THESIS

AN OUTLIER DETECTION APPROACH FOR PCB TESTING BASED ON PRINCIPAL COMPONENT ANALYSIS

Submitted by

Xin He

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2011

Master's Committee:

Advisor: Anura P. Jayasumana

Yashwant K. Malaiya
Steven C. Reising

ABSTRACT

AN OUTLIER DETECTION APPROACH FOR PCB TESTING

BASED ON PRINCIPAL COMPONENT ANALYSIS

Capacitive Lead Frame Testing, a widely used approach for printed circuit board testing, is very effective for open solder detection. The approach, however, is affected by mechanical variations during testing and by tolerances of electrical parameters of components, making it difficult to use threshold based techniques for defect detection. A novel approach is presented in this thesis for identifying boardruns that are likely to be outliers. Based on Principal Components Analysis (PCA), this approach treats the set of capacitance measurements of individual connectors or sockets in a holistic manner to overcome the measurement and component parameter variations inherent in test data.

Effectiveness of the method is evaluated using measurements on different types of boards. Based on multiple analyses of different measurement datasets, the most suitable statistics for outlier detection and relative parameter values are also identified.

Enhancements to the PCA-based technique using the concept of test-pin windows are presented to increase the resolution of the analysis. When applied to one test window at a time, PCA is able to detect the physical position of potential defects. Combining the basic and enhanced techniques, the effectiveness of outlier detection is improved.

The PCA based approach is extended to detect and compensate for systematic variation of measurement data caused by tilt or shift of the sense plate. This scheme

promises to enhance the accuracy of outlier detection when measurements are from different fixtures. Compensation approaches are introduced to correct the 'abnormal' measurements due to sense-plate variations to a 'normal' and consistent baseline. The effectiveness of this approach in the presence of the two common forms of mechanical variations is illustrated. Potential to use PCA based analysis to estimate the relative amount of tilt and shift in sense plate is demonstrated.

ACKNOWLEDGEMENTS

After finishing the thesis, it is time to thank all the people who have helped me so much in the past, which makes this thesis possible and enjoyable for me.

First of all, I would like to give my deepest appreciation to my research advisor, Professor Jayasumana, for providing me such a great opportunity to work with him. I couldn't thank him enough for his knowledgeable guidance, valuable suggestions, and continuous encouragement throughout my Master study. His tremendous help is always around whenever I need. His vision and knowledge have not only stimulated my passion on research but also taught me many ways in my personal life. I am deeply grateful to the efforts, patience, and enthusiasm that he shared with me along this entire process.

I would also like to thank Professor Malaiya who devoted lots of time into my research. His help is complementary to what Dr. Jayasumana did. From the communication with him, I benefited from both his intelligence and his humor, which will be important to my future life.

Thanks Professor Reising for being on my thesis committee and for their valuable time spent on reading and commenting on my thesis. His passion on research and language produced extremely positive influence on my study and work.

I would like to acknowledge Dr. Kenneth Parker for providing me the opportunity to access Agilent dataset and test background. Thanks to Mr. Steve Hird, Mr. Kent Carey of Agilent, and the Agilent University Relations program for their generous support.

I have also benefited greatly from the interaction and collaboration with all the members of the CNRL group at Colorado State University. Their intelligence, dedication, and fantastic personalities make it a learning experience for me everyday. Thanks for their helpful discussion and encouragement when I felt frustrated.

Last but not least, I am so grateful to my parents who helped me develop the interest in science during my childhood. Thank for their love, encouragement, and support during my whole life, especially when I am thousands of miles away from them in the last few years. Finally, I would like to thank my wife, who bought me endless happiness and took care of my life in every single aspect whole heartedly.

TABLE OF CONTENTS

Chapter 1. Introduction

In board manufacture, the defects that occur include: open solder joints; insufficient, excess or malformed solder joints; lost devices; shorts; excess solder; dead devices; incorrect device placements; polarized devices wrongly placed; and misaligned parts [1].

The Capacitive Lead Frame Testing technique (known as TestJet® or its enhanced version VTEP® in industry) is an effective method used in printed circuit board (PCB) testing [1]. With this technique, open solder defects can be detected without having to power the board under test by measuring the capacitance between a pin and a tester sense plate.

The TestJet technique tests for open pins in connectors and sockets on boards, using the capacitance formed between a device pin and a suspended sensing plate [1]. During TestJet test, the pin under test is connected to an AC signal source while all other pins are connected to ground. An open defect on the tested pin changes the measured capacitance to an abnormal one. In TestJet test results, a normal test reading means that there is no defect, or at least open solder defect, and a low test reading is an indication that the tested signal pin itself is open [2][3]. However, in large-scale manufacturing, parameter variations from component to component and board to board affect the lead capacitance values. Furthermore, the variation from test fixture to test fixture, and the variation from test system to test system affect the measurements and their accuracy. A good board tested in one environment eg, fixture and temperature may appear to be a 'bad' one in another test environment. Sometimes a test needs to set a new limit when executed in a new test environment. With evolving technologies, and consequently, the increasing

densities of components and boards, the margins available for deciding among faulty and fault-free devices are shrinking. That limits the ability of TestJet method to detect defects.

Threshold Setting is used currently with capacitive testing to differentiate the normal values from the abnormal capacitance values. However, nowadays, designers are improving board functionality without increasing the board size. More circuits and the related pins on a similar sized board are becoming smaller and smaller. Higher signal speed is also a driven factor for the smaller component on a board. The shrinking component size on PCB results in lower coupling capacitance between signal pins and the sense plate making the measurements decrease a lot. Then it is more challengeable to set a threshold value to screen defective devices. For example, original good measurements 50-60 fF could hold much more room for board to board variation than 8-10 fF measurements.

Relative thresholds based on standard deviation are also made ineffective by these factors. Non-optimal threshold settings can result in higher false fails or false passes. The challenge is further compounded due to the fact that each pin tested has a threshold that is different from others, yet often correlated to them. Furthermore, mechanical parameters such as spacing between the plate and the connector/device vary from one mounting on the tester to another. Similarly, the capacitance value corresponding to a pin may vary from board to board due to the fact that components are from different vendors and different batches. These and other factors combine to make the selection of appropriate thresholds a challenging task.

A printed circuit board, also referred to as a board, is a unique assembly that contains many devices. We use test data from Agilent boards with connectors for DDR2 RAM to evaluate our scheme. A device (such as a connector or socket) is a unique item with a collection of numbered pins, which are subject to testing with the exception of VDD and ground pins. Multiple devices with unique names may exist on a single board. For example, j3 and j24 are standard connectors used in DDR2 RAM boards.

In this thesis we present a novel method for PCB testing, based on Principal Components Analysis (PCA), to improve the efficiency and decrease the potential false classification rate of the TestJet technique. It changes the testing paradigm from one that compares values against fixed thresholds to one that detects outliers. Thus, if the majority of the boards are fault-free, the outliers, which by definition are significantly different from the rest of the boards, are likely to be abnormal or even defective. The method relies on an ensemble of measurements, allowing it to identify correlations among pin capacitances. Thus it can adapt to board-to-board, device-to-device, fixture-to-fixture, and test system to test system variations more effectively than traditional techniques.

PCA based outlier detection has been investigated and found effective for testing of ICs in [4][5][6][7]. It is a successful statistical test technique for the detection of faulty ICs whenever analog test measurements are involved, e.g., IDDQ, delay, power etc. With PCBs however, the defect characteristics of faults and their manifestation in measured values are significantly different from those with ICs. For example, in PCBs, the effects of defects are more localized, and the tests are able to capture spatial distributions. These spatial distributions indicate the recognition of correlations among measurements of

adjacent pins. The concept of test pattern associated with ICs is not applicable for PCB capacitance measurements. Furthermore, the measured values can vary over a wide range from pin to pin in good boards and connectors.

This thesis presents and evaluates a PCA based outlier detection scheme for PCBs, where the set of measurements per device or a connector is used in a holistic manner to detect the outliers. An extension of that method is presented in which the analysis is carried out separately for small subsets (windows) of pins. Latter scheme, the localized method, exhibits better sensitivity for connector testing due to the fact that the effects of an open pin are likely to be limited to a small set of neighboring pins. Furthermore, it makes the identification of the specific pin affected easier as the abnormal pin is localized to within the window size. When good boards are tested under different fixtures which are subject to some mechanical misalignment of the sense plate, the measures values may vary depending on the degree of misalignment. This thesis considers approaches to compensate for the measurement differences for different fixtures that will avoid incorrect detection of outliers.

As an Intern at 'Cadeka Microcircuits, Loveland, CO', we were able to familiarize with an actual industry test process. The system at Cadeka is used for chip test with ATE (automatic test equipment) and test code. For each die on the wafer, several different measurements are applied. Since upper/lower limits are set according to Product Requirement Sheet, the die can be signaled as Fail or Pass from the comparison with these limits. Average value, standard deviation, CP (process potential index) and CPK (process capability index) are automatically calculated by the program which could be

used for further analysis. Each die measurement also contains location (x, y) information. Then, a plot of wafer with Pass/Fail dies could be created by C language. Based on the plot, areas that most problems happened will be observed. Normally, the dies on the out edge of wafer have high possibility of failure. The research presented in this thesis addresses the scenarios that will be encountered as the component and manufacturing tolerances shrink due to miniaturization, at which time upper/lower limits will be hard to enforce.

Chapter 2 provides a background to capacitive lead frame testing and compares it with other techniques. Chapter 3 outlines the PCB based outlier detection scheme. The TestJet measurement data used for evaluation of the effectiveness of the scheme is described in Chapter 4. Chapter 5 evaluates the proposed scheme in detail. Then in Chapter 6 we evaluate the proposed technique with many different test data. The comparison is made in Chapter 7 between our PCA method and traditional threshold-based approach, where the threshold is set as a multiple of standard deviation of measurements. In Chapter 8, we also present and evaluate a modification to the basic strategy, to enhance its sensitivity. Window size selection is discussed in later part of the Chapter. We also investigate, in Chapter 9, the effects of common mechanical variations in sense plate on the PCA based outlier detection algorithm, to obtain new methods compensating the measurements variation. Chapter 10 discusses the measurement distribution from single board and multiple boards. Conclusions and future research are presented in Chapter 11 of this thesis.

Chapter 2.
Capacitive Frame Lead Testing

2.1. Background

The fault spectrum of PCBs changed a lot with **Surface-Mount Technology (**SMT) manufacturing, which caused the open solder defects to become the top problem for numerous manufacturers. In fact, the development of SMT polarized capacitors and SMT connectors have made visually checking the correct orientation and connection difficult or impossible to do [8]. The undetected defects always lead to un-repairable damage to the device after several hours usage by consumer.

Missing component, wrong component, mis-oriented, dead component, wrong device alignment, short between pins , solder open… are possible board defects. Some of them could be tested by the powered test method and un-powered test method which will be discussed in detail in Chapter 4. For the open-solder defect, Capacitive Lead Frame Testing is now an effective method

As an effective method to detect open-solder defect, Capacitive Lead Frame Testing (known as TestJet® or its enhanced version VTEP® in industry) was researched and developed from mid-1980's. The technique was suggested by the parasitic diode detection and parasitic transistor detection techniques, a parasitic capacitor can formed on the lead frames of integrated circuits, connectors, capacitors and some switches. The kind of reliable parasitic capacitance could be predicted well. Then If there is no variability from measurement system and topology, the parasitic capacitance makes the test technique effective in high volume manufacture [8].

2.2. Principles

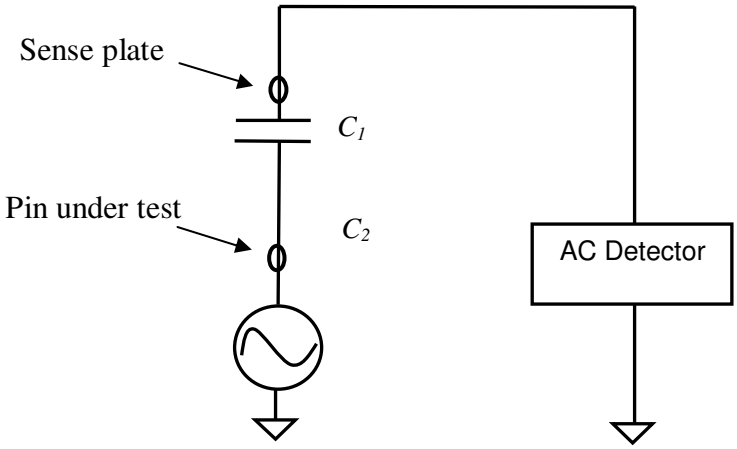


Figure 2- 1. Simple equivalent circuit of connector without open defect [1]

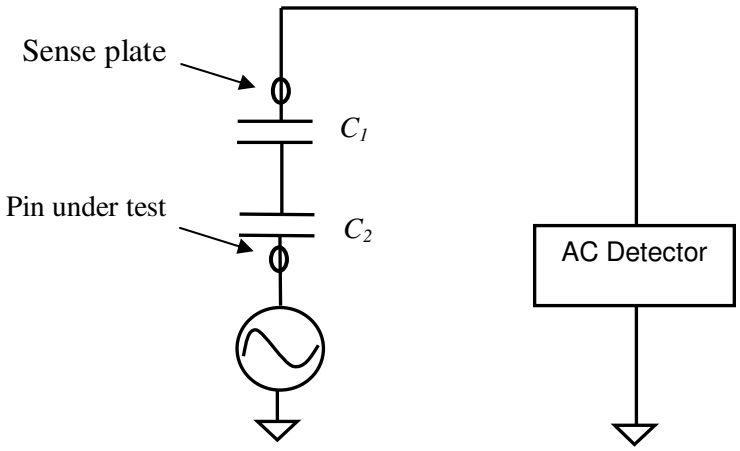


Figure 2- 2 Simple equivalent circuit of connector with open defect [1]

Discriminating measurement of good connection from that of an open solder connection is the main principle of the Capacitive Lead Frame Testing. Figure 2-1 shows simple equivalent circuit of the connector without open defect. In contrast, when an open defect exists, as shown in Figure 2-2, another capacitance C_2 will exist in series with C_1 in

the circuit. The equivalent circuits above ignore the parasitic inductor, resistor and mutual capacitor.

When an open defect exists, the new overall capacitance $C_{open} = \frac{C_1 * C_2}{(C_1 + C_2)}$ which is normally 2 to 10 factors smaller than the original C_1 can cause a detectable difference. The difference can be used to tell the open solder in the ICs or PCBs without knowing what the devices actually does [1].

Capacitive Lead Frame Testing uses fixture implementation shown in Figure 2-3 to form a detectable capacitance from the signal measurement device to the devices under test. The Capacitive Leadframe testing measures the capacitance between the test pins and a sense plate to identify defects. When an AC signal stimulates the tested pin, the sense plate suspended over the connector will transfer a capacitively coupled signal into a buffer and then to the tester, where the signal is converted to a measure of capacitance. The measured capacitance may be fairly small, often less than 100 femtofarads (fF). If there is an open solder defect existing between the board and the connector pin, the capacitance detected by the tester often decreases significantly, to perhaps 10 fF or even less. Variations in these measurements must be accounted for to avoid false pass/fail indications [2].

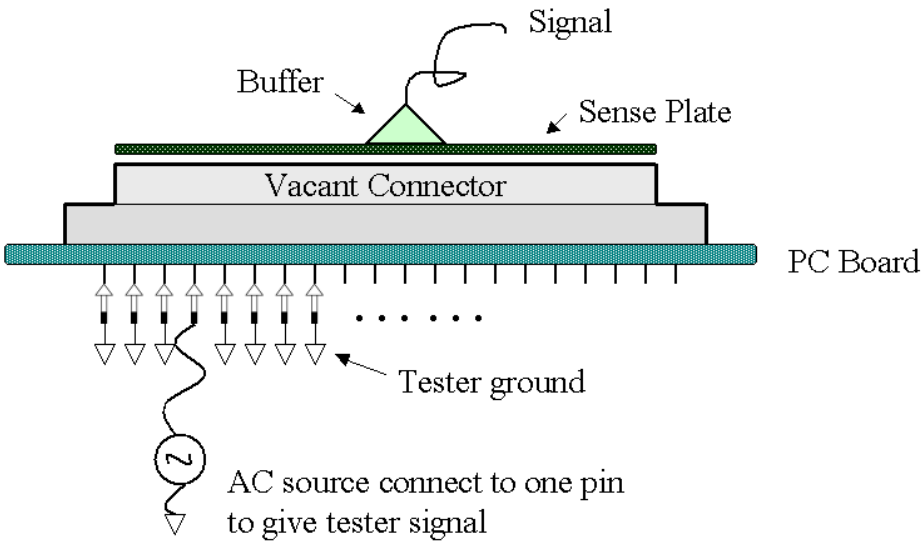


Figure 2- 3. The 'test jet' structure [3]

2.3. Advantages and Challenges

The Capacitive Lead Frame Testing technique has demonstrated good resolution for solder defects [1]. Capacitive Lead Frame Testing technique is a way to test the signal pins without having to give power to the whole chip or PCBs, which belongs to the 'Unpowered opens test' genre [1] as described in Chapter 4 .

The technique needs neither complicated programming nor debug to digital, analog or mix-signals devices, which is a relatively easy testing method [8]. The Capacitive Lead Frame Testing is a reliable diagnosis to the open solder joint defect at pin level. It allows manufactures to optimize the process for best output. The technique has reduced time and damage related with fixing. It also reduced the expensive and reputation-damaging field failures. Maximum fault coverage was obtained from its ability to test both sides of PCBs. The technique can also test the IC's with un-grounded heat sinks [8]. However, the IC's with internal ground planes can't be tested normally. If the internal circuits are above the ground plane, then the open defects in ICs will be testable. Also,

the technique can't test the internal integrity of IC's including bond wire attachment and silicon integrity. In addition, the range of the technique has been extended to capacitor and parallel capacitor combinations which are used in many circuit topologies today [8].

Capacitive Lead Frame testing of connectors will be ineffective on fixed pins such as power and grounded pins. This is due to the redundant pins and bypass capacitance [2].

2.4. Improved Approach

Ground pins are very important in connectors because they can assure the signal integrity of the differential data signal pairs, which is especially important for the high speed signals [2]. If a defect exists in the ground pins problems such as loss of signal integrity margin, increased bit error rate, increased electromagnetic interference etc. will happen.

Because of the problems from redundant pins and bypass capacitance, Capacitive Lead Frame Testing is not as effective for ground pins and powered pins as for signal pins. A new approach based on Capacitive Lead Frame Testing called "Network Parameter Measurement" was developed to solve this problem [2].

In fact, the Network Parameter Measurement used same fixtures as Capacitive Lead Frame Testing shown in Figure 2-3. In the connector circuit, there are normally series-equivalent-resistors, series-equivalent-inductors, and series-equivalent-capacitors on path from one pin to the other one. Sense capacitors exit between the high pin end and the sense plate. In addition, a mutual inductance exists between neighboring pins [2]. The relationship can be found between any neighboring pins. In Figure 2-4 simple equivalent circuit models are used to show the principle.

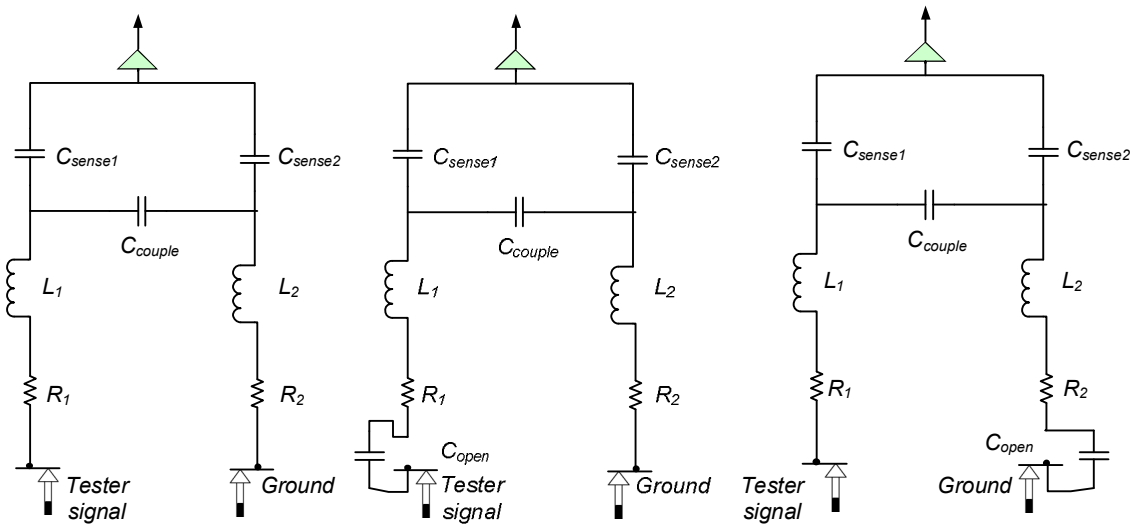


Figure 2- 4. Equivalent circuit models with (a) non-defective pins (b) open defect on tested pin (c) open defect on neighboring pin [3]

In Figure2-4 above, difference among parts (a) (b) and (c) is the open-capacitance. When open capacitance exits in different position, the impedance seen from sense plate; the voltage / current signal sent to the buffer amplifier will all be changed. To make the analysis more clear, the capacitors C_{sense1}, C_{sense2} and C_{sense3} with delta structure in Figure 2-4 is transferred to be C_a , C_b and C_c with Y structure in Figure 2-5. C_d is the defective capacitance added by the open solder.

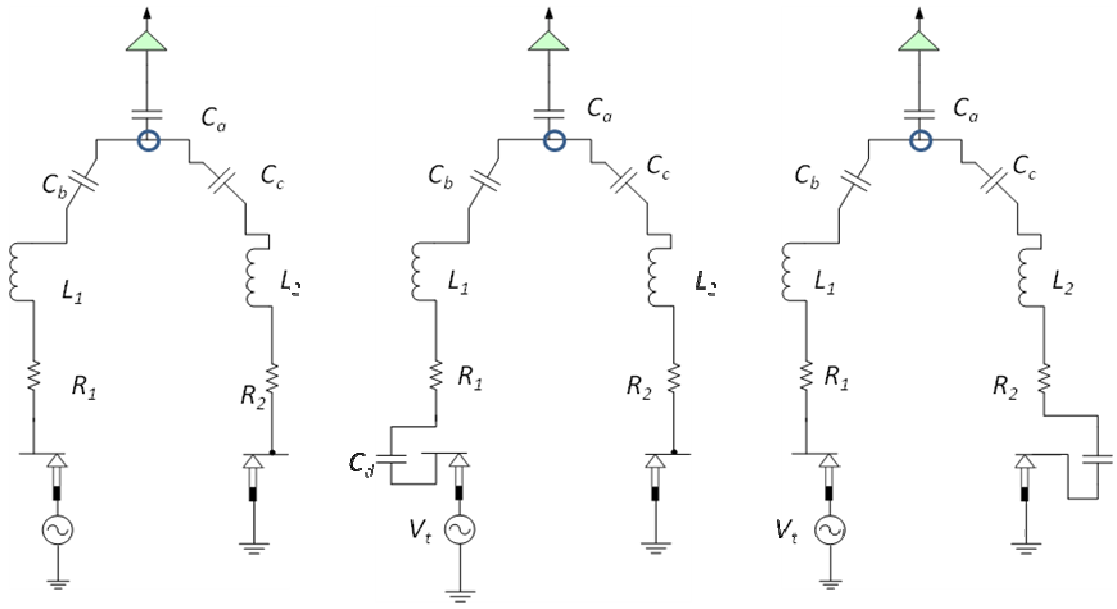


Figure 2- 5. Equivalent circuit model to neighboring pins with delta to Y transfer with (a) non-defective pins (b) open defect on tested pin (c) open defect on neighboring pin

The real values of the equivalent resistance, capacitance, and inductance can be obtained from the Agilent Pspice model. In the model, typical equivalent resistance value is around 0.0006 ohms, equivalent capacitance is about 0.002 pF and equivalent inductance is about 0.3 nH.

