ^2 statistic for that class includes the two of its neigbouring classes. The exception is however for the first and the last class, where only a neighbor class is included. Hence the degree of freedom for all classes is 3, while for the first and the last class is 2. Whenever the expected class observation is less than 2, the χ^2 statistic is not computed as they will not provide any significant results. This is because an expected value of one observation is usually at the tail of the χ^2 distribution and any major deviation of the squared error distribution will be already identified and removed by the qualitative reliability analysis procedure.

Table II shows the quantitative analysis results for all *qualified* products other than the outliers identified in the qualitative analysis. For each class with a mid-value $Cls.$, beginning from the 2^{nd} quartile, the observed $Obs.$, and the expected $Exp.$ number of observation are shown. The corresponding d^2/e values and the $\chi^2_{df=3}$ statistic are calculated. The probability value $pvalue$, is determined for $\chi^2_{df=3}$, with degrees of freedom df, 2 for the first and the last class and 3 for the remaining classes from the χ^2 statistical table [25]. A probability value greater than 0.05 indicates that the deviation

TABLE II
QUANTITATIVE ANALYSIS RESULTS FOR THE QUALIFIED PRODUCTS

Cls.	Obs.	Exp.	$(d-e)^2/e$	$\chi^2_{df=3}$	p value
0.11	10	14	1.14	1.24	0.35
0.13	9	10	0.10	1.74	0.4
0.15	6	8	0.50	4.76	0.1
0.17	1	6	4.16	4.66	0.1
0.19	4	4	0.00	5.16	0.05
0.21	6	4	1.00	2.33	0.3
0.23	5	3	1.33	4.33	0.1
0.25	0	2	2.00	3.83	0.25
0.27	3	2	0.50	3.50	0.20
0.29	2	1	1.00	-	-
0.31	2	1	1.00	-	-
0.33	1	1	0.00	-	-
0.35	0	1	1.00	-	-
0.37	1	0	*	-	-
0.39	1	1	0.00	-	-
0.41	2	0	*	-	-
0.43	0	0	*	-	-
0.45	1	1	0.00	-	-

of the observed value of the squared error from the expected χ^2 value is sufficiently small that chance alone accounts for it, while a probability value less than or equal to 0.05 means that some factor other than chance is operating for the deviation to be significantly large. Hence a p value equal to 0.05 for the qualified products belonging to the class interval with a mid-value 0.19 indicates that there is only 5% chance that this product, belong to the χ^2 distribution. Although the observed and the expected value for this class match, the χ^2 statistic of a class is influenced by its two adjacent neigbours. Hence the chance that the products from the neighbouring classes can be hypothesed to be relatively less affected. Since the p value for the remainder of the classes are greater than 0.05, it can be concluded that the deviation of the squared error is due to chance and can be probablistically considered reliable within the realm of the functional test conducted on the product.

IV. CONCLUSIONS

A multivariate reliability classifier model capable of qualitatively and quantitatively analyze the reliability of analogue electronic products has been described. One of the major advantages of the model is that it simultaneously accommodates the strengths of using correlating and uncorrelated test variables for identifying outliers. To avoid any bias to the results caused by commonly occurring shifts in test-response measurements, robustly estimated distance metric has be used by the model to identify marginal outliers. The functional model of the reliability classifier has been chosen in a way that the results of the qualitative analysis can be further analyzed quantitatively to measure the level of reliability of the product in a probabilistic sense. The combined other models like the linear regression model and the principal component model do not identify the outliers exclusively, in contrast to our multivariate reliability classifier model. Other applications of the reliability classifier like identification and conditioning specific functional test for product grading and early detection of failures remains to be investigated.

REFERENCES

[1] “Zero defects guideline,” Automotive Electronics Council, http://www.aecouncil.com/, Aug 2006.

[2] W. Riordan, R. Miller, and E. S. Pierre, “Reliability improvement and burn in optimization through the use of die level predictive modeling,” in *Proc. IEEE International Reliability Physics Symposium*, 2005, pp. 435–445.

[3] W. R. Daasch, J. McNames, R. Madge, and K. Cota, “Neighborhood selection for iddq outlier screening at wafer sort,” *IEEE Design and Test of Computers*, vol. 19, no. no. 5, pp. 74 – 81, Sep./Oct. 2002.