For example, consider the parameter set R_1=0.6 mΩ R_2= 0.6 mΩ L_1=0.3 nH L_2=0.3 nH C_{sense1}=0.002 pF C_{sense2}=0.002 pF C_{sense3}=0.002 pF and we set V_{test}=2 V f=8000 Hz

$$C_a = C_{sense2} + C_{sense3} + \frac{C_{sense2} * C_{snese3}}{C_{sense1}} \tag{2.1}$$

$$C_b = C_{sense1} + C_{sense3} + \frac{C_{sense1} * C_{snese3}}{C_{sense2}} \tag{2.2}$$

$$C_c = C_{sense2} + C_{sense1} + \frac{C_{sense2} * C_{snese1}}{C_{sense3}} \tag{2.3}$$

The condition of the pin, i.e., no-open defect, open defect at tested pin, or open defect at neighboring pin changes the impedance measured across the tested pin, which will lead to the change of voltage (V_o) in the fork point (the circled point in Figure 2-5). The change of V_o will directly result in change of the current (I_{sig}) flowing into the buffer amplifier.

The voltage and current signal change in the three cases is illustrated with mathcad plot in Fig 2-6, Fig 2-7 and Fig 2-8.To simplify the description, the impedance of buffer is set to be zero here.

For non-defective connector:

$$Z_n = \frac{\frac{1}{j*\omega*C_a}\cdot(\frac{1}{j*\omega*C_c}+j*\omega*L_2+R_2)}{\frac{1}{j*\omega*C_a}+(\frac{1}{j*\omega*C_c}+j*\omega*L_2+R_2)}+R_1+j*\omega*L_1+\frac{1}{j*\omega*C_b} \quad (2.4)$$

Z_n is the impedance seen from the signal generate into tested pin without defects.

$$V_o = \frac{V_t}{Z_n}*\frac{\frac{1}{j*\omega*C_a}\cdot(\frac{1}{j*\omega*C_c}+j*\omega*L_2+R_2)}{\frac{1}{j*\omega*C_a}+(\frac{1}{j*\omega*C_c}+j*\omega*L_2+R_2} \quad (2.5)$$

Then, $$I_{sig} = \frac{V_o}{\frac{1}{j*\omega*C_a}} \quad (2.6)$$

$$|V_o| = 0.5\ \text{V} \qquad |I_{sig}| = 1.005e-10\ \text{A}$$

When the pin under test opens:

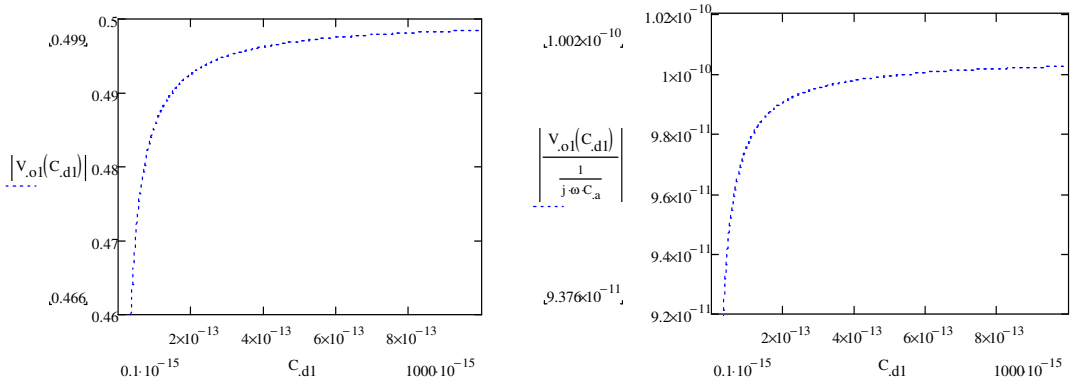
$$Z_{d1}(C_d) = \frac{\frac{1}{j*\omega*C_a} \cdot (\frac{1}{j*\omega*C_c} + j*\omega*L_2 + R_2)}{\frac{1}{j*\omega*C_a} + (\frac{1}{j*\omega*C_c} + j*\omega*L_2 + R_2)} + R_1 + j*\omega*L_1 + \frac{1}{j*\omega*C_b} + \frac{1}{j*\omega*C_d}$$

(2.7)

$Z_{d1}(C_d)$ is the impedance seen from the signal generate into tested pin with open solder defect at tested pin.

$$V_{o1}(C_d) = \frac{V_t}{Z_{d1}(C_d)} * \frac{1}{j*\omega*C_a} \tag{2.8}$$

Since the parameter C_d is a variable in the formula above, we can use mathcad to plot the relationship between C_d and the signal amplitude. As shown in Figure 2-6, voltage and current signals all increase with the increase of C_d.



(a) (b)

Figure 2- 6. Signal amplitude changes with C_d when the tested pin open (a) the voltage amplitude at fork point (b) current amplitude flowing into the buffer

When the neighboring pin opens:

$$Z_{d2}(C_d) = \frac{\frac{1}{j*\omega*C_a}\cdot(\frac{1}{j*\omega*C_c} + j*\omega*L_2 + R_2 + \frac{1}{j*\omega*C_d})}{\frac{1}{j*\omega*C_a} + (\frac{1}{j*\omega*C_c} + j*\omega*L_2 + R_2 + \frac{1}{j*\omega*C_d})} + R_1 + j*\omega*L_1 + \frac{1}{j*\omega*C_b}$$

(2.9)

$Z_{d2}(C_d)$ is the impedance seen from the signal generate into tested pin with open solder defect at neighboring pin.

$$V_{o2}(C_d) = \frac{V_t}{Z_{d2}(C_d)} * \left[\frac{\frac{1}{j*\omega*C_a}\cdot(\frac{1}{j*\omega*C_c} + j*\omega*L_2 + R_2 + \frac{1}{j*\omega*C_d})}{\frac{1}{j*\omega*C_a} + (\frac{1}{j*\omega*C_c} + j*\omega*L_2 + R_2 + \frac{1}{j*\omega*C_d})} \right] \qquad (2.10)$$

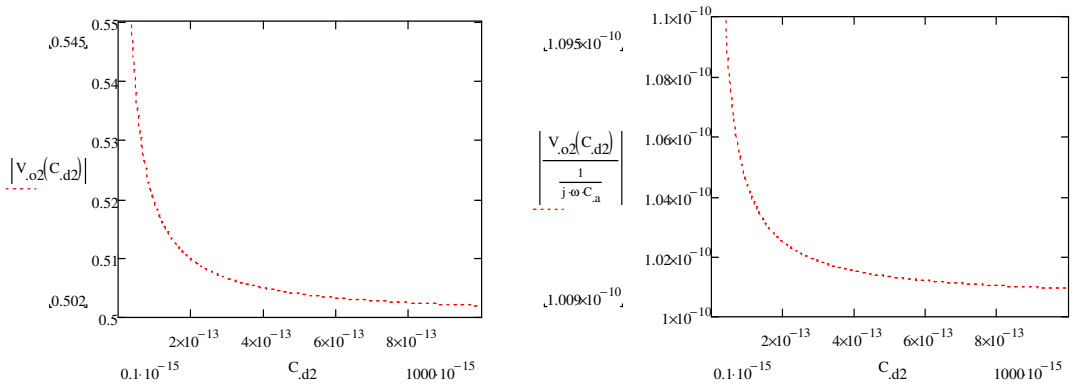


Figure 2- 7. Signal amplitude change with C_d when the neighboring pin opens (a) the voltage amplitude at fork point (b) current amplitude flowing into the buffer

In Figure 2-7, the voltage and current signal amplitude decrease a lot with the increase of C_d.

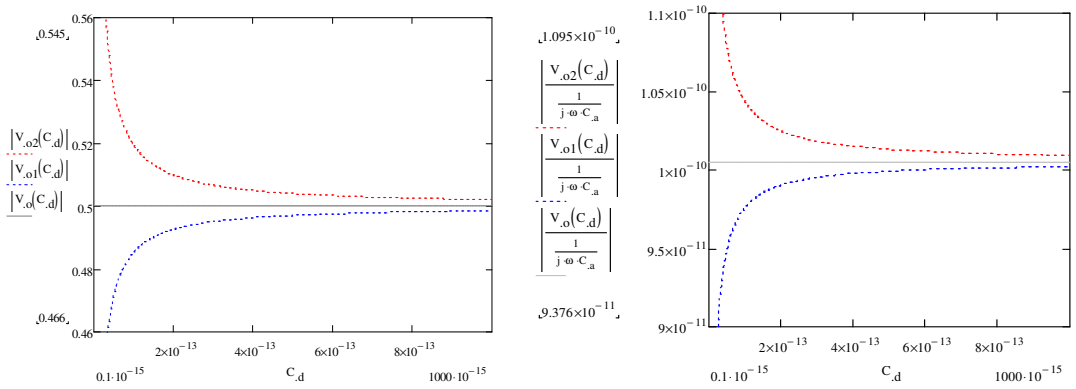


Figure 2- 8. Combined signal amplitude change in the three situation (a) the voltage amplitude at fork point (b) current amplitude flowing into the buffer

Based on formulas above we could compare the three cases as can be seen in the three plots of Figure 2-8. Since there is no C_d in the non-defective situation, the signal amplitude in the situation won't change with C_d. That will result in the straight line in grey color in Figure 2-8 (a) and (b). The red waveforms in Figure 2-8 (a) and (b) are signal amplitude with open tested pin, which is always higher than the non-defective signal amplitude. The blue waveform is the signal amplitude with open neighboring pin, which is always lower than the non-defective signal amplitude. Comparing the three waveforms in the Figure 2-8 (a), (b) we can see that open tested pin causes significantly lower signal amplitude, while the open neighboring pin can cause higher signal amplitude than the normal value.

With the three waveforms in the Figure 2-8 (a) we can also see that difference among waveforms is large when C_d is small. In the plate capacitor $C_d = \frac{\varepsilon * A}{d}$, where ε is the dielectric constant, A the area of plate, d is the distance between two plates. Since the other two parameters can be considered to be constant, when C_d is small d must be large.

Then, when the solder open is large enough, which would lead to a small open defective capacitance, the signal amplitude deviation from normal value will be clear.

The buffer in fact is a signal amplifier, which has an impedance value. If virtual ground wasn't assumed at the buffer amplifier, we can assume constant amplifier impedance Z_{amp} connected to the upper end of C_a (Figure 2-5). Then V_{signal} can be calculated from voltage divider rule as:

$$V_{signal} = V_o * \frac{Z_{amp}}{\frac{1}{j*\omega*C_a} + Z_{amp}} \tag{2.11}$$

The signal voltage has similar trend to that of V_o discussed previously.

In fact, the NPM is not limited to the grounded and power pins on connector. It can be applied to the signal pins as a subset of the capacitive Lead Frame Testing, which will explain the high abnormal readings in the test data in the following Chapters.

Chapter 3.
Principal Components Analysis Based Outlier Detection

Threshold settings such as relative thresholds based on standard deviation are used currently with capacitive testing to differentiate the normal values from the abnormal capacitance values. However, the shrinking size of the features and the resulting lower capacitance of signal pins make it more and more challenging to set such a threshold values. Non-optimal threshold settings can give rise to higher false fails or false passes. The challenge is further compounded due to the fact that each pin tested has a threshold that is different from others, yet often correlated to them. Furthermore, mechanical parameters such as spacing between the plate and the connector/device vary from one mounting on the tester to another. Similarly, the capacitance value corresponding to a pin may vary from board to board due to the fact that components are from different vendors. These and other factors combine to make the selection of appropriate thresholds a challenging task.

As an effective tool in this thesis we present a Principal Components Analysis (PCA) based technique to analyze the Capacitive Lead Frame test data and to detect defective boards,. PCA has been a well-known multi-dimensional correlated data analysis method for more than 100 years. PCA has been successfully applied to visualizing data, data exploration, outlier detection, compressing data etc. In electronic testing arena, PCA has been used to detect outlier Integrated Circuits [5][6].

In this Chapter, a background of the PCA will be provided, which includes the transformation formulae and test statistics for outlier detection.

3.1. Introduction to PCA

One of the best-known multivariate analysis methods, Principal Components Analysis (PCA) was introduced by Pearson at the beginning of the 18th century. Then, it was further developed by Hotelling in 1933[9].In Principal Components Analysis, a multi-dimensional interrelated data set is transformed to be a much lower dimensional data set while retaining as much of the information as possible. This kind of dimension reduction can be achieved by transforming the original data set into a series of uncorrelated Principal Components (PCs). In the series of PCs, data is ordered from high variance to low variance, where the first few PCs can contain the most information [9].

The PCA transformation can be specified by following steps: First, find the direction that achieves the largest projection variance from the data projection; After finding the direction above, we continue to look for another direction, which is orthogonal to that one and contains as much of the remaining variance from the data projection as possible. That is shown in Figure 3-1. Then, we look for the third one, and so on. In fact, the direction can be considered as the linear combinations of the original data set. The process continues until the remaining linear combinations or Principal Components are found.

The aim of the direction seeking is to capture the variability in the original data set [10]. Of course, the PCA should be applied to inter-correlated data. If no such correlation exists in the data, PCA representation won't have an advantage over the original representation.

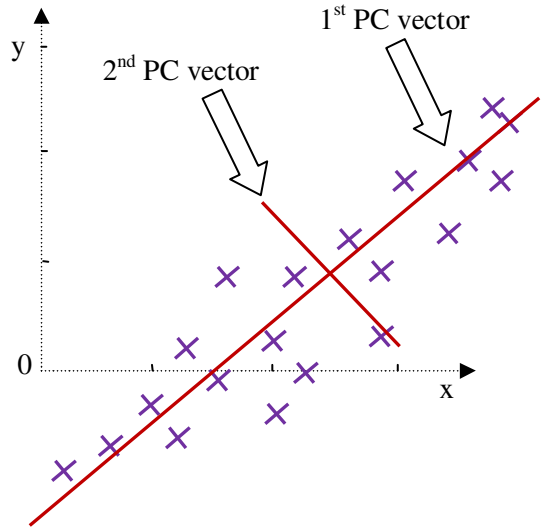


Figure 3- 1. Test observations with the first two Principal Components working as axis

3.2. Mathematical Representation of PCA

Let the $M_{m\times n}$ be the matrix of capacitance measurements, where m is the number of boards, each with n measurements corresponding to the n tested pins. Let M_c be the centered matrix where mean value of its column is subtracted from each element. The mean value of the column is the "measured capacitance of a pin averaged over all boards". For example the mean value of the i_{th} column can be defined by

$$M_{mean\_i} = (M_{1i} + M_{2i} + \ldots + M_{ni}) / n \qquad (3.1)$$

where M_{ni} is the n_{th} component in the i_{th} column.

With centered raw data matrix, there are two methods that may be employed to compute the Principal Components.

The first one, which is widely used in literature, involves computing the covariance matrix C of the centered matrix first:

$$C = M_c * M_c^T \tag{3.2}$$

Sometime the correlation matrix is also needed when data types inside the original matrix are much different from each other. The correlation matrix R is the normalized covariance matrix such as C, in which element $R_{mn} = \frac{C_{mn}}{S_m S_n}$ (3.3)

The S_m and S_n are square roots of variance to corresponding to column and row of element (m,n). C_{mn} is the element (m,n) in covariance matrix C above [5].

After the above steps, PC calculation can be carried out by Eigen Value Decomposition (EVD): $E = V * L * V^{-1}$ (3.4)

where E is either the correlation matrix R or the covariance matrix C. PC score can be calculated from [5]

$$Z = M_c V \tag{3.5}$$

The second approach applies Singular Value Decomposition (SVD) to the centered real data matrix M_c. SVD technique computes U, S, V such that

$$M_c = USV^T \tag{3.6}$$

where the columns of U and V are called singular vectors, and S is the diagonal matrix containing the singular values [10][11][12][13]. In factorized style:

$$M_c = USV^T$$

$$= \begin{pmatrix} u_1 & u_2 & \dots & u_m \end{pmatrix} \begin{pmatrix} \delta_1 & & & & \\ & \delta_2 & & & \\ & & \dots & & \\ & & & \delta_r^T \dots & \\ & & & \dots & \end{pmatrix} \begin{pmatrix} v_1^T \\ v_2^T \\ \dots \\ v_n^T \end{pmatrix} \tag{3.7}$$

In the Equation (3.7), 'r' is the rank of M_c. When r<m, other elements following δ_r^T are equal to 0.

The columns of U and V are called singular vectors. S is a diagonal matrix that contains the singular values. In outer-product form [10] it is:

$$M_c = \sum_{i=1}^{n} \delta_i u_i v_i^T \tag{3.8}$$

From the factorization above, the PC score or Z-score matrix $Z_{m\times n}$ is given by equation (3.5) where $Z = M_c V$

Each board is now characterized by n PC scores or Z scores, represented by the n columns of Z. The first coordinate (called the first principal component) account for the direction that contains most variance of data projection then the second one, and so on. In fact, the variance of each column in the Z score matrix is automatically ordered by the algorithm from high to low. Figure 5-3 in Chapter 5 illustrates an example of Z score matrix. As can be seen, Z score variance values decrease from high to low along the column number.

In SVD, algorithm forces the first component to go through origin while maximizing the variance projected. Use of the centered matrix M_c ensures that the first coordinate is not forced to pass through the origin, and thus can catch the real maximum projected variance from the data [10][34][35][36]. Figure 3-2 shows an example of a PC plot in two-dimensional data. In Figure 3-2 (a), the largest Principal Component is calculated from non-centered dataset which starts from origin. As can be seen, this principal component doesn't really catch the largest projected variance from data. When the PC is

calculated with centered data the largest PC will be the one expected. As shown in Figure 3-2 (b), start point of the first PC vector is not limited to the origin, and is able to capture the largest projected variance.

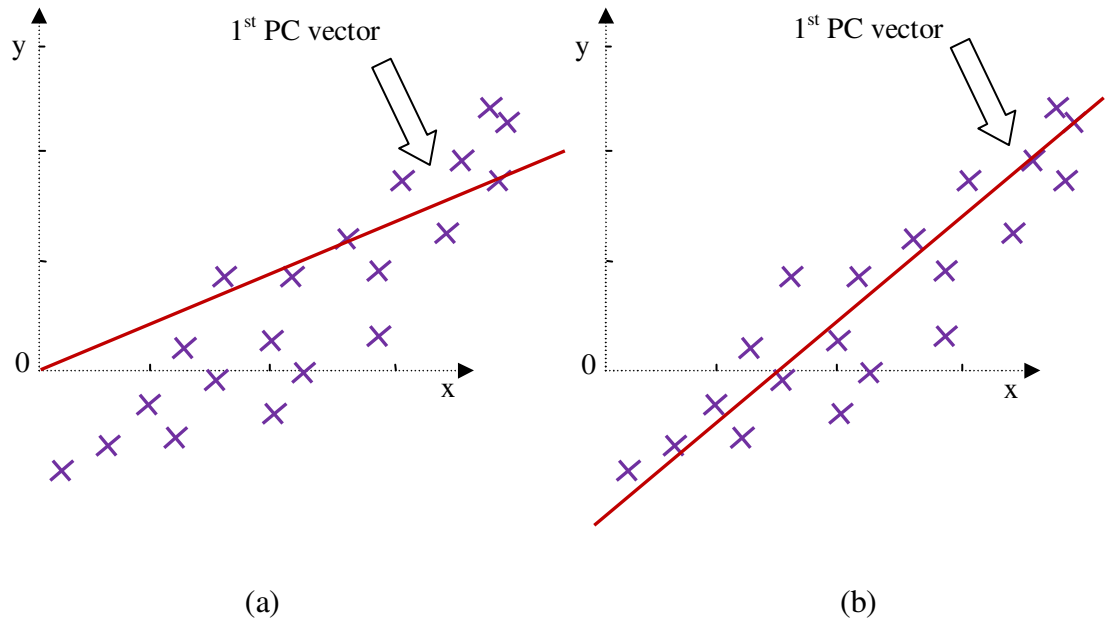


Figure 3- 2. Comparison the first PC vectors calculated from (a) non-centered data and (b) centered data

Comparing the two methods EVD and SVD, SVD is the more robust, reliable and precise method with no need to compute the input covariance/correlation matrix. In numerical analysis fields, SVD is well known for its convergence and stability properties and it also works well with ill conditioned matrices [13]

3.3. Outlier Detection with PCA

Outliers are observations (test results of boards or connectors) that are numerically distant from the rest of the observations (of boards or connectors). In n-pin (n variables) test results, the outliers will be much different from the rest of the devices in the n-

dimensional space. In multivariate data, outliers can also be observations that never show extreme values in any one dimension. That is because of the general data structure (or plot trend) of outlier does not confirm with the rest of the observation. Such kind of outliers will not be detected by inspecting variables one by one [11].

Gnanadesikan and Kettenring (1972) discussed the plots that use only a few PCs to detect outliers. Such kind of plots (as shown later in Figure 5-4 and Figure 5-5) use the Principal Components as axes; represent the observations (devices) as scatter points to roughly show the outliers [11].

3.4. Test Statistics

To detect the outliers, i.e., boards or connectors with test measurement patterns that are significantly different from the rest of the boards or connectors, a distance measure is used. Since the variance of the first and last few PCs vectors contain different information, the first few PCs and last few PCs can detect different types of outliers. Different test statistics such as $d_{1i,}\, d_{2i,}\, d_{3i}$ and d_{4i} have been defined which can be applied for detecting different type outliers [11].

$$d_{1i}^2 = \sum_{k=p-q+1}^{q} z_{ik}^2 \tag{3.9}$$

The calculation of statistic d_{1i}^2 (suggested by Rao [11]) is based on the Principal Components. If there is no outlier, the independent observations d_{1i} ($d_{1i} = \sqrt{d_{1i}^2}$) should follow a Gamma distribution. In the formula *of* d_{1i}, Z_{ik} is the value of the k_{th} PC for the i_{th} board; *p and q* define the sequence numbers of the first and last PC used for the evaluation, which also determine the number of PCs selected.

$$d_{2i}^2 = \sum_{k=p-q+1}^{p} \frac{z_{ik}^2}{l_k} \tag{3.10}$$

A weakness of d_1 is that it gives insufficient weight to the last few PCs, especially when the one with large variance and the one with very small variance are used together. So d_{2i} was proposed in [11] as an alternative to d_{1i}. The PCs inside of d_{1i} are normalized by the variance of its column. In the d_{2i} formula above, l_k is the variance of the k_{th} PC.

Gnanadesikan and Kettenring [11] also use the d_{3i} as test statistics. d_{3i} can detect outliers in the data with large effect on first few PCs.

$$d_{3i}^2 = \sum_{k=1}^{p} l_k z_{ik}^2 \tag{3.11}$$

Hawkings also defined a d_{4i} statistic that works effectively in many experiments [11].

$$d_{4i} = \max_{p-q+1 \le k \le p} \left| \frac{z_{ik}}{\sqrt{l_k}} \right| \tag{3.12}$$

In some literatures, $d_{0i}^2 = \sqrt{\sum_{k=1}^{p} \frac{z_{ik}^2}{l_k}}$ [8] (3.13), and $X_i = \log 10(\max_{p-q+1 \le k \le p} \left| \frac{z_{ik}}{\sqrt{l_k}} \right|)$ [4] (3.14)

are also used as test statistics. As can be seen, d_{0i} is very similar to the square root of d_{2i}, except its p and q are set to be the same value. The X_i value could be viewed as the logarithm of the d_4 value.

The four d_i statistics presented above have been suggested for outlier detection with Principal Components [11]. Using the information of several PCs (subset of the Z score matrix), all of them can change the multivariate analysis into single-variable analysis.

Since the first and last few PC vectors contain different information, the first few PCs and last few PCs can detect different types of outliers [11]. The best test statistics for a given problem depends on the data type and the purpose of the test.

The basic concept for PCA based test technique was described above. We use two different schemes based on the same technique for testing the PCBs. First, we use a *global method*, which uses the entire data set for the board to identify outliers. In this case, the PCA based algorithm would be applied to the whole measurement matrix. This method takes into account the variations such as tester to tester variation or fixture to fixture variation more effectively as effect of such variations manifest over the entire set of measurements. However, a weakness of this method is the fact that an open in one pin influences the capacitance values of only few other pins in its neighborhood. Thus the overall effect on the test statistic is like to be somewhat smaller, as the analysis is based on the dataset for the entire set of pins.

Therefore in *localized method*, we propose an outlier detection scheme based on a small window of pins at a time. The original measurement matrix M is first sorted according to the pin number in relative physical layout area. Then the matrix is first vertically separated into different smaller matrices (test windows). A calculation similar to that in global analysis is applied to each window. To cover the entire connector or the set of pins, the test is performed by carrying out the evaluation in each window to cover the entire set of pins, one window at a time. Thus we have two options, overlapping

windows or non-overlapping windows. We will address the global and localized method in detail in Chapters 5,6,7,8.

Chapter 4
PCB Test Data Sources

This chapter introduces the test background and datasets of PCB measurements used in this work. Section 4.1 provides the background for board testing. Then the dataset used for further analysis in later chapters is illustrated in Section 4.2.

4.1. Background for Boards Test

Some of the defects that may be introduced during board manufacture include the following: opened solder; insufficient, excess or malformed solder joints; missing devices; shorts which can be caused by device mis-registration; excess solder; dead devices; incorrect device placement; polarized devices wrongly placed; and misaligned parts [1]. To ensure that the boards are defect-free before they are shipped to customers, different types of tests are applied to detect such defects. The different board test steps are discussed next.

Structural tests, functional tests and system tests are used for board testing. As is shown in Figure 4-1, these test steps are employed one by one after the manufacture. The following paragraphs will introduce these steps in detail.

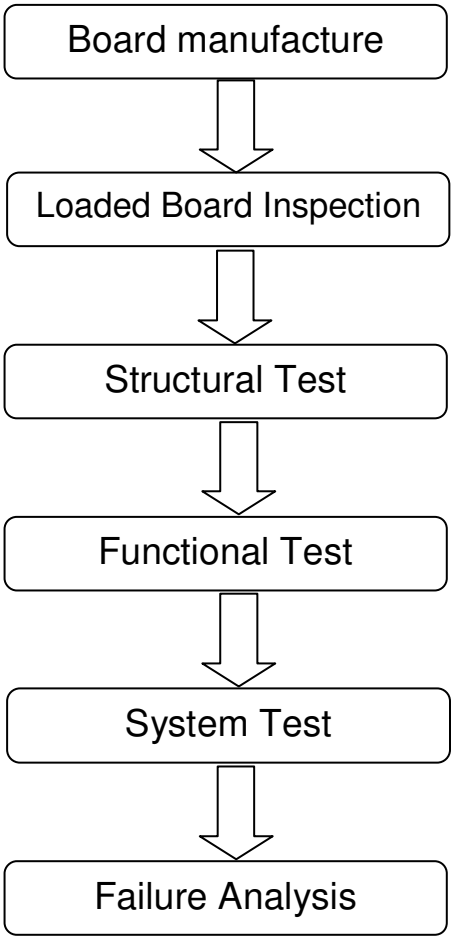


Figure 4- 1. Board test steps

Loaded Board Inspection is a test step that applies some operation wave such as X-ray to illuminate the area of interest and capture an image. By comparing the image with specifications, the quality of the device under inspection can be judged. Loaded Board Inspection includes Automatic Optical Inspection (AOI) and Automatic X-ray Inspection (AXI). AOI and AXI use two different operation wavelengths, visible light and X-ray illumination. *AOI* does not utilize penetrating radiation. The light illuminates the board from different directions. Then the radiation is reflected from the exposed board surface. If the part cannot be illuminated, it cannot be tested. *AXI* on the other hand can detect the

inner parts of objects like copper, board materials and integrated circuits. Both 2-dimensional and 3-dimensional inspection can be achieved by AXI. Two-dimensional test is simpler but the image suffers from the degeneration of resolution when there are components on both sides of the board. This problem can be solved by 3-dimesional tests, but at the cost of more mechanical complexity [1, 21].

Different from other tests, Loaded Board Inspection doesn't need much programming information. However, it can find the defects that cannot be found by ICT such as the alignment and solder quality problems.

Structural Test is the test step that inspects the internal board structure to ascertain whether the board is built correctly. It is able to identify defects such as wrong components, incorrectly installed components and missing components. Structural tests includes *In-Circuit-Test (ICT)* and *Boundary Scan.* ICT utilizes electrical probes to test PCBs for defects such as opens. Most of the time it is a low frequency test technique. During the test, the board is not operated as it is normally intended to [1]. ICT employs unpowered and powered test.

With *Unpowered test* the whole PCB doesn't need to be activated, i.e., powered on, during the test. Signals may be applied only to the part under test to obtain responses. For testing a short in a specific component, for example, a load resistor in series with a small voltage source is connected. Then a limited voltage or current stimulus is applied to the part under test. Since the voltage across the load resistor will be monitored, once the voltage exceeds a threshold voltage, a short defect may be considered to exist [1].

When the unpowered tests are applied to PCB components such as resistors, inductors and capacitors, which have associated nominal values and a tolerance, the

unpowered test is also termed as '*unpowered analog test*' to measure these component values. Finding open signal pins on connectors or ICs with Capacitive Lead Frame testing and Network Parameter Measurement (NPM) are examples of *unpowered analog tests*.

Powered In-circuit Digital Test is a test employing a digital sequencer to test digital devices such as ICs on the boards. Digital input signals are applied to the devices while monitoring digital responses at the same time. In practice, there may be several similar processes running at the same time.

Programmable analog parameters such as voltage, slew rate, and receiver high/low voltage comparison windows are needed for drivers and receiver / comparator circuits. *Powered In-circuit Digital Test* also needs to utilize lots of memory to store digital stimulus, response [1].

Powered Mixed-signal Test is a test used when both analog and digital test are needed. The digital subsystem and analog subsystem on the In-Circuit tester can coordinate the tests in this case. For example, to test a Digital-to-Analog converter on a board, the digital subsystem may simulate digital data corresponding to a sine wave, while the analog subsystem measures the frequency or distortion of the signal.

In-circuit-test is effective in detecting the presence, correctness, orientation, liveness (whether the component is 'dead'), shorts and open defects. Since the visual inspection can only determine whether the device is present and appears correct, it cannot tell if the device is dead or defective. In-circuit-test can apply electrical tests to these components. However, ICT is essentially useless for detecting defects associated with device alignment and joint quality [1]. The In-Circuit-Test steps for printed-circuit boards are presented in Figure 4-2.

ICT has fast test speed, provides good defect detection at component level and facilitates automatic test development. However, ICT has a high cost and may cause high board stress due to high probe density.

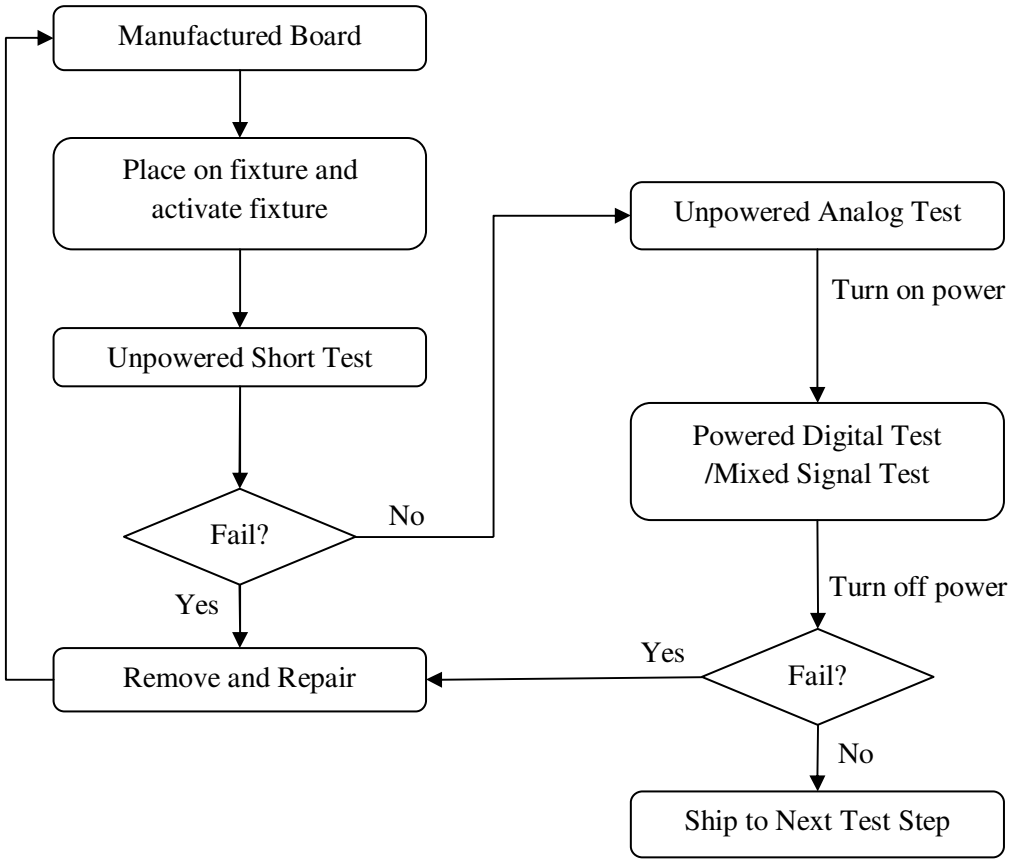


Figure 4- 2. ICT board test steps [1]

Boundary-Scan Tests is a method to test PCB wire lines or IC sub-blocks without applying physical probes. During the test, build-in Boundary-Scan devices in digital ICs are utilized to perform testing. When not used, these circuits and devices turn back to normal functions.

Normally, Boundary-Scan Tests can also be considered as a subset of Design-For-Test rules, which will facilitate the board testing [1, 31]. The specifications in IEEE STD

1149.1[1] developed by the Joint Test Action Group (JTAG) for Boundary-scan, Boundary-scan test describes this semi-automated and fast test method. It also requires minimum test access, which helps the board testing when test probes are compromised by the layout density. Defects such as opens and shorts in digital ICs can be detected by this method [1，22-26].

Functional Test is a test to confirm that a board meets its design criteria. The board under the test needs to operate toward its original design purposes on a test platform for verifying its performance specification. A board that passed the structural tests may not pass the *Functional Test* step [4, 31].

Normally, the functional test has to and is capable of providing very good test coverage. However, it cannot provide good diagnostics; in other words, it is not able to tell the exact defects or the defect locations. Functional tests need a long test time and a costly design effort [32].

System Test is a type of test where the boards are inserted into the final system and then the system is turned on to see if the whole system works well with the product. This test can only give pass/ fail but can't give detailed defect information [1].

In the practice, most of the tests above will collaborate to filter the bad boards. In modern test systems, the total test time for all of the tests above may only last 30 to 50 seconds.

4.2. PCB Measurements Datasets

This thesis uses Capacitive Lead Frame Test (called TestJet in industry) data for several different PCBs obtained using Agilent testers. When performing TestJet

measurements, the tester uses relays to connect the board pin under test to an AC source running at 8192 Hertz and around 250 mV peak-to-peak. The voltage is so low that it minimizes the chance of diodes on the board turning on. All the pins except for the pin-under-test are connected to the ground.

The sense plate transfers the signal to a buffer amplifier and then to a signal analyzer, which detects the capacitances in femto-farads scale with sensed signal. The sensor may get an average value over several periods of AC signal.

Measurements in the test dataset we use correspond to connectors residing on boards tested in a working production line. For each connector on the board, all but the grounded/VDD pins were tested and capacitance measurements obtained. A board could be tested more than once to test the repeatability, so there can be multiple data records for a single board (each record is termed as a *boardrun*).

The characteristics of datasets used in this thesis are summarized in Table 4.1.

Dataset Name	# of Connectors in data	Tested Pins	Boards Measurement
D0	5	670	22
J24	1	145	15
J25	1	150	15
J27	1	147	15
J28	1	151	15
J31	1	77	22
D1	7	1053	20
J7	1	132	20
J10	1	132	20
J13	1	133	20
J16	1	131	20
J41	1	130	20
J45	1	133	20
J47	1	129	20
J50	1	133	20
D3	4	594	83
Data3_j24	1	145	83
Data3_j25	1	150	83
Data3_j27	1	148	83
Data3_j28	1	151	83
LVLD1	2	262	6
J3_norm	1	142	6
J10_norm	1	120	6
LVLD2	2	262	5
J3_all	1	142	5
J10_all	1	120	4
LVLD3	2	262	4
J3_projection	1	142	4
J10_projection	1	120	4
LVLD4	2	262	4
J3_all	1	142	4
J10 all	1	120	4

Table 4- 1. Datasets used in this thesis

D0, D1, D3, LVLD1, LVLD2, LVLD3, LVLD4 identify seven different datasets used in this thesis. For each dataset, there may have several different connector measurements. For example, in dataset D1, the data in Data3_j24 and Data3_j27 are from connector j24 and j27.

In each test dataset there are some unique tested boards. When needed, some boards have been tested several times. We call each measurement set corresponding to one set of measurements of a board as a boardrun. Thus the same board may be associated with multiple board runs. Datasets in D3 are composed of four different types of connector test measurements. Each dataset measurement contains 47 unique boards. However, some boards were tested multiple times resulting in the 83 board measurements termed as boardruns in the following.

D1, a relatively comprehensive set, includes 17 unique boards with a total of 20 boardruns. Each of the board in Data_D1 includes eight connectors: j7, j10, j13, j16, j41, j45, j47 and j50. The multiple connectors lead to over 1000 pins tested per board.

Data in LVLD1, LVLD2, LVLD3 and LVLD4 all contain measurements of connectors j3 and j10, which are DDR2 memory card connectors. However, j3 on this particular board is mounted at a 45 degree angle, i.e., the board that is plugged in will be at a 45 degree angle to the main board. J10, like other connectors, is perpendicular to the board. For example, Data_j3 is from j3 connector and corresponds to 6 unique boards.

The four measurement datasets in D0, namely j24, j25, j27 and j28 are based on the tests for 240-pin DDR memory card connector. However, j30 is a 140-pin connector. The test data sets described above will be used for the analysis in the following chapters. All of the datasets were provided by Agilent Technologies.

Chapter 5.
PCA Analysis of Board Measurements with d Statistics

This chapter illustrates the use of the PCA based outlier detection approach. Different test statistics are applied to data sets Data3_j24 and Data_D1. In Section 5.1, dataset Data3_j24 is introduced and analyzed to investigate the effectiveness of PCA. Then in Section 5.2 different test statistics based on PCA are applied. To make the comparison between test statistics clearly, different subsets of most-significant and least-significant PCs are used. To support the conclusion drawn in Section 5.2, Section 5.3 applies the method to another dataset Data_D1. Selection of test statistic parameters is also discussed.

5.1. Analysis of Measurements of a Single Connector with Principal Components

Figure 5-1 is the plot of 83 boardruns for Data3_j24 while Figure 5-2 shows these same measurements with clear outliers removed. As can be seen from Figure 5-1, while the general pattern remains the same, the measurement variation among different boardruns is often detectable. The six boardruns 17, 18, 19, 20, 21, 22 would be identified as outliers by manual inspection of data. Of course there are other outlier boardruns besides the six above that are harder to identify (e.g., boardruns 4, 5, 14, 15……). Figure 5-2 shows the same data, but without the six outliers. Test results for pins 210 to 240 (the right part of plot) have a significantly higher variation compared to the others. A close inspection of the plots for different boards shows strong fluctuations in measurements with values that are far different from each other.

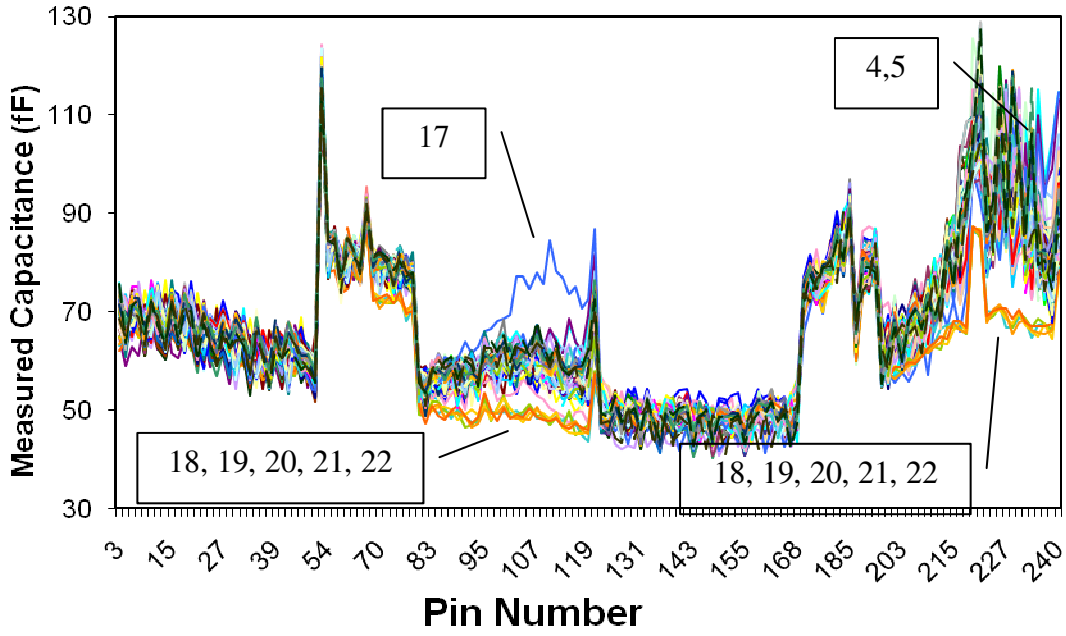


Figure 5- 1. Plot of raw measurements of Data3_j24

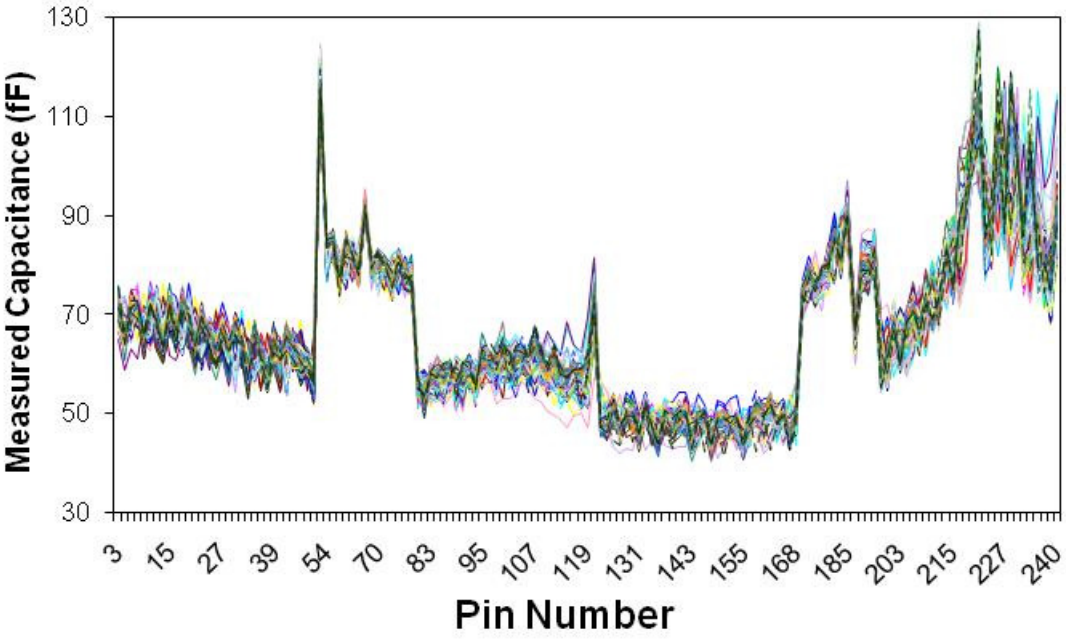


Figure 5- 2. Plot of raw measurements of Data3_j24 with outliers removed

To present the point clearly here, boardrun 17 and boardruns 18, 19, 20, 21, 22 would be used as clear outliers, which can be inspected from the difference between Figure 5-1 and Figure 5-2.

5.1.1. PCA Result for Data3_j24

Principal Components can be utilized for outlier detection. The Principal Components are calculated with centered SVD algorithm. The calculated Z score matrix contains PC vectors.

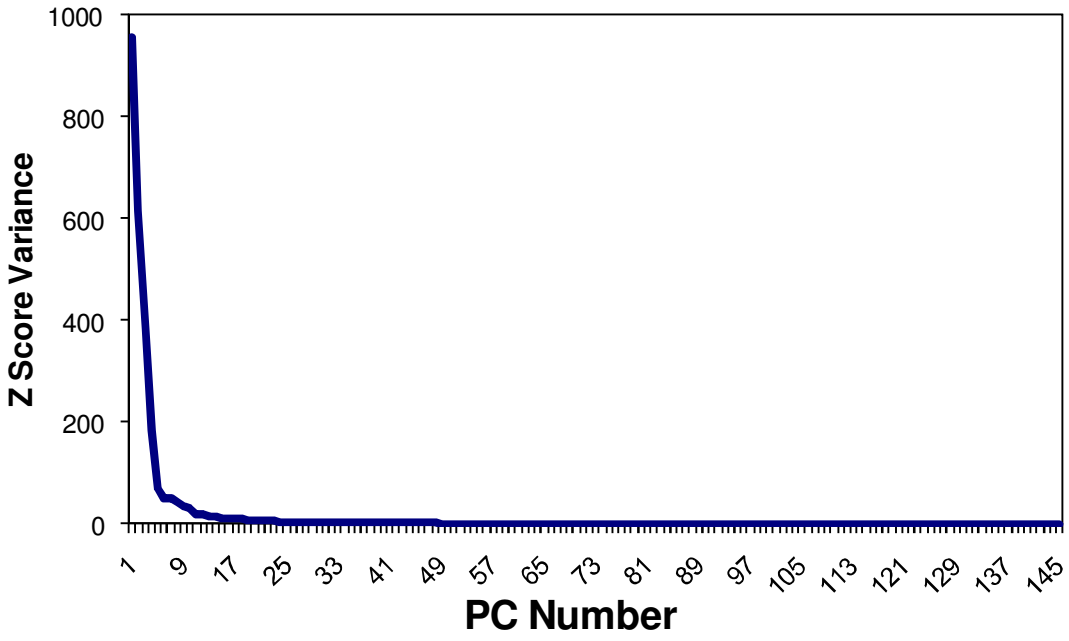


Figure 5- 3. Plot of Z score variance values for Data3_j24

Figure 5-3 shows the variance of the Z score values with PC numbers. As can been seen, variance of the Z score values change gradually from largest variance to a minimal one with the PC number. Since the first and last few PCs contain different information, the first few vectors and last few vectors can be used to detect different types of outliers.