[4] T. J. Powell, J. Pair, M. S. John, and D. Counce, “Delta iddq for testing reliability,” in *Proc. IEEE VLSI Test Symposium*, 2000, p. 439.

[5] J. P. de Gyvez, G. Gronthoud, and R. Amine, “Vdd ramp testing for rf circuits,” in *Proc. IEEE International Test Conference (ITC)*, 2003, pp. I651 – 658.

[6] L. Fang, M. Lemnawar, and Y. Xing, “Cost effective outliers screening with moving limits and correlation testing for analogue ics,” in *Proc. IEEE International Test Conference (ITC)*, 2006.

[7] F. Angiulli, S. Basta, and C. Pizzuti, “Distance-based detection and prediction of outliers,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 18, no. 2, pp. 145–160, 2006.

[8] M. Khan and S. Akter, “Multiple-case outlier detection in least-squares regression model using quantum-inspired evolutionary algorithm,” in *Proc. IEEE International Conference on Computer and Information Technology (ICCIT)*, 2009, pp. 7–12.

[9] D. Janakiram, A. M. Reddy, and A. V. U. P. Kumar, “Outlier detection in wireless sensor networks using bayesian belief networks,” in *IEEE First International Conference on Communication System Software and Middleware (Comsware)*, 2006, pp. 1–6.

[10] S. Bay and M. Schwabacher, “Mining distance-based outliers in near linear time with randomization and a simple pruning rule,” in *Proc.IEEE International Conference Knowledge Discovery and Data Mining (KDD03)*, 2003.

[11] S. Ramaswamy, R. Rastogi, and K. Shim, “Efficient algorithms for mining outliers from large data sets,” in *Proc.IEEE International Conference Managment of Data (SIGMOD)*, 2000, pp. 427–438.

[12] S. Krishnan, R. Jonker, and L. van de Logt, “Variance reduction for supply ramp based cheap rf test alternatives,” in *IEEE European Test Symposium*, 2007, pp. 55 – 62.

[13] R. Gnanadesikan and J. R. Kettenring, “Robust estimates, residuals, and outlier detection with multiresponse data,” *Biometrics*, vol. 28, pp. 81–124, 1972.

[14] P. O’Neill, “Production multivariate outlier detection using principal components,” in *Proc. IEEE International Test Conference (ITC)*, 2008, pp. 1 – 10.

[15] P. C. Mahalanobis, “On the generalised distance in statistics,” in *Proc. of the National Institute of Sciences of India*, vol. 2, no. 1, 1936, pp. 49 – 55.

[16] T. E. Smetek and K. W. Bauer, “Finding hyperspectral anomalies using multivariate outlier detection,” in *Aerospace Conference*, 2007, pp. 1–24.

[17] P. J. Rousseeuw and K. van Driessen, “A fast algorithm for the minimum covariance determinant estimator,” *Technometrics*, vol. 41, pp. 212–223, 1999.

[18] P. Filzmoser, “Identification of multivariate outliers,” *Austrian Journal of Statistics*, vol. 34, no. 2, pp. 127–138, 2005.

[19] R. Turakhia, W. Daasch, J. Lurkins, and B. Benware, “Changing test and data modeling requirements for screening latent defects as statistical outliers,” in *IEEE, Design & Test of Computers*, vol. 23, no. 2, 2006, pp. 100 – 109.

[20] A. C. Rencher, *Methods of Multivariate Analysis*, 2nd ed. John Wiley & Sons, Inc., 2002.

[21] T. Jollifee, *Principal Component Analysis*. Springer-Verlag, New York, 2002.

[22] A. M. Law and W. D. Kelton, *Simulation Modeling and Analysis*, 2, Ed. McGraw-Hill, Inc., 1991.

[23] M. Evans, N. Hastings, and B. B. Peacock, *Statistical Distributions, 3rd ed. New York: Wiley, p. 57, 2000.*, 3, Ed. New York: Wiley, 2000.

[24] H. Chernoff and E. L. Lehmann, “The use of maximum likelihood estimates in 2 tests for goodness-of-fit”. the annals of mathematical statistics,” *The Annals of Mathematical Statistics*, vol. 25, pp. 579–586, 1954.

[25] R. Fisher and F. Yates, *Statistical Tables for Biological Agricultural and Medical Research*, 6, Ed. Oliver & Boyd, Ltd., Edinburgh.