To test the effectiveness of PCA for identifying outliers, we use a series of scatter plots of different combinations of PCs vectors. With first two PCs vectors as x-axis and y-axis respectively, boardrun number has been plotted according to these values in Figure 5-4. In the figure, boardruns 17, 18, 19, 20, 21, 22 are relatively far away from the cluster of other devices. Similar phenomenon can also be detected in the 3-Dimensional or higher dimensional plots. For example, in Figure 5-5, an additional PC vector, the third PC vector, works as the z-axis.

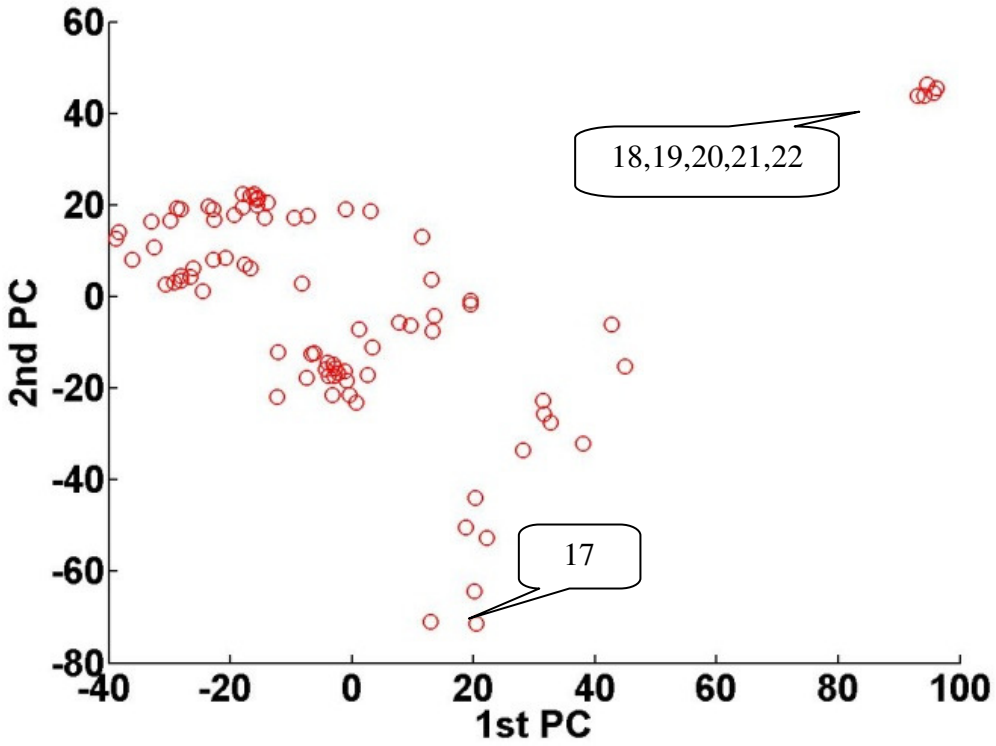


Figure 5- 4. Boardruns plot for Data3_j24 with the first two PCs as axes

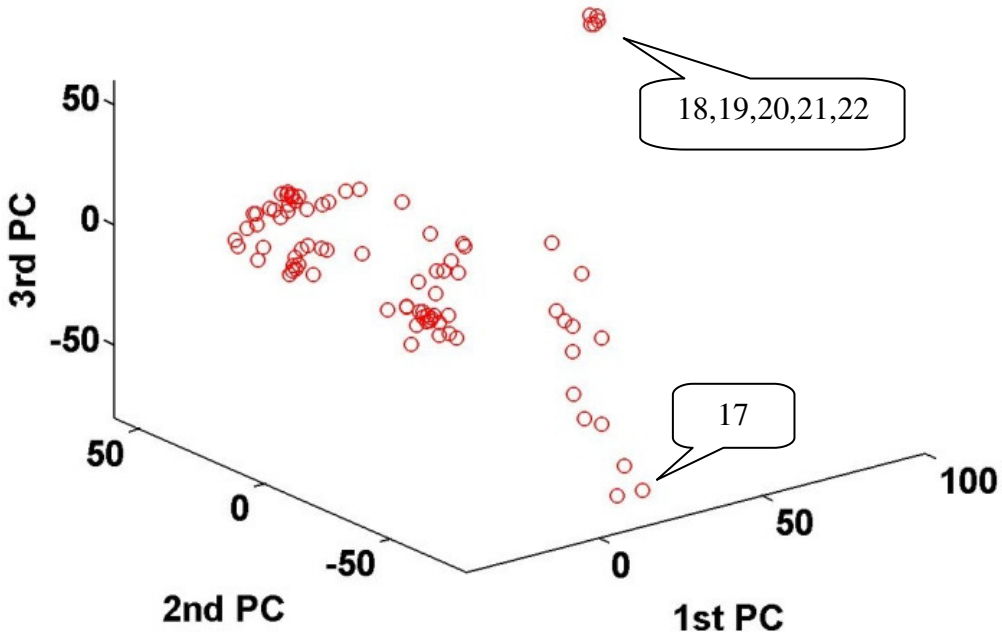


Figure 5- 5. Boardruns plot for Data3_j24 with the first three PCs as axes

From the two plots above, clear outliers like 17, 18, 19, 20, 21 and 22 could be observed as clearly different from the others, which means Principle Components appear to be effective in multi-dimensional outlier detection. However, the scatter plot analysis above is based on inspection. The multi-dimensional analysis can be reduced to a one-variable analysis with test statistics such as d_1, d_2 d_3, *and* d_4.

Calculation of test statistics d_1, d_2, d_3, d_4, with given Principal Components is applied to all of the boardruns in Data3_j24. Then the boardruns are sorted according to the respective test statistic value (d value). Since the Cumulative Distribution Function (CDF) plot can clearly show the difference between the test statistics value with boardrun numbers, all of the boardruns are plotted onto the CDF curve with respect to their own

test statistics values. The outlier boardruns would stand out at the high-end of the CDF curve.

5.2. Selections of Test Statistics and PCs

In this section, we investigate the selection of appropriate PCs for outlier detection in PCBs. The set Data3_j24 is tested with four different test statistics using different combinations of PCs.

5.2.1. '*d*' Statistics with Most Significant PCs

In the following analysis the four test statistic are evaluated and compared with the most significant 1, 3, 5, 10 PCs respectively. .

In Figure 5-6, d_1 value is represented on x-axis, six clear outliers, namely 17, 18, 19, 20, 21, and 22, stand out at the high end of the CDF curves. The clear break after the six devices number shows that these six boardruns are far away from the other devices in the multi-dimensional test result. Some other boardruns after the clear break like 15, 14, 5, 4, 13, 11, 16, 8, 9 and 10 can also be grouped as potential outliers.

CDF plots in d_2 and d_4 scale are somewhat similar. In d_2 CDF curve (Figure 5-7), the six boardruns only appear at the high end when the first PC is used. In the other 3 situations, boardrun 4, 5 show even higher d_2 values than the boardruns 18, 19, 20, 21, 22. In addition, the break after device 18, 19, 20, 21, 22 is not clear enough to separate them from others. The boardrun numbers sorted in d_4 scale only match the inspection result when the first PC is used, as shown in Figure 5-8 (a). However, the break after the six boardruns makes all other boardruns compact together which results in the potential

outliers to be undetectable. Analysis with the first 3, 5, 10 PCs give boardruns 5 and 17 higher d_4 values than boardrun 18, 19, 20, 21, 22 do.

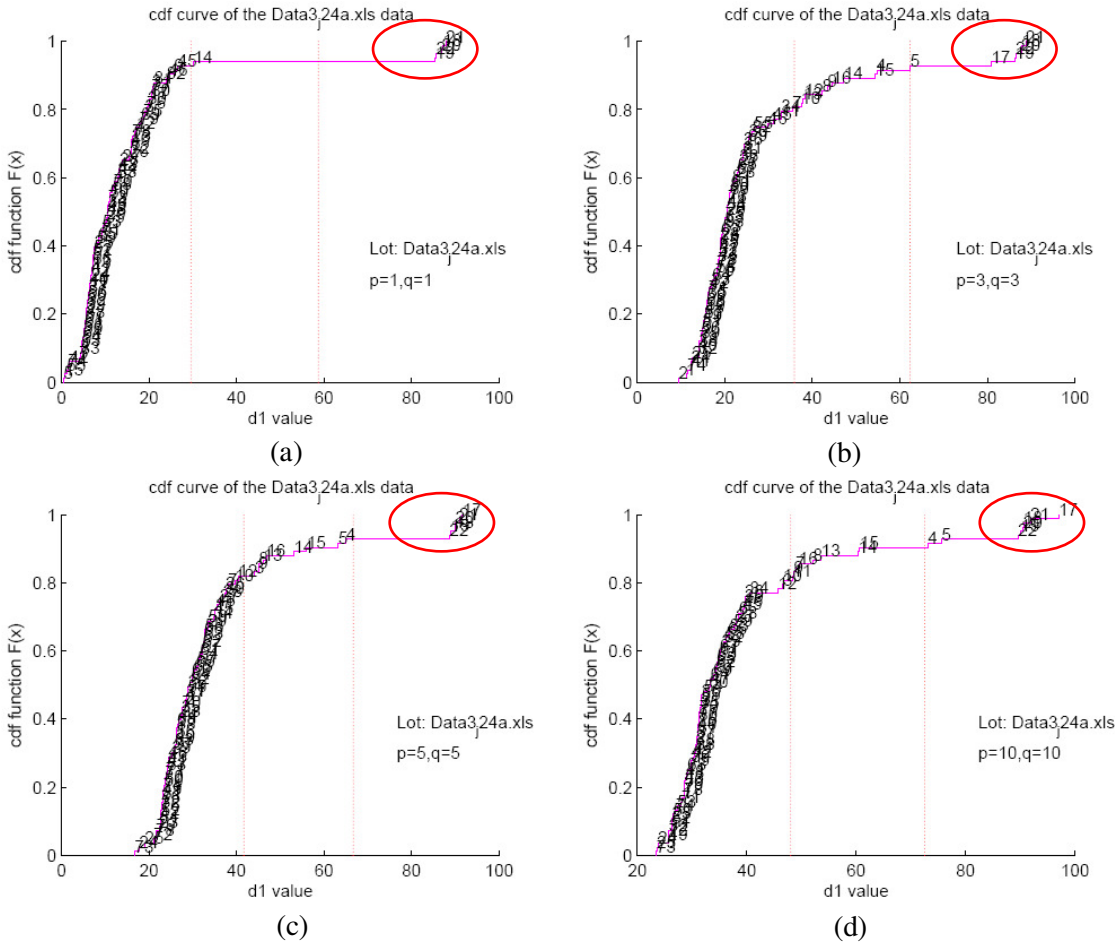


Figure 5- 6. CDF plot to all boardruns in d_1 scale with (a) Only the first PC is used (b) Only the first three PCs are used (c) Only the first five PCs are used (d) Only the first ten PCs are used

The boardruns circled are 17, 18, 19, 20, 21 and 22.

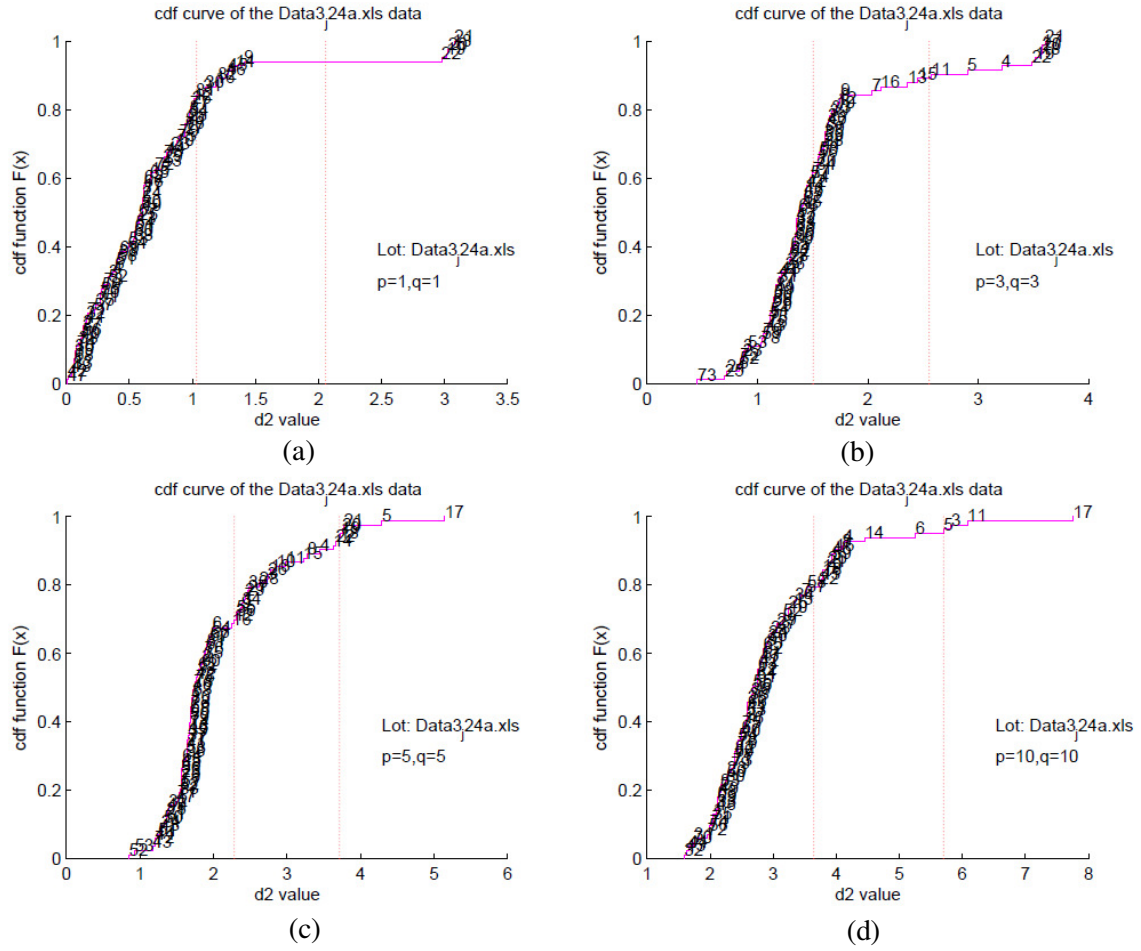


Figure 5- 7. CDF plot to all boardruns in d_2 scale with (a) Only the first PC is used (b) Only the first three PCs are used (c) Only the first five PCs are used (d) Only the first ten PCs are used

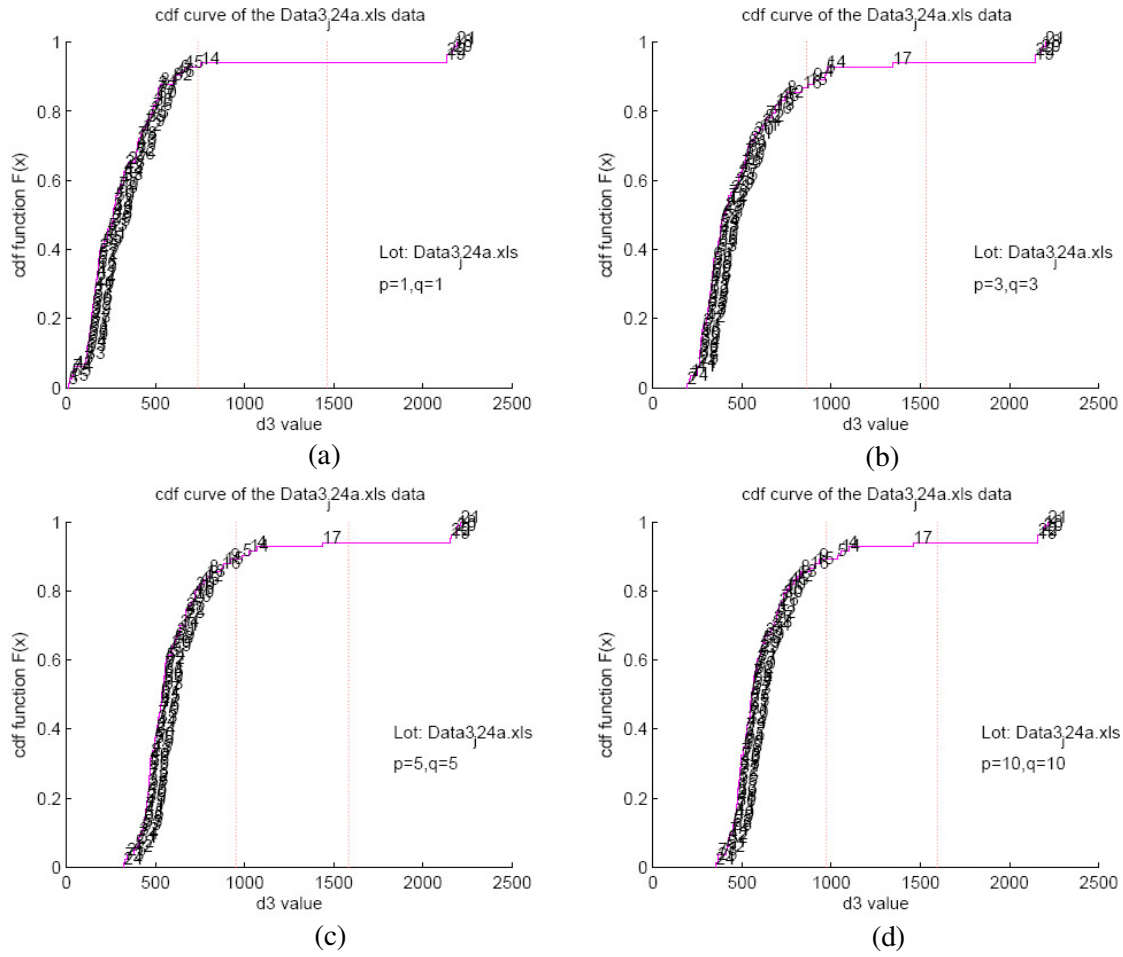


Figure 5- 8. CDF plot to all boardruns in d_3 scale with (a) Only the first PC is used (b) Only the first three PCs are used (c) Only the first five PCs are used (d) Only the first ten PCs are used

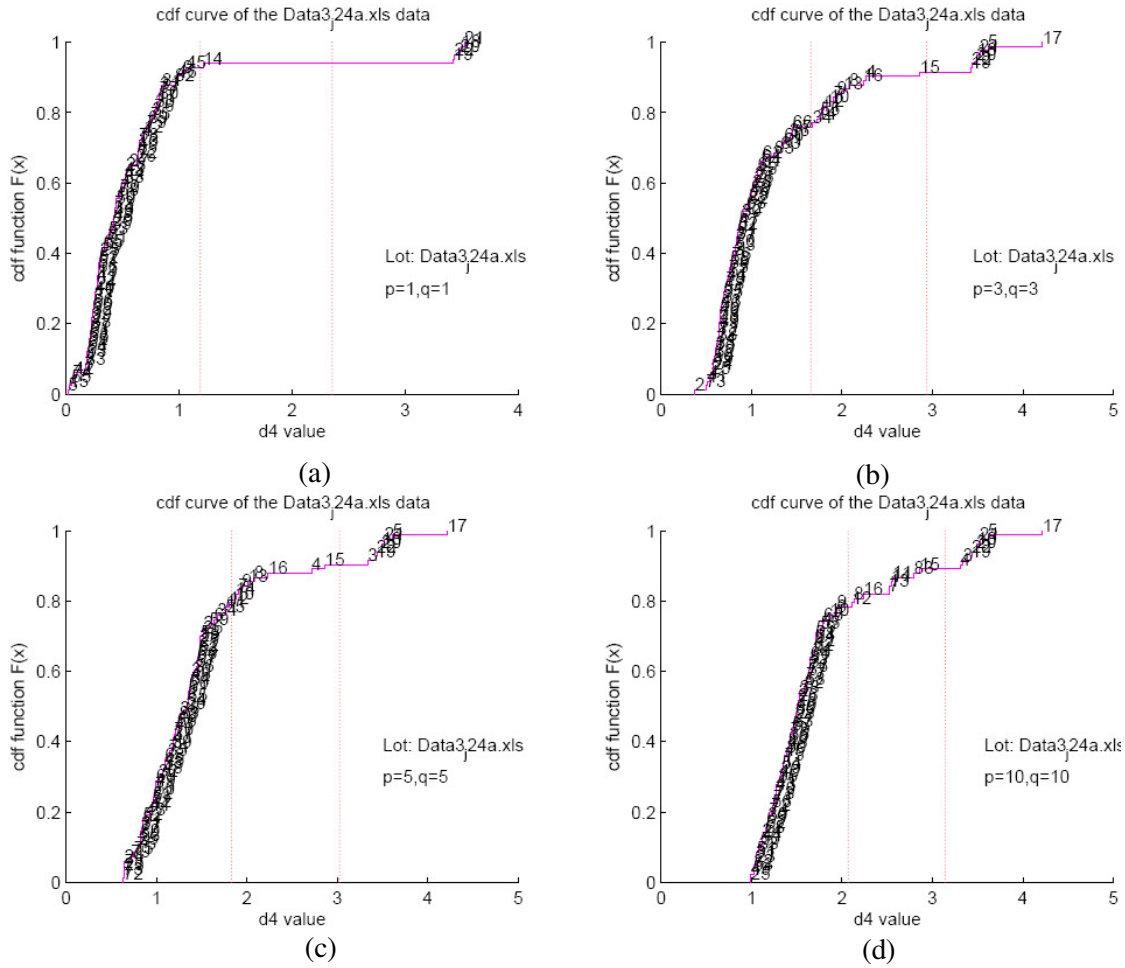


Figure 5- 9. CDF plot to all boardruns in d_4 scale with (a) Only the first PC is used (b) Only the first three PCs are used (c) Only the first five PCs are used (d) Only the first ten PCs are used

According to the sorting results in Figure 5-8, all of the plots using d_3 give very similar and matched results. This test statistic was also discussed in [11].

5.2.2. '*d*' Statistics with Least Significant PCs

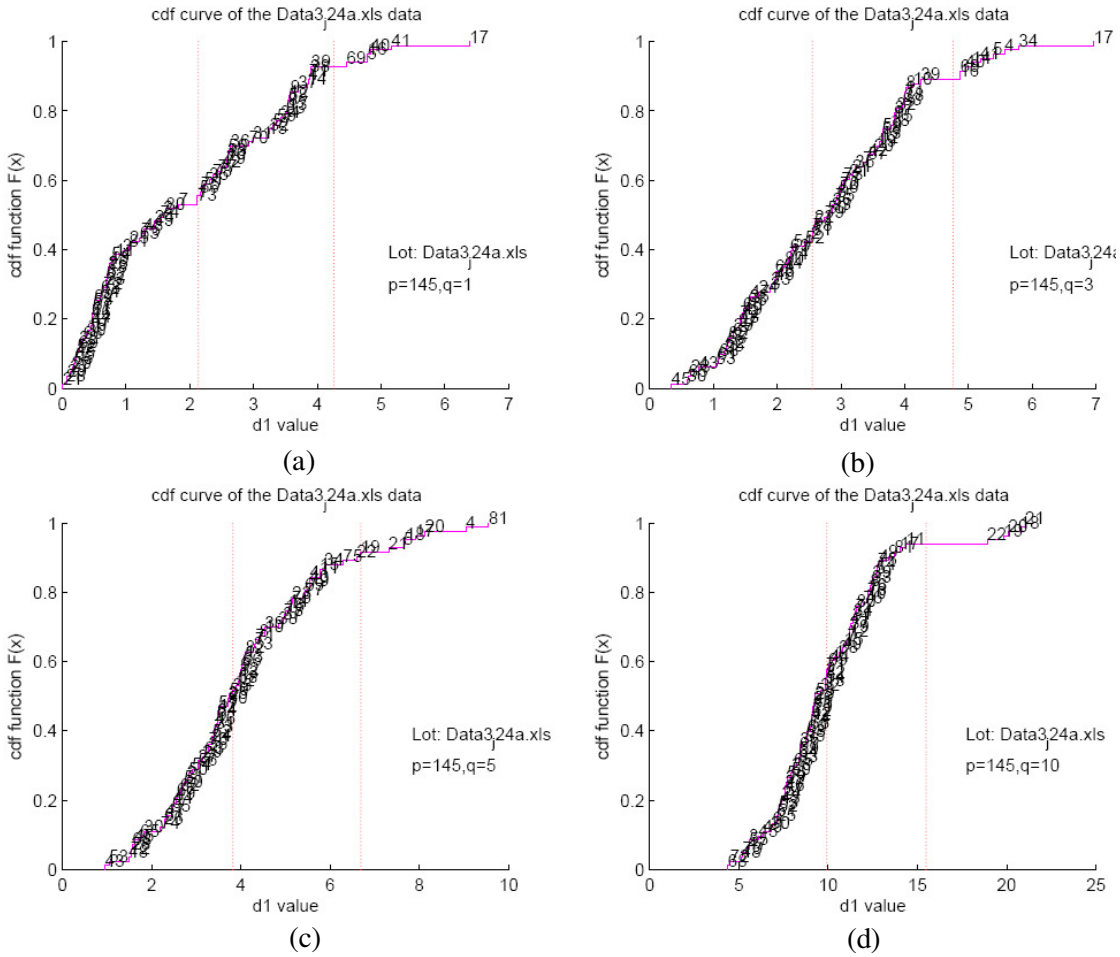


Figure 5- 10. CDF plot to all boardruns in d_1 scale with (a) Only the last one PC is used (b) Only the last three PCs are used (c) Only the last five PCs are used (d) Only the last five PCs are used

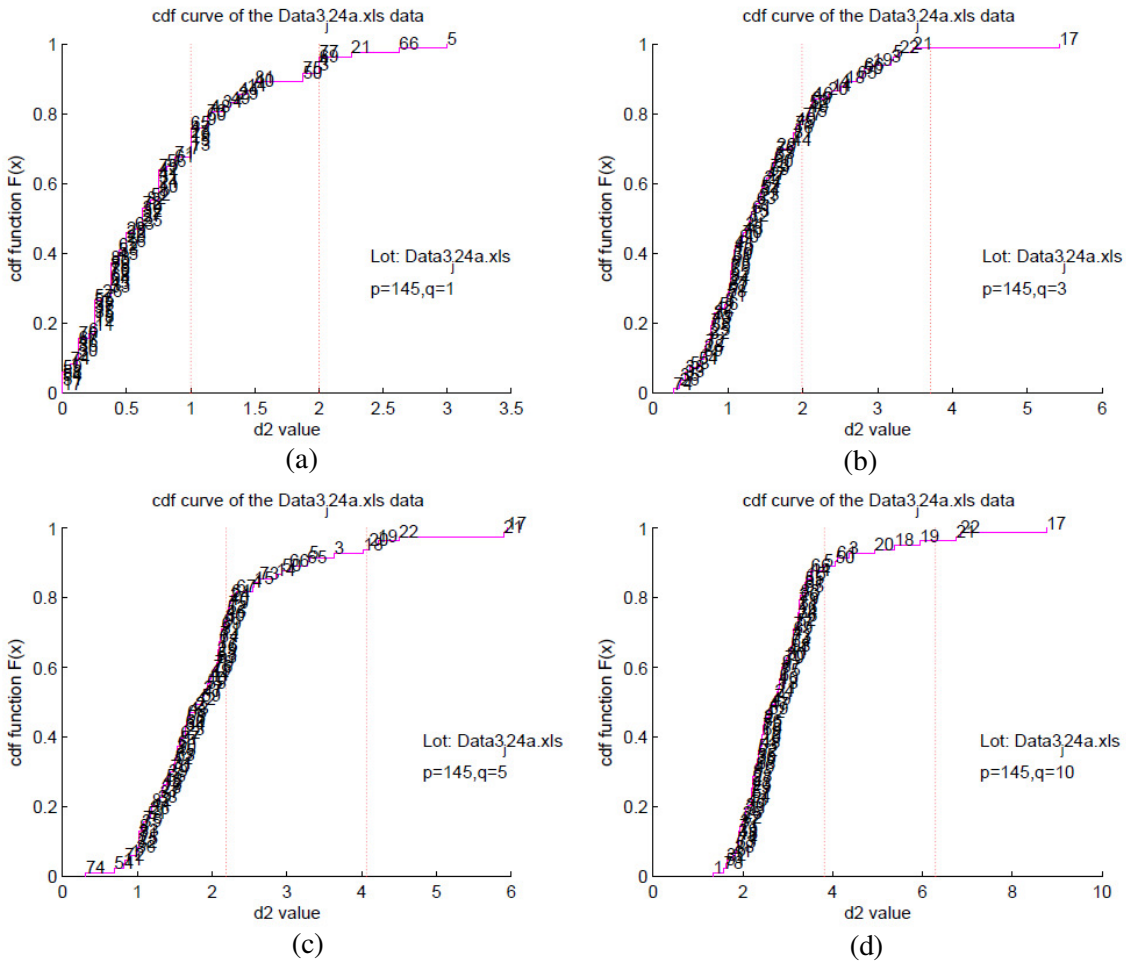


Figure 5- 11. CDF plot to all boardruns in d_2 scale with (a) Only the last one PC is used (b) Only the last three PCs are used (c) Only the last five PCs are used (d) Only the last five PCs are used

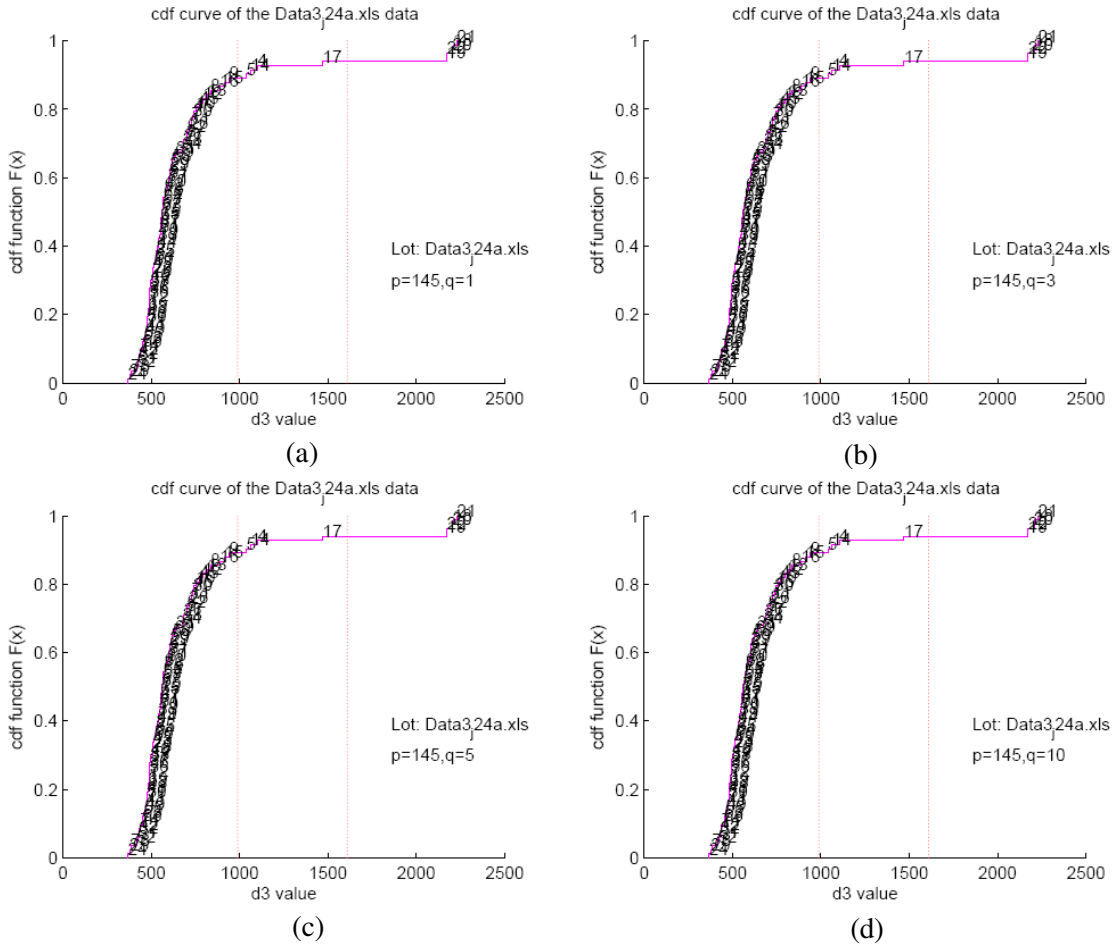


Figure 5- 12. CDF plot to all boardruns in d_3 scale with (a) Only the last one PC is used (b) Only the last three PCs are used (c) Only the last five PCs are used (d) Only the last five PCs are used

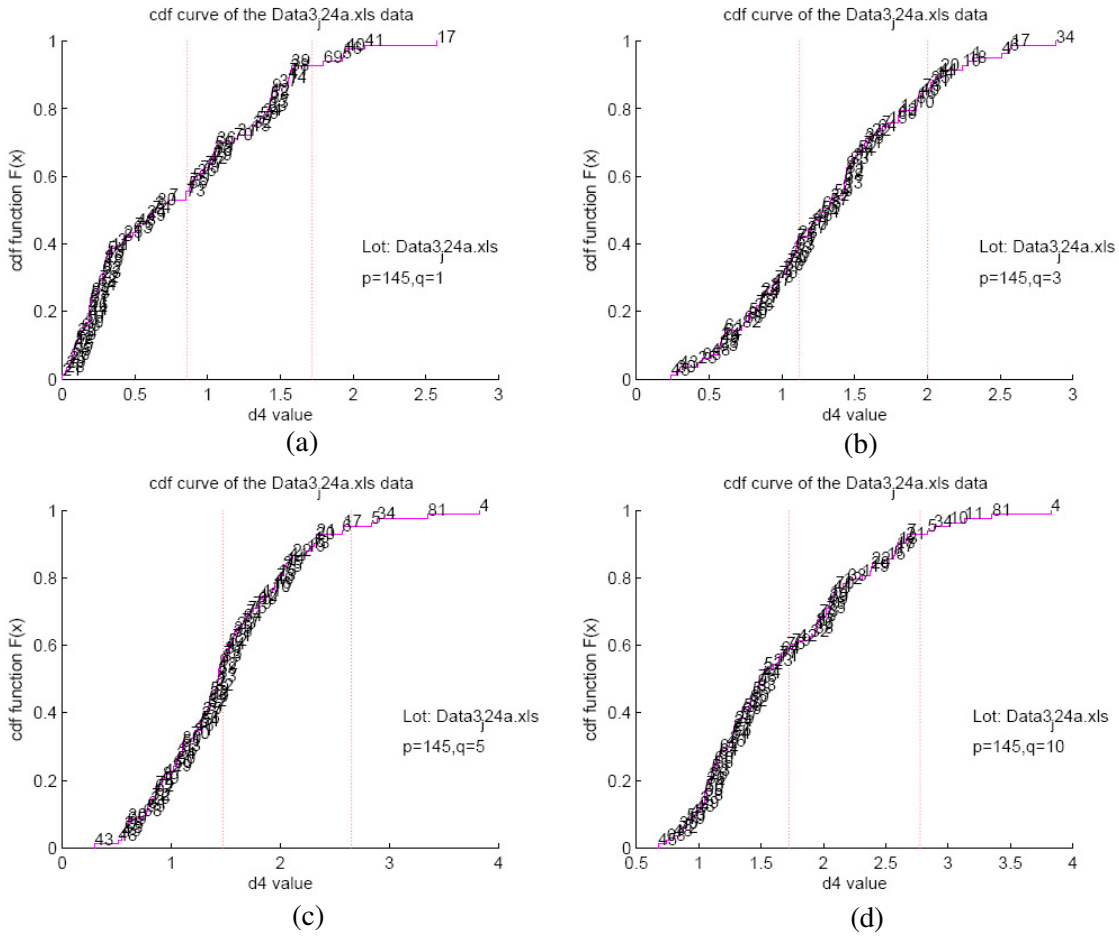


Figure 5- 13. CDF plot to all boardruns in d_4 scale with (a) Only the last one PC is used (b) Only the last three PCs are used (c) Only the last five PCs are used (d) Only the last five PCs are used

The analyses in this section utilized the last (least significant) 1, 3, 5, 10 PCs for the test statistics comparison. With d_1 as scale, the first plot in Figure 5-10 shows boardrun 17 at the high end. The last one shows boardruns 18, 19 20, 21, 22 at the high end but without 17. The sorting result in the third plot is unacceptable.

In d_2 scale plot of Figure 5-11, boardrun 17 stands at the highest end of the first plot, but other boardruns following it are not the outliers identified by inspection. Boardrun 4

shows highest d_4 value when using the last 3, 5, 10 PCs. The sorting results in d_4 scale are similar to those in d_2 scale.

All of the results above imply that the PCB test affects the more significant PCs a lot. Literatures such as [4] show d_3 is not stable enough to detect outliers. Rather than using d_3, d_1 will be the focus of analyses in this thesis.

5.2.3. Test Statistics and Parameter Selection

As discussed in Section 5.2.1 and Section 5.2.2 above, we find that the boardrun numbers standing at the high end of d_1 scale and d_3 scale CDF curves with the most significant PCs give relatively consistent results, which match the inspection results. Majority of the sorting results in d_2 and d_4 scale shows un-matched plot. For example, the potential outlier boardruns after the break do not standout clearly.

Extensive experiments presented above indicate that the most significant PCs provide good test results in PCB testing. Figure 5-3 shows the variance of different principal components, which indicates that the first 15 principle components carry almost all the variance information of the dataset. Regardless of the type of PCB test dataset, the more significant components appear to be more useful for detection of outliers, which can be supported by other analyses of this thesis. The set of PCs that are suitable for defect detection is problem specific. For example, in case of IDDQ measurement data, less significant principal components are more useful in detecting the outliers [4] [11].

Figure 5-14 shows the CDF plot in d_1 scale using the 5 most significant PCs for all the boardruns in Data_j24. To avoid clutter, we have labeled only a select set of boardruns in the CDF. The detailed sequence of sorted boardrun numbers are provided

below the figure. The plot of the CDF shown in Figure 5-14 clearly identifies the outlier boardruns, which appear at the right extreme separated from the remaining boardruns. Boardruns 17, 18, 19, 20, 21 and 22 all clearly stand out at the high end of the CDF curve where they indicate much larger d_1 values than the other boardruns. The boardruns 18, 19, 20, 21 and 22 are five repeated tests of the same board. In the plot, the six boardruns show a clear difference from others which means that they are far different from others based on the holistic view provided by the PCA analysis. With the first few PCs used, the outliers detected in the d_1 value of CDF plot match the observations from the raw data plot. With another clear break at the high tail of CDF plot, some other boardruns like 5, 4 are also identified as possible outliers depending on the degree of filtering desired. A careful inspection of the right end of Figure 5-1 indicates these two boardruns to show abnormal value.

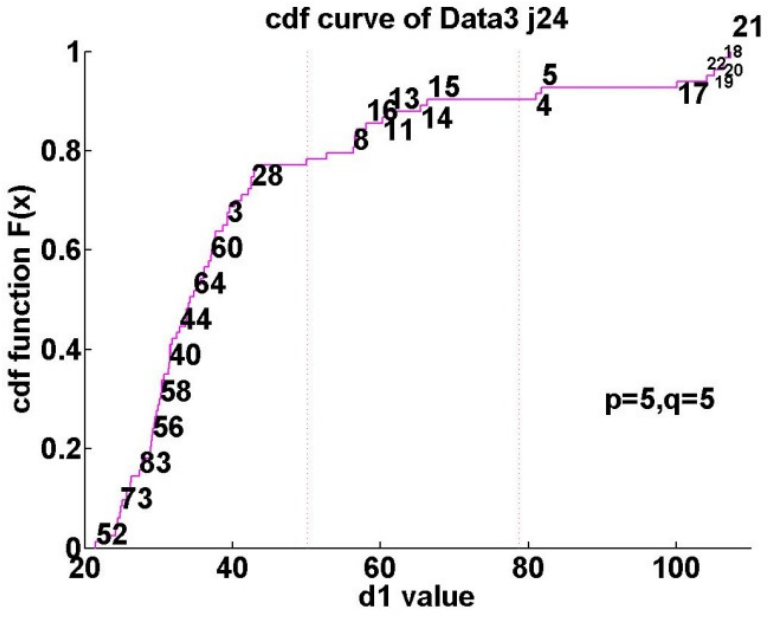


Figure 5- 14. More detailed CDF plot to Data3_j24 in d_1 scale with the 1st 5 PCs

The boards number from left to right:

52,51,50,53,49,32,**73**,24,48,25,74,72,**83**,71,57,47,1,42,**56**,70,6,68,38,37,**58**,43,59,41,39,55,**40**,2,23,78,33,35,**44**,69,79,54,75,36,**64**,80,76,31,77,65,**60**,29,81,63,61,62,**3**,67,66,82,27,45,**28**,46,26,30,34,12,**8**,7,10,9,**16,11,13,14,15,4,5,17,22,19,18,20,21**

In Figure 5-15, with the last three PCs used for d_1, outliers standing on the right part of the break are subset of the outliers and potential outliers in Figure 5-14. Those following them are not the outliers. If the analysis was used, we cannot detect all the outlier in the test. The order of clear outliers and potential outliers will also be mixed. Since similar or worse results also occur in other analysis with least significant PCs, the *d* value calculated with least significant PCs is not suitable for PCB testing here.

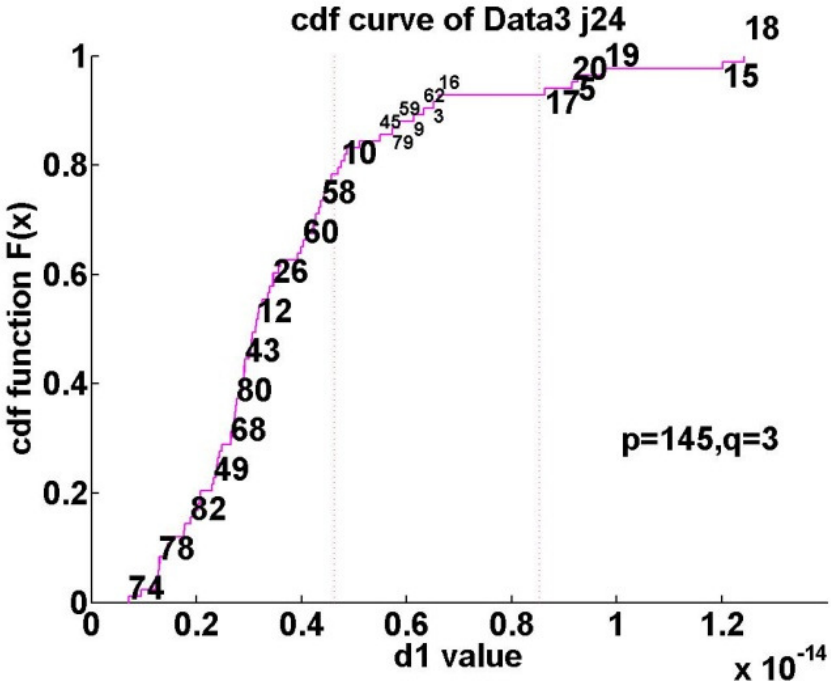


Figure 5- 15. More detailed CDF plot to Data3_j24 in d1 scale with the last 5 PCs

In Figure 5-3, we saw that the variance of Z score will decrease a lot after the Z score column number (PC number) is greater than 10. Then numbers of PC vectors selected should be smaller than 10. However, to include more information, a number of PCs larger

than three should be considered. From experience p=q=3 to p=q=7 with d_1 statistics effectively detect the outliers.

5.3. Further Test with Multiple Connector Measurements

To verify the conclusion above, similar tests are applied to a more comprehensive dataset Data_D1. In contrast to Data3_j24, which is for a board with a single connector, the Data_D1 with 20 boardruns is for a board with 8 tested connectors, which leads to over one thousand tested pins per board. Compared with measurement matrix of Data3_j24 of dimension 145 × 83, the dimension of matrix of Data_D1 is 1053 × 20.

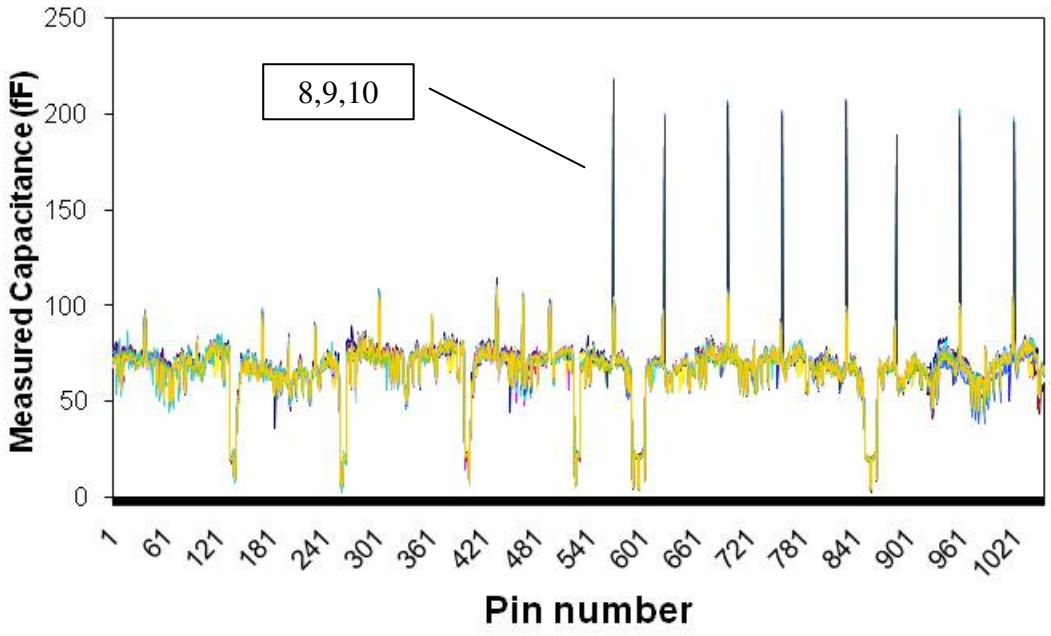


Figure 5- 16. Plot of raw measurements of Data D1

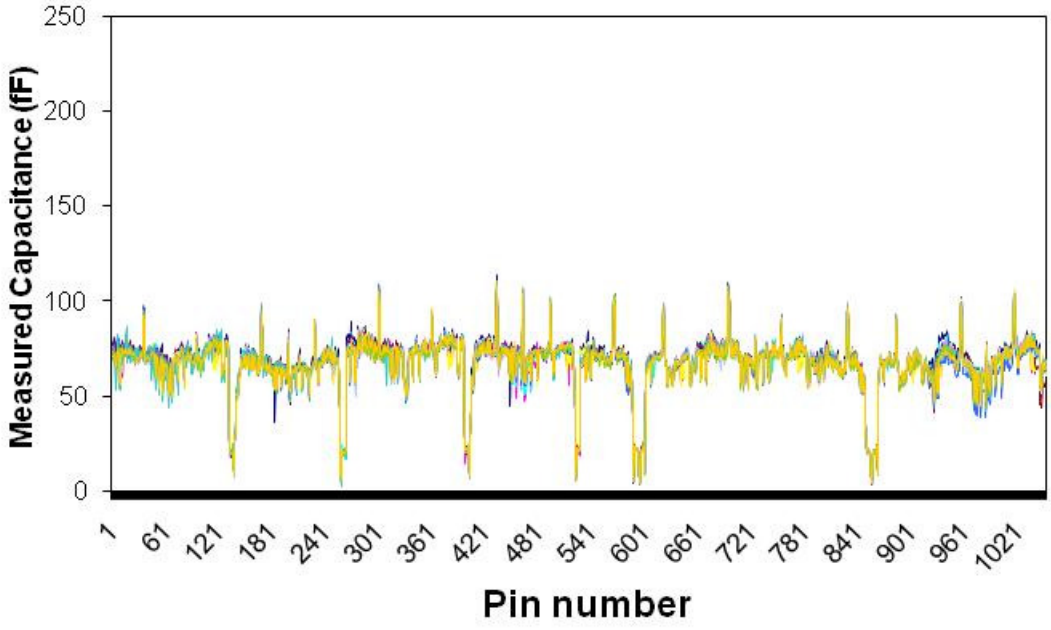


Figure 5- 17. Plot of raw measurements of Data D1 without clear outliers

Inspections to Figure 5-16 and Figure 5-17 show that boardruns 8, 9 and 10 give unusual measurements at some of the 1053 pins, which appear on the right half of Figure 5-16. These clear outlier measurements make them clearly different from other boardruns. Results consistent with this can be obtained from

Figure 5-18 and Figure 5-19 which use the first few PCs as axes.

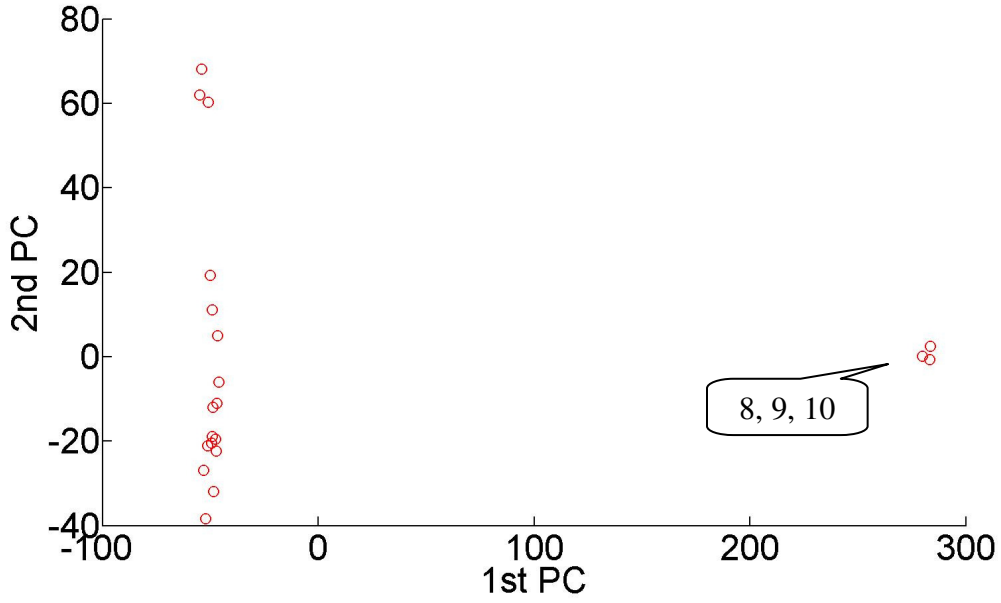


Figure 5- 18. Boardruns plot for Data D1 with the first two PCs as axis

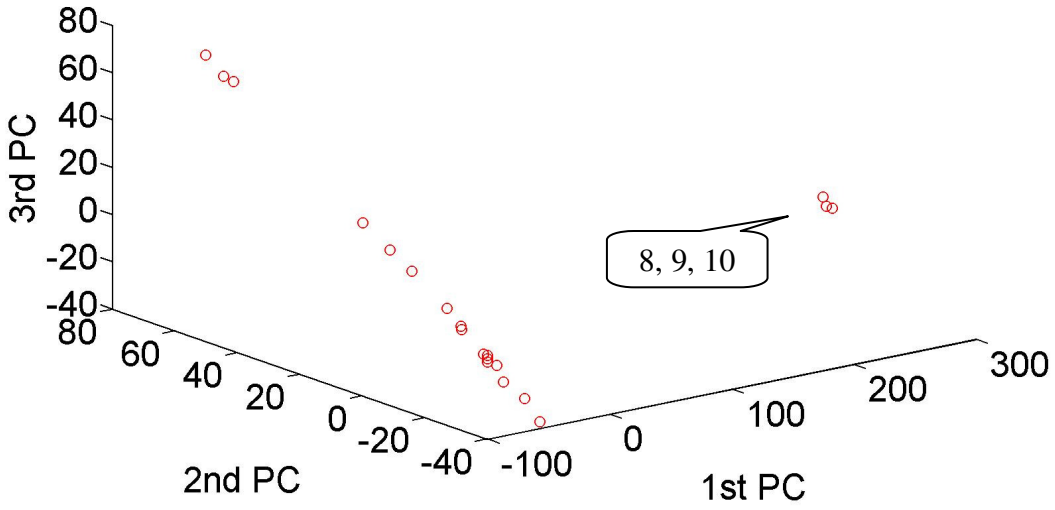


Figure 5- 19. Boardruns plot for Data D1 with the first three PCs as axis

5.3.1. *'d'* Value Test with Most Significant PCs

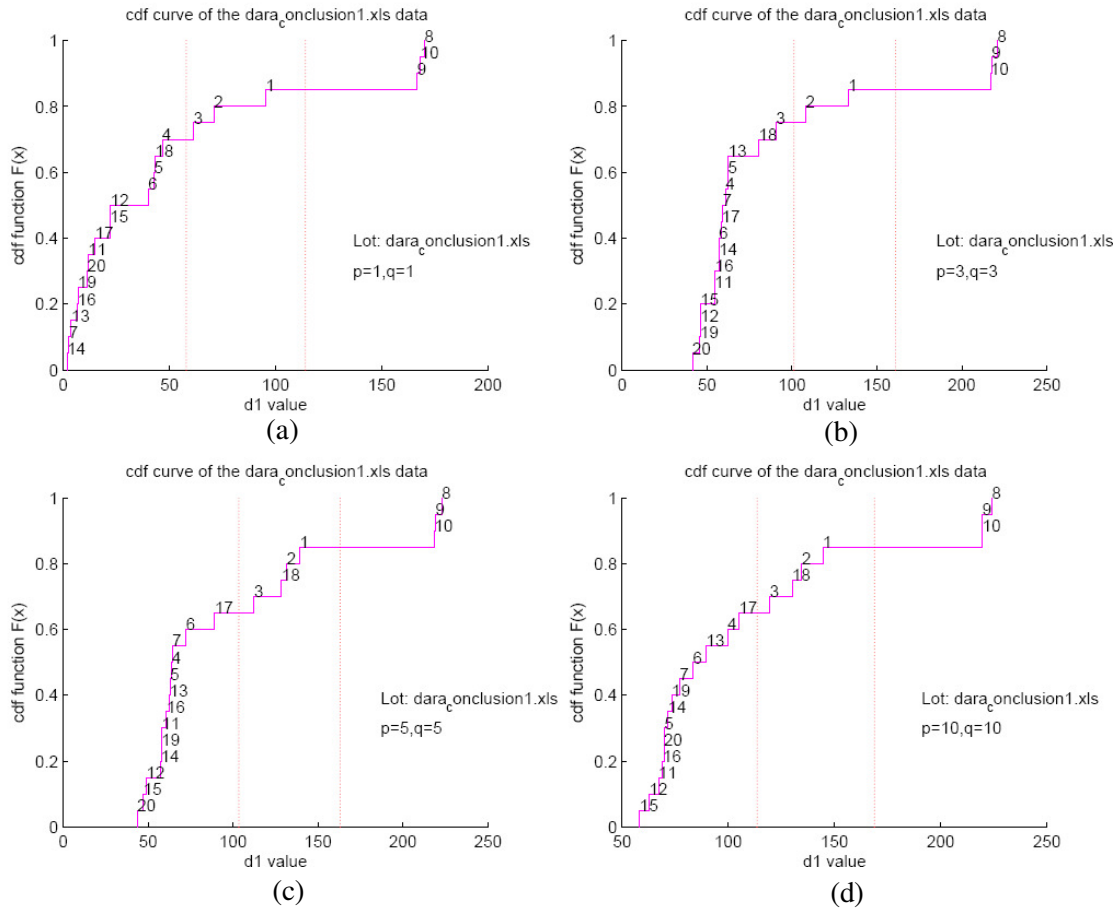


Figure 5- 20. CDF plot to all boardruns in d_1 scale with (a) Only the first PC is used (b) Only the first three PCs are used (c) Only the first five PCs are used (d) Only the first ten PCs are used

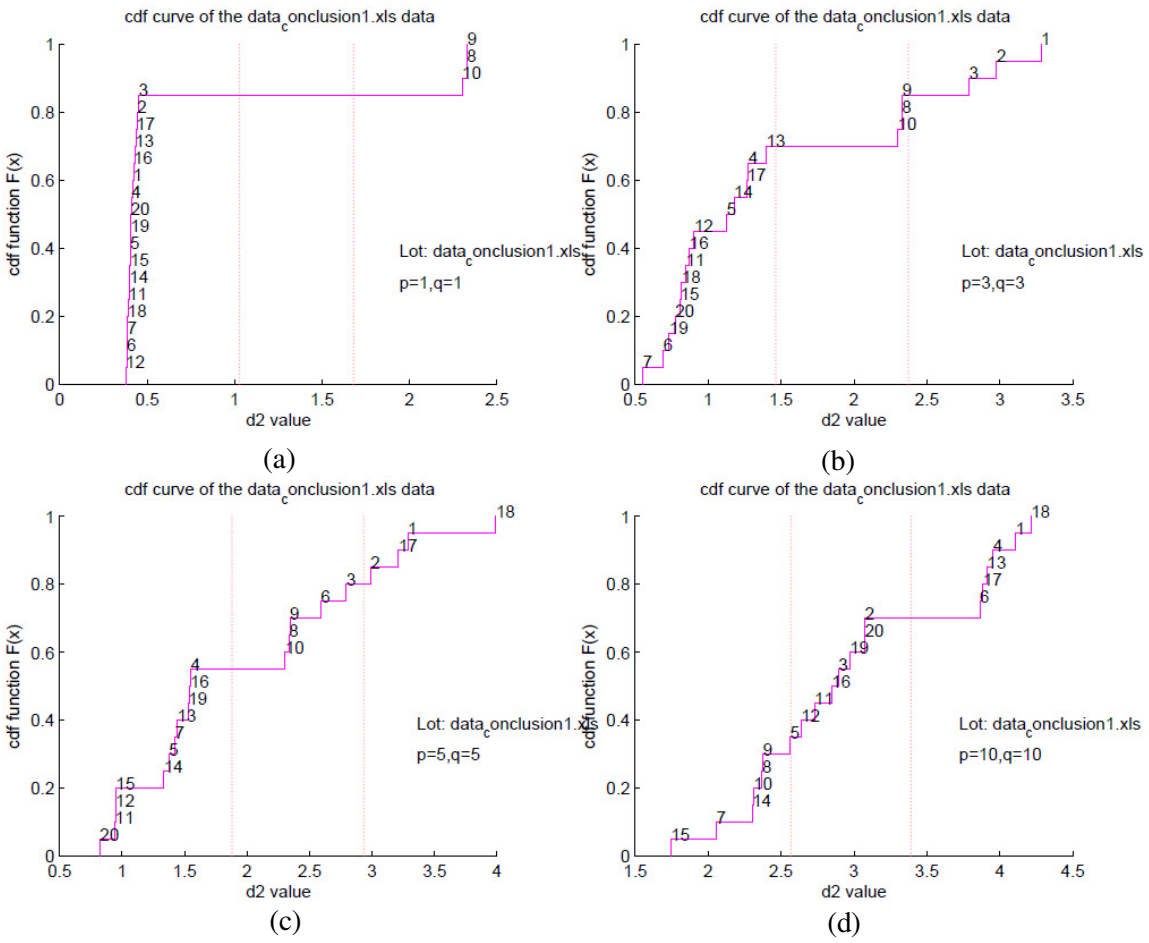


Figure 5- 21. CDF plot to all boardruns in d_2 scale with (a) Only the first PC is used (b) Only the first three PCs are used (c) Only the first five PCs are used (d) Only the first ten PCs are used

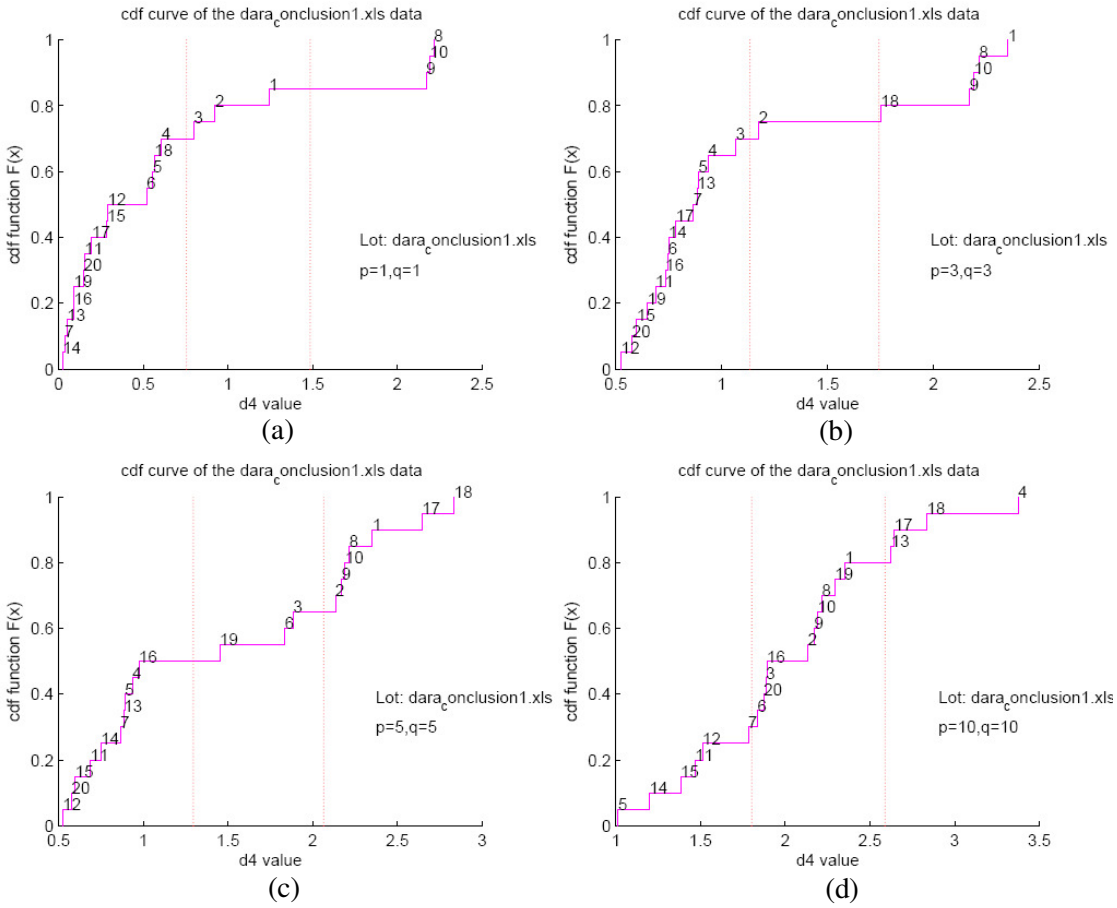


Figure 5- 22. CDF plot to all boardruns in d_4 scale with (a) Only the first PC is used (b) Only the first three PCs are used (c) Only the first five PCs are used (d) Only the first ten PCs are used

5.3.2. *'d'* Value Test with Least Significant PCs

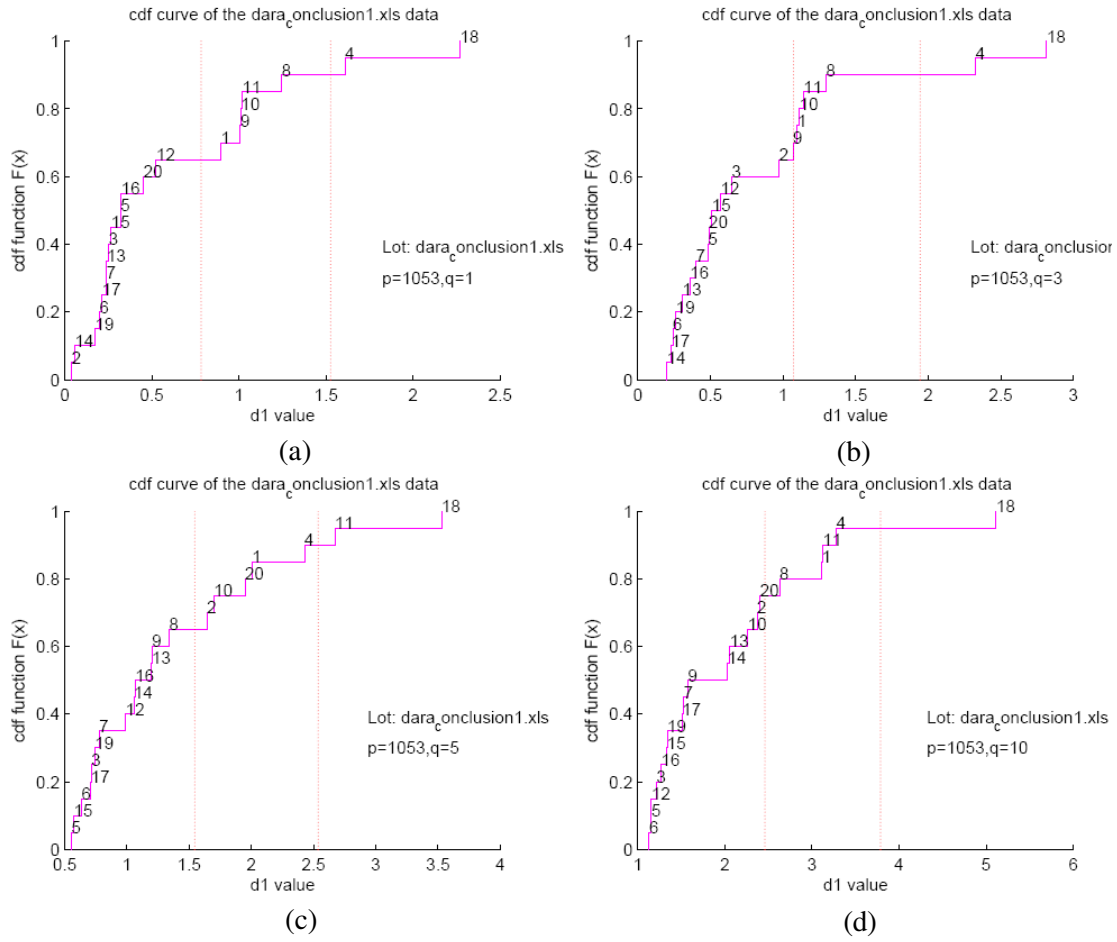


Figure 5- 23. CDF plot to all boardruns in d_1 scale with (a) Only the last one PC is used (b) Only the last three PCs are used (c) Only the last five PCs are used (d) Only the last five PCs are used

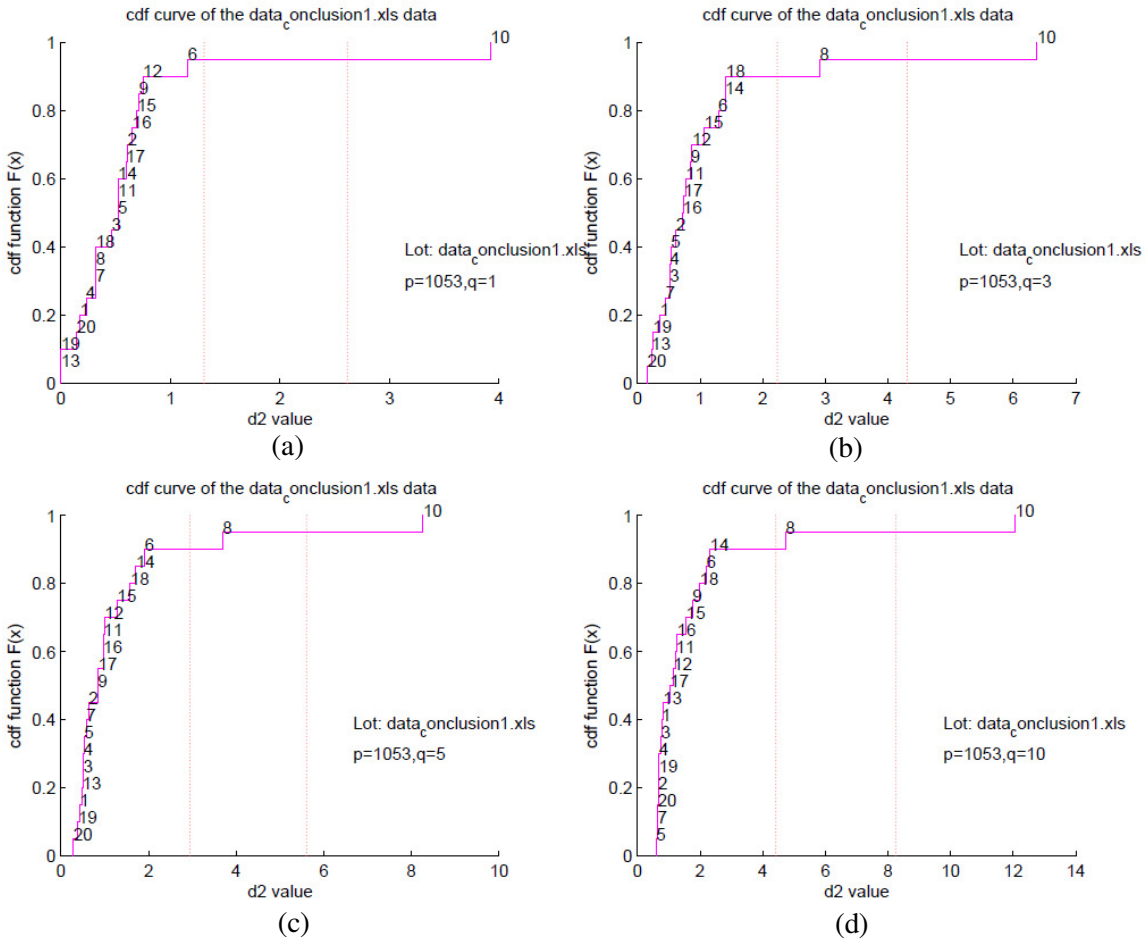


Figure 5- 24. CDF plot to all boardruns in d_2 scale with (a) Only the last one PC is used (b) Only the last three PCs are used (c) Only the last five PCs are used (d) Only the last five PCs are used

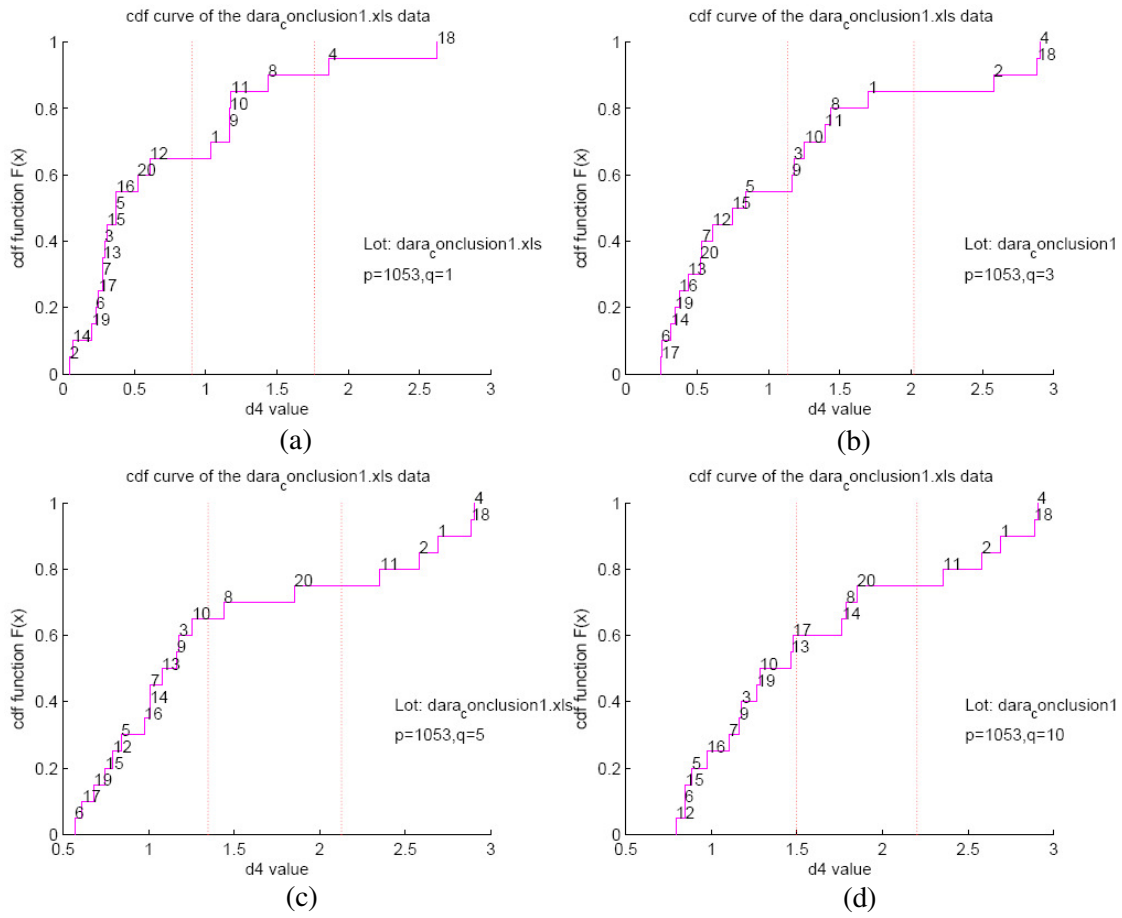


Figure 5- 25. CDF plot to all boardruns in d_4 scale with (a) Only the last one PC is used (b) Only the last three PCs are used (c) Only the last five PCs are used (d) Only the last five PCs are used

From what has been discussed above, the three boardruns 8, 9, 10 are clearly different from others with inspection. However, not all CDF plots in Section 5.2.2 show them as outliers. When using most significant PCs in the calculation, CDF plot in d_1 scale clearly shows the three outliers at the right end, which coincide the three boardruns with abnormal measurements. The sorting in d_1, d_2, d_4 scale with least significant PCs cannot show any of the three boardruns 8, 9 and 10 at the CDF high tail. The d_1 values with the first few PCs can detect clear outliers or abnormal boardruns effectively.

Chapter 6.
Principal Component Analysis to Agilent Datasets with Test Statistics d_1

With most significant PCs, the d_1 test statistics selected in Chapter 5 is effective in detecting outliers. This chapter applies analysis with d_1 statistics to other Agilent datasets to further evaluate test effectiveness.

The analysis is first applied to dataset D0 which includes measurements of connectors j24, j25, j27, j28 and j31. Then the three datasets Data3_j25, Data3_j27, Data3_j28 from D3 are analyzed one by one.

Data j24 includes fifteen j24 connector measurements where connector j24 has 145 tested signal pins. Figure 6-1 below shows measurements at each of the tested pins with fifteen boardruns. Based on the inspection to the 15 measurement curves, all of them have relatively close measurement values.

However, there is no denying the fact that the measurements are not identical. We can still find difference among boardruns 6, 31, 3, 15, 12 and 14 in the red oval area.

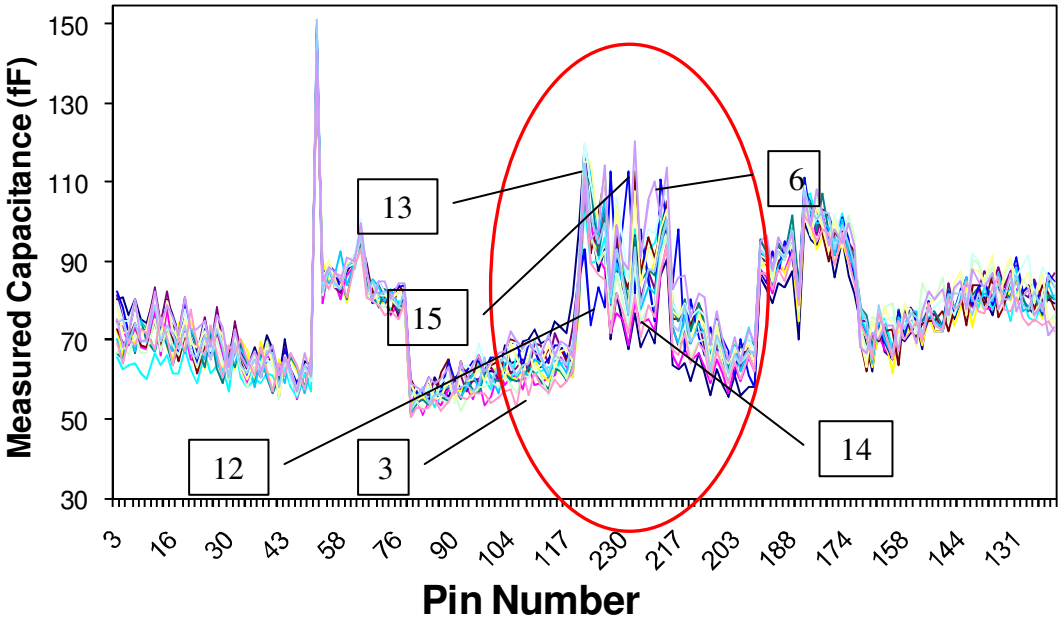


Figure 6-1. Raw measurements plot of j24

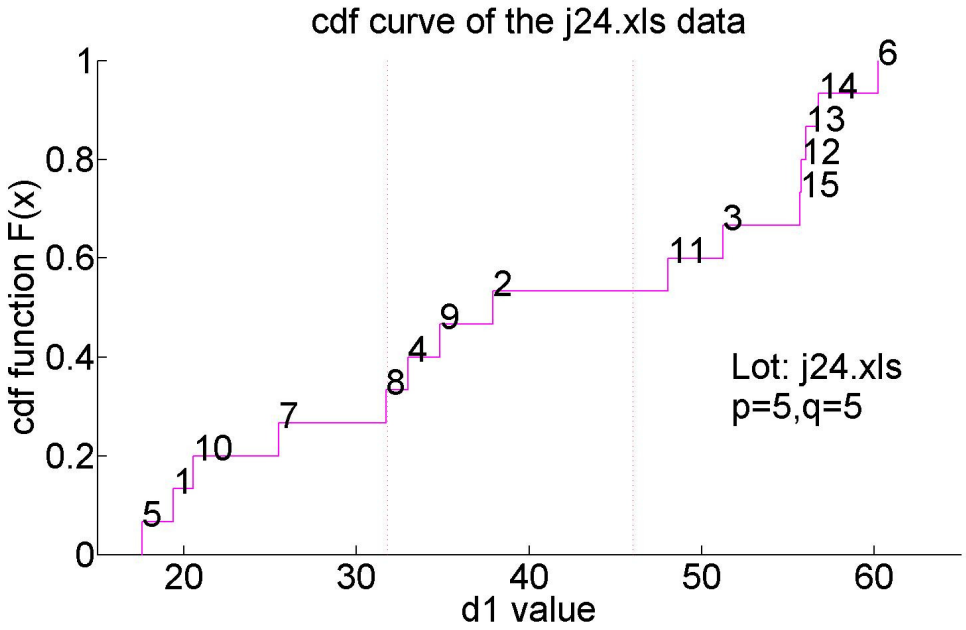


Figure 6-2. CDF plot to j24 in d_1 scale

The 15 boardruns in data j25 are from j25 connectors, each of which contains 150 tested pins. The measurement curves show more close measurements than those in the j24 data.

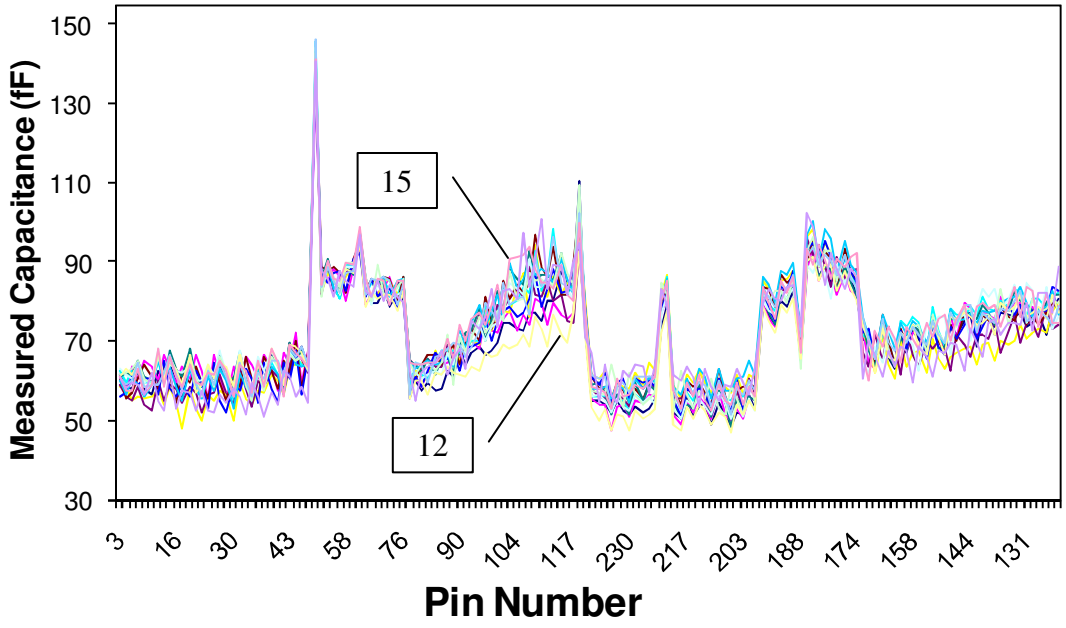


Figure 6- 3. Raw measurements plot of j25

From the inspection to the j25 measurement plot in Figure 6-3, boardruns 12 and 15 (the purple one) show relatively detectable difference from others at multiple pins. The CDF curves give test result matching inspection clearly in Figure 6-4. The two boardruns are provided with the highest d_1 values. Because of the relatively close measurements in both dataset, outlier criteria should be made by specialist. Those with highest d_1 value may be ruled out as outliers. Outlier criteria will be further discussed in 'Local Analysis' section of Chapter 8.

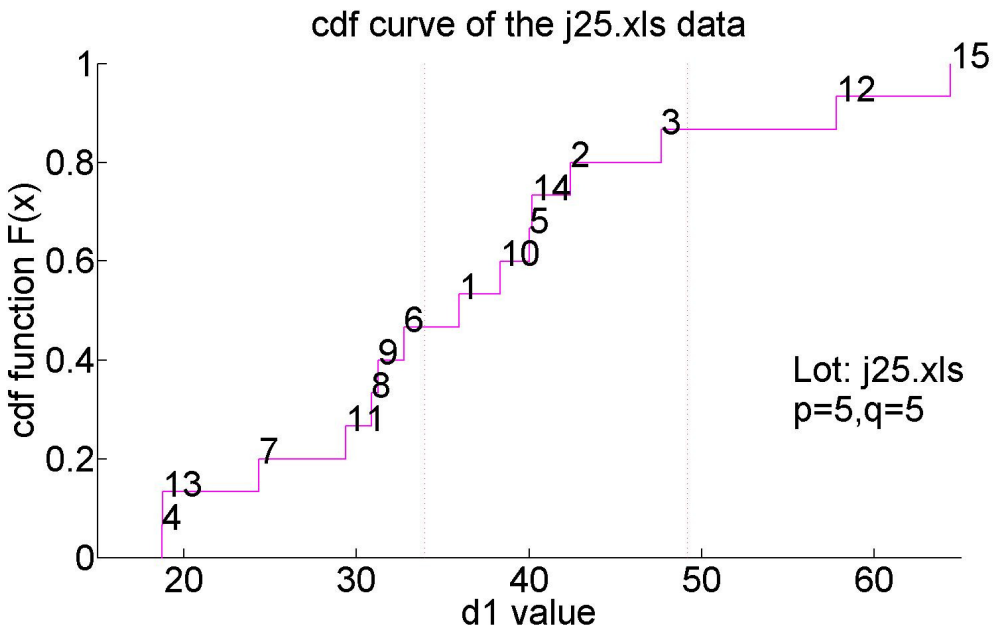


Figure 6- 4. CDF plot to j25 in d_1 scale

Compared with dataset j24 and j25, j27 is one containing clear outliers. In Figure 6-5, each of the fifteen j27 boardruns show measurements of 147 tested pins. The pink curve, which is boardrun 2, shows a difference from others like horizontal shift. Boardrun 3, yellow curve, shows much higher and lower values in the middle part of the plots. Following the two boardruns, boardrun 1, which is the dark blue curve, also shows clear peaks. If all of these factors are taken into account, the 3 boardruns can be viewed as outliers, which are also clearly shown in the CDF plot in Figure 6-6.

Since the overall variance plays essential roles in d_1 calculation, the datasets containing higher variance (or possible outliers) will show higher d_1 scale range. That can be shown from the maximum d_1 values of dataset j24, j25 and j27 in CDF plots.

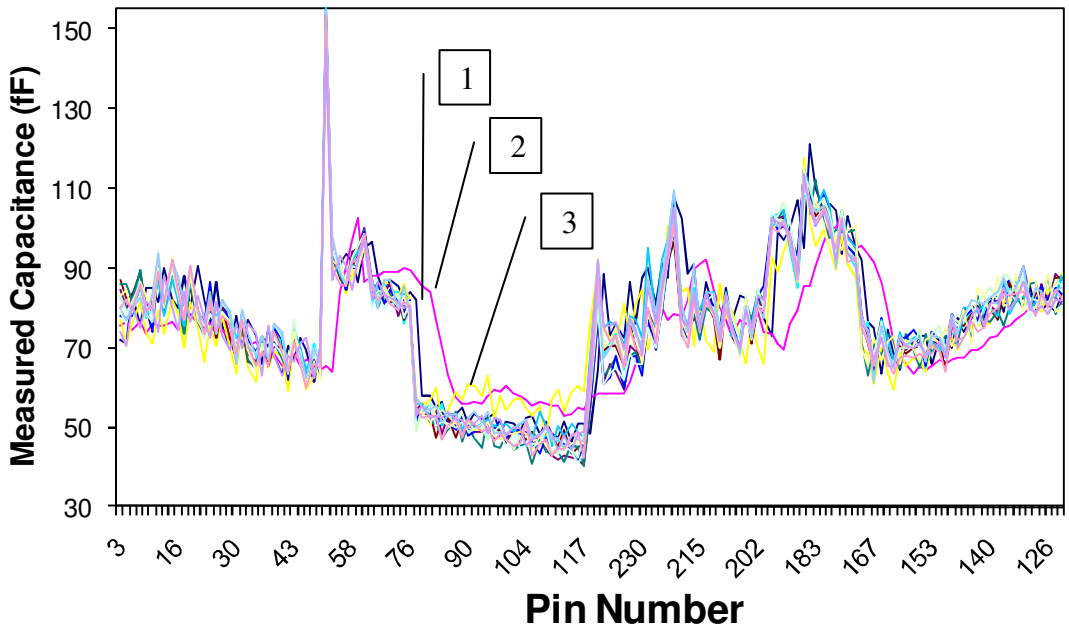


Figure 6- 5. Raw measurement plot of j27

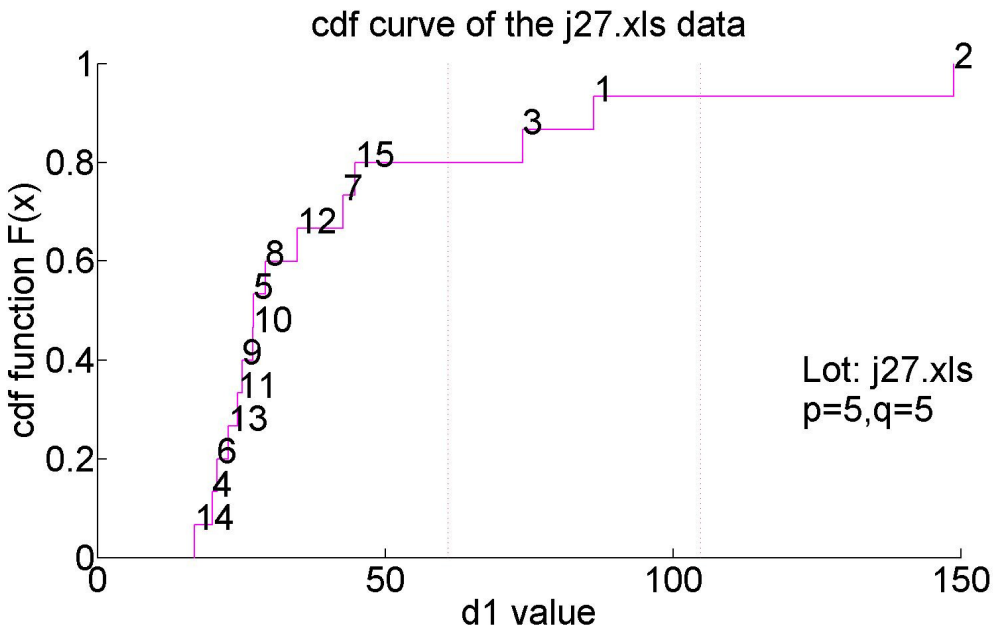


Figure 6- 6. CDF plot to j27 in d_1 scale

Dataset j28 is one with 15 boardruns and 151 tested pins each. Like measurements of dataset j24, most abnormal peaks appear in the middle part of the figure.

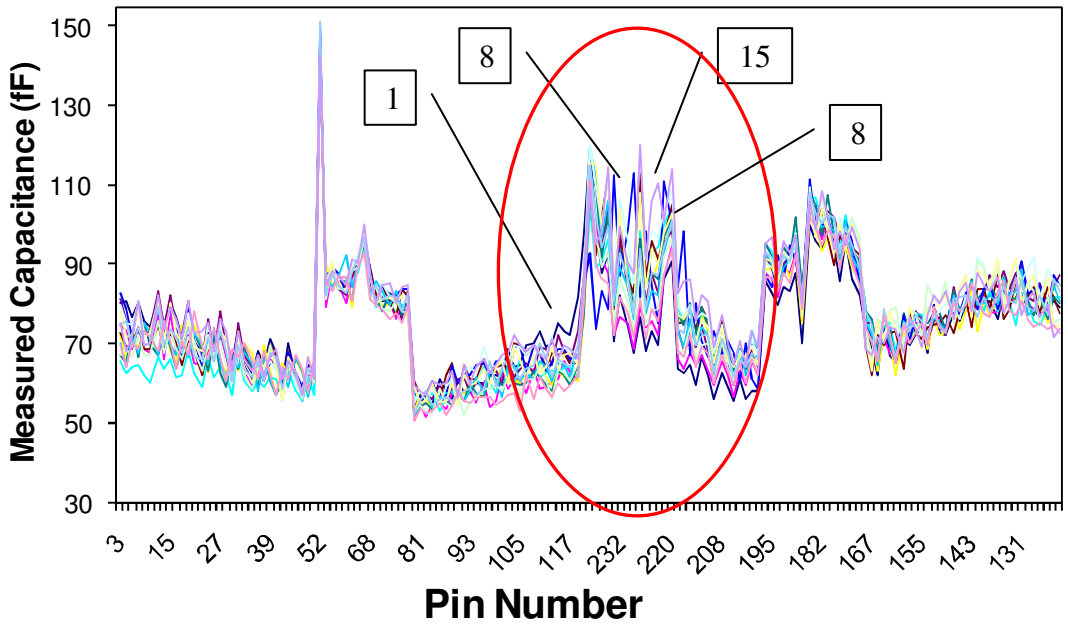


Figure 6- 7. Raw measurements plot of j28

As be inspected from Figure 6-7, boardruns 8, 1 and 15 all show clear peaks in the circled area. Figure 6-8 provides the test result with CDF curve in d_1 scale.

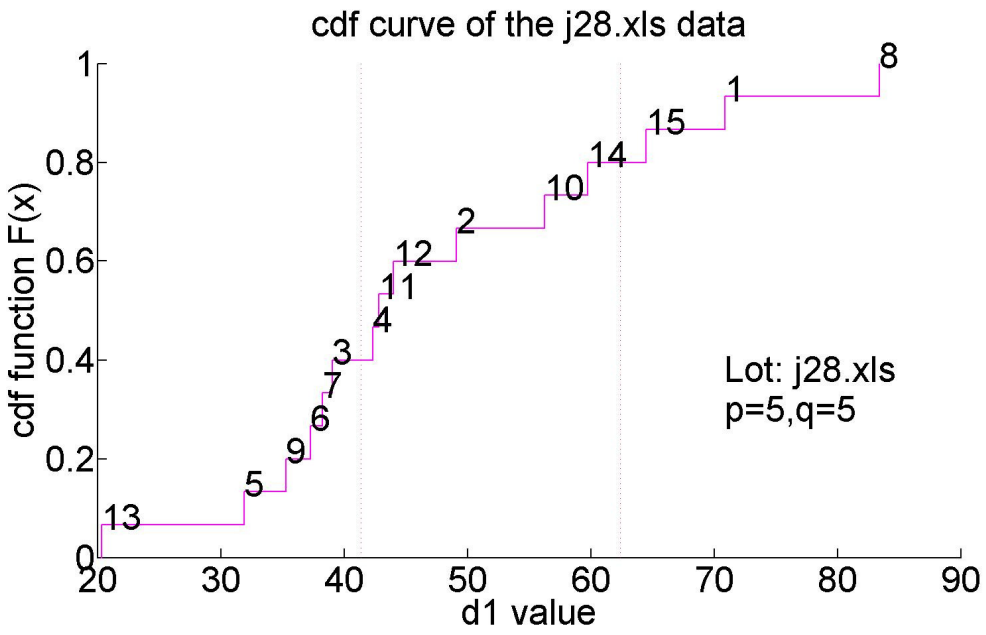


Figure 6- 8. CDF plot to j28 in d_1 scale

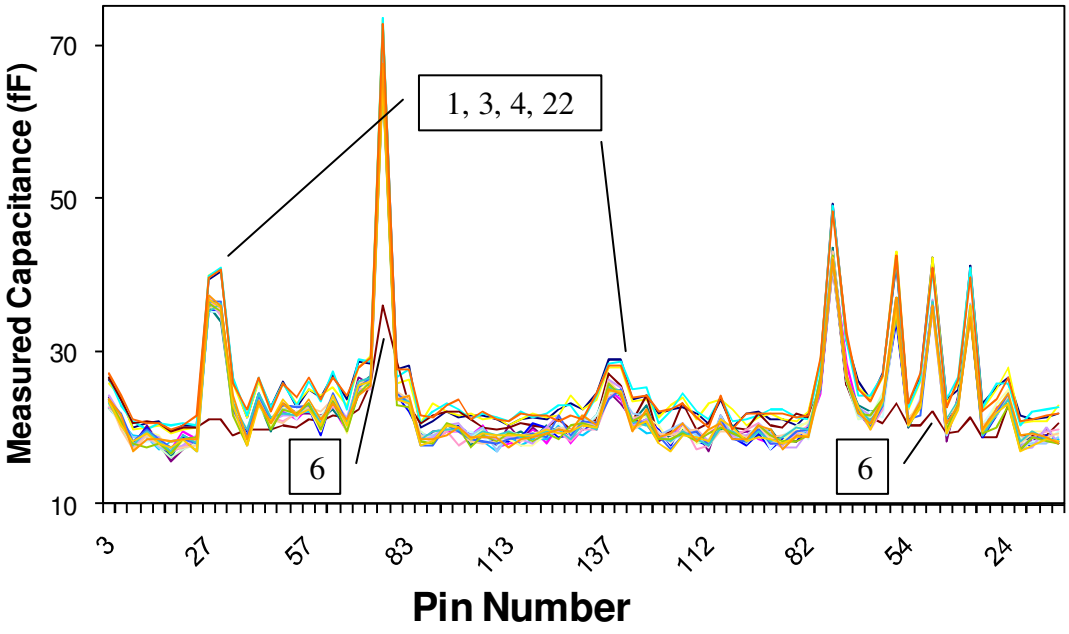


Figure 6- 9. Raw measurements plot of j31

Dataset j31 corresponds to a j31 standard connector on a set of 22 boardruns. Visual inspection of Figure 6-9, which shows the row measurement values in dataset j31, identifies the boardruns 1, 3, 4 and 22 as clear outliers.

Next we use the first 5 Principal Components (PCs) to calculate the d_1 value. Figure 6-10 shows the CDF plot for d_1 for all of the boardruns in dataset j31. Boardruns 6, 1, 3, 4 and 22 all clearly stand out at the high end of CDF curve where they indicate much larger d_1 values than the other boardruns. Remaining 17 boardruns are clustered together on the left side of the plot. The five boardruns show a clear break from others which means that they are far different from others based on the holistic PCA analysis. This technique can effectively filter the abnormal boardruns.

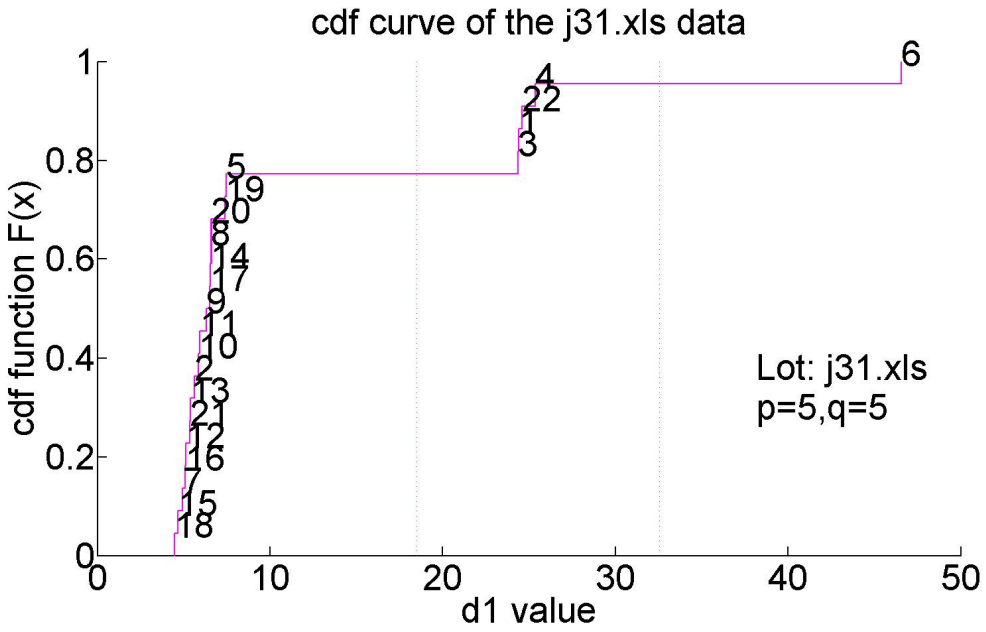


Figure 6- 10. CDF plot to j31 in d_1 scale

In Figure 6-10, the clear outliers in this dataset didn't give as high d_1 values as that of j27. The reason is partially because of the tested pin for each connector is only 77

which are much lower than the 151 in j27. When needed, the comparison should be made between datasets with same devices and pins.

In Data3, there are three datasets: Data3_j24, Data3_j25, Data3_j27 and Data3_j28, which are measurements to the respective connectors. All the three datasets have 83 boardruns. As stated in the previous analysis to Data3_j24, some boardruns in D0 are in fact repeated tests to the same board. There are 47 unique boards in each dataset. To avoid the repeated illustration to Data3_j24 (analyzed in Chapter 5), analysis will begin from Data3_j25.

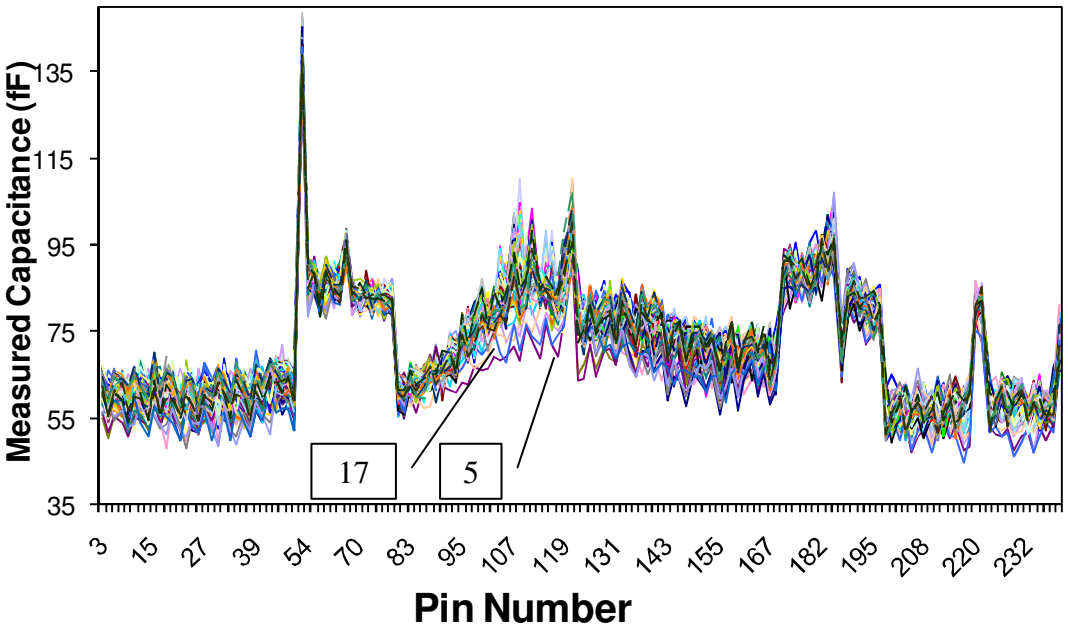


Figure 6- 11. Raw measurements plot of Data3_j25

Each boardrun in Data3_j25 below has 150 tested pins. Compared with Data3_j24, the difference among measurements is very small. The blue curve is boardrun 5, and the violet one is boardrun17. The boardruns measurement difference from others, lead them

to different from others in CDF plot of Figure 6-12. The maximum d_1 value in CDF plot is smaller than 80, which shows that the measurement difference inside of the dataset is much smaller than that in Data3_j24.

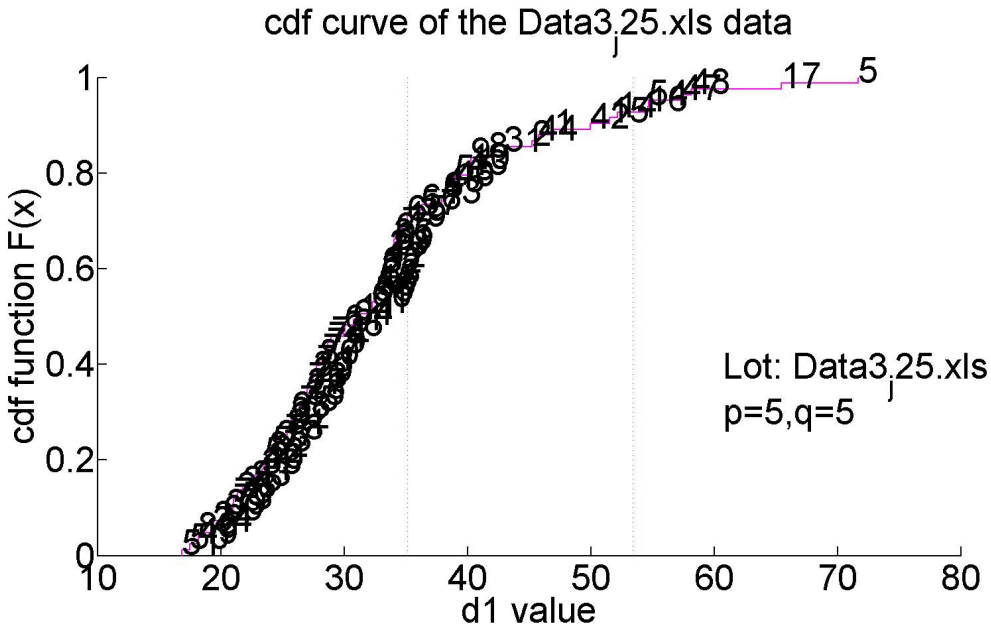


Figure 6- 12. CDF plot to j31 in d_1 scale

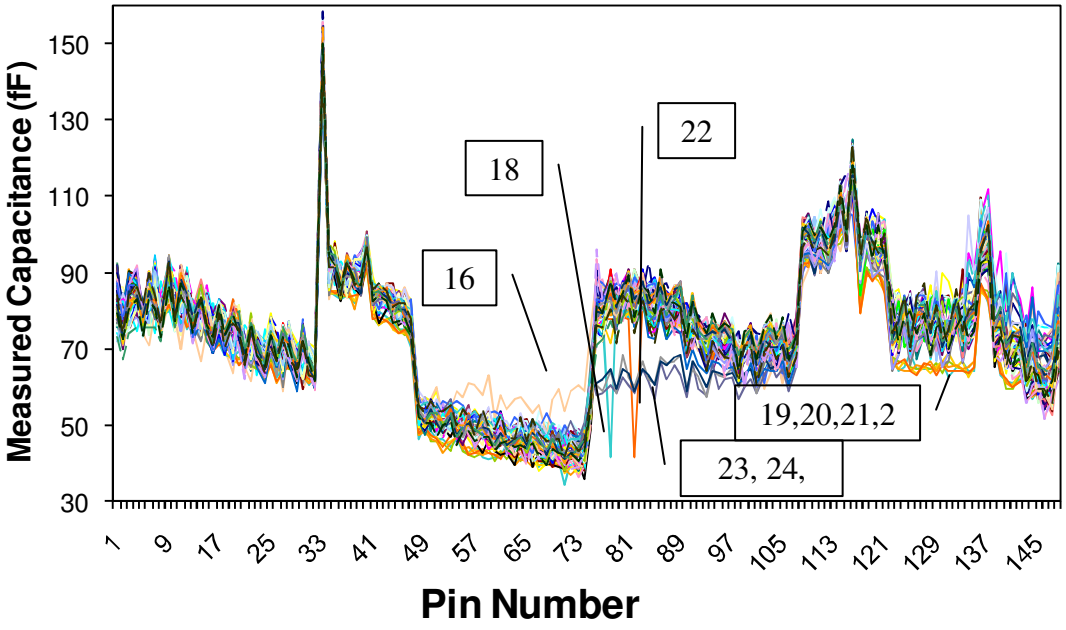


Figure 6- 13. Raw measurements plot of j27

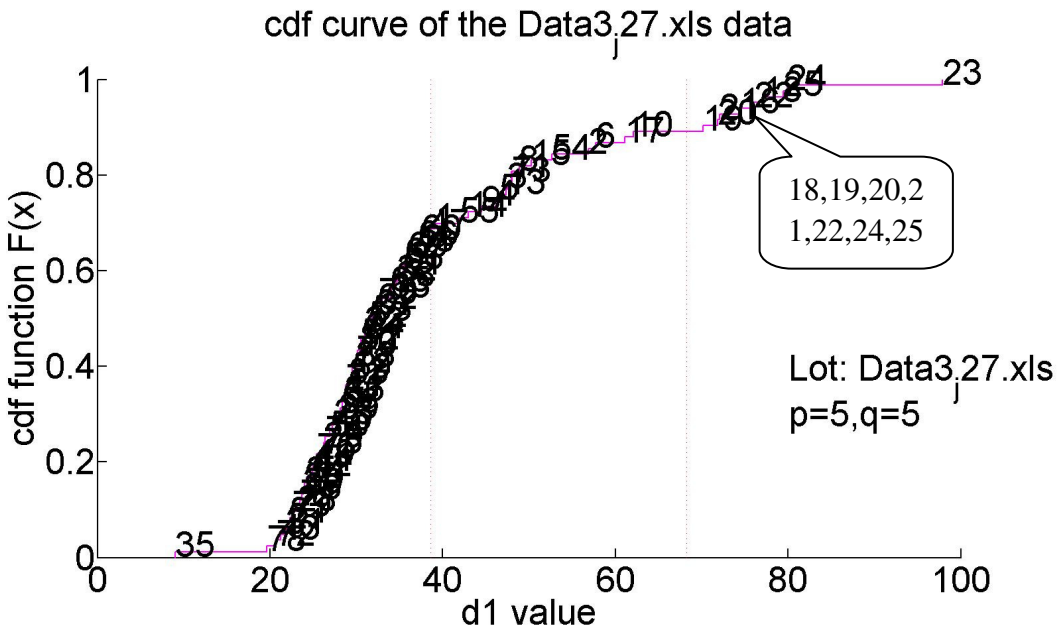


Figure 6- 14. CDF plot to Data3_j27 in d_1 scale

There are 148 tested pins on each of the 83 boardruns in Data3_j27 shown in Figure 6-13. In fact, d_1 values of boardrun 18, 19, 20, 21, 22, 23, 24, 25 are relatively higher

than those of other boardruns as shown in Figure 6-14. This difference can also be inspected from Figure 6-13.

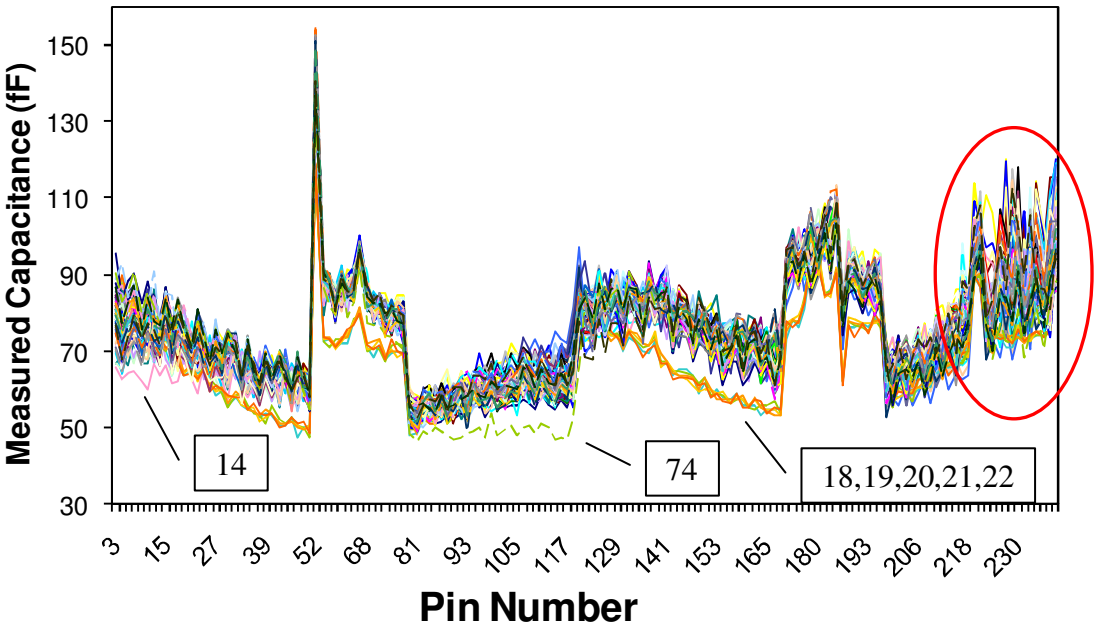


Figure 6- 15. Plot of raw measurements of j28

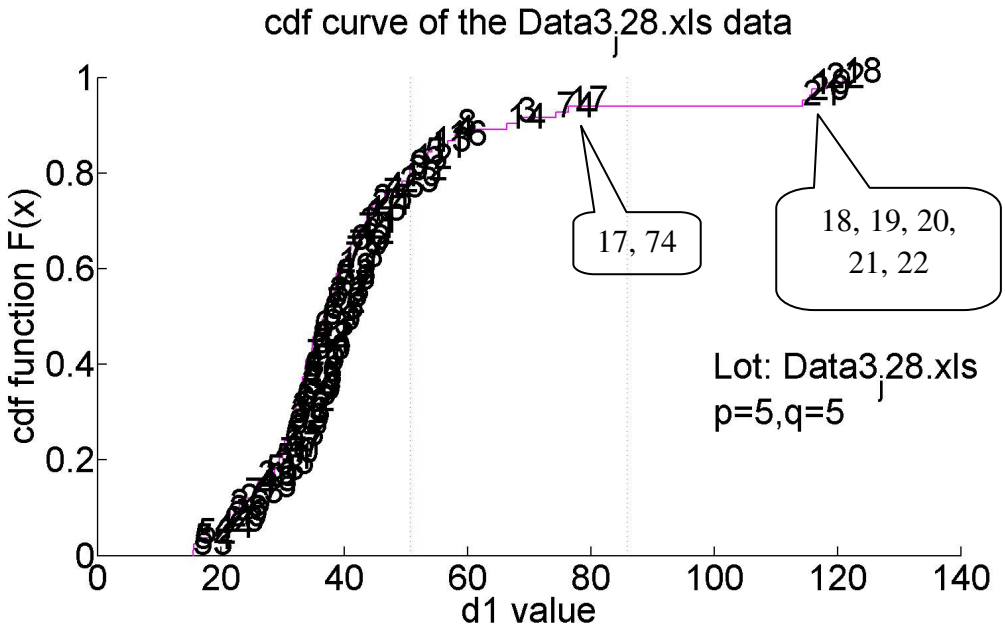


Figure 6- 16. CDF plot to Data3_j28 in d_1 scale

Similar phenomena are inspected in the Data3_j28. The measurement values of the 151 tested pins of each boardrun are plotted in Figure 6-15. Right part of Figure 6-15 shows lots of peaks that cannot be inspected clearly. However, measurements of boardruns 18, 19, 20, 21, 22 show very clear gap from others at almost half of the tested pins. That also explains their high d_1 values in Figure 6-16. These boardruns can be regarded as outliers. In fact, measurement 74 also shows very clear gaps in the middle part of the Figure 6-15. The reason why it doesn't show at the extreme part in Figure 6-16 is that our analysis is based on one test window holistic view analysis. For boardrun 74, since the pins of showing deviation are not as many as those of boardruns 18, 19, 20, 21, 22, its d_1 value will not so high. Rather than discussing it here, I will solve this kind of problem with localized analysis in later chapters.

The analyses in this chapter employ d_1 as the statistic to detect the outliers. The method effectively detects the difference among the measurements, which further confirm the conclusion drawn from Chapter 5.

Chapter 7.
Comparison between PCA and Traditional Method in PCB Outlier Detection

In the traditional approach, the standard deviation of measurement is used to identify the pins considered to be sufficiently different from the others. In the standard deviation method (STDev method), mean and the standard deviation are calculated using the data for all tested pins. The mean value serves as an expected good value. High and low limits are set equal to the average plus/minus α×STDev, for an appropriately selected α based on the degree of screening desired. For the Data3_j24 illustrated in Chapter 5, the boardruns that would be considered as abnormal based on this method are shown in Table 7-1 for different values of α. For example, the upper and lower limits for pin 1 of 83 boardruns are $Average_1 + \alpha \times STDev$ and $Average_1 - \alpha \times STDev$.

α	**Abnormal PCB Detected**	**Boards No.**
6	0	
5.5	1	17
5	1	17
4.5	2	14, 17
4	5	3,11,14,17,83
3.5	11	3,4,5,6,8,11,14,15,17,59,83
3	18	3,4,5,6,8,9,11,14,15,16,17,18,19,20,21,51,59,83
2.5	35	3,4,5,6,7,8,9,11,12,13,14,15,16,17,18,19,20,21,22,34, 36,47,48,51,53,57,58,59,60,63,68,73,80,81,83

Table 7- 1. Abnormal boards detected by STEDev method

With the increase of the value of α, fewer and fewer boardruns are included into the suspected group of boardruns. Since normally $\alpha \geq 4$ is often used as the threshold, when $\alpha = 4$, the STDev method detects boardruns 3, 11, 14, 17 and 83 as the abnormal ones. However, the PCA based outlier detection method identifies boardruns 18, 19, 20, 21, 22 and 17 as outliers. Note that with the PCA based outlier detection, the decision is significantly more justifiable as the CDF plot clearly separates the outliers. With STDev method, the value of α has to be guessed.

PCA methods	STED method	Tester's Selection
18,19,20,21,22,17	3,11,14,17,83	18,19,20,21,22,17

Table 7- 2. Comparison between different outlier detection methods

In Table 7-2, we compare the three methods for data from Data3_j24. The last column in Table 7-2 illustrates outliers detected by manual examination of data. We can see that the outliers detected by the PCA methods match well with those selected by manual inspection of raw data to detect outliers. The STDev method does not treat the different pins fairly; even though two pins may be similar, one may happen to have a wider range due to one or few boards causing a higher deviation for that pin.

Figure 7-1 identifies the specific pins of boardruns 3, 17, 83, 14, 11 with STDev method with $\alpha = 4$. The STDev method can effectively detects both connector measurement 17 and 83 as outliers. However, the method unfairly treated others: in fact, connector measurement 83 at pin 155 showed very close measurement value to the mean value. The

far different standard deviation between the two pins makes the threshold values between them are much different.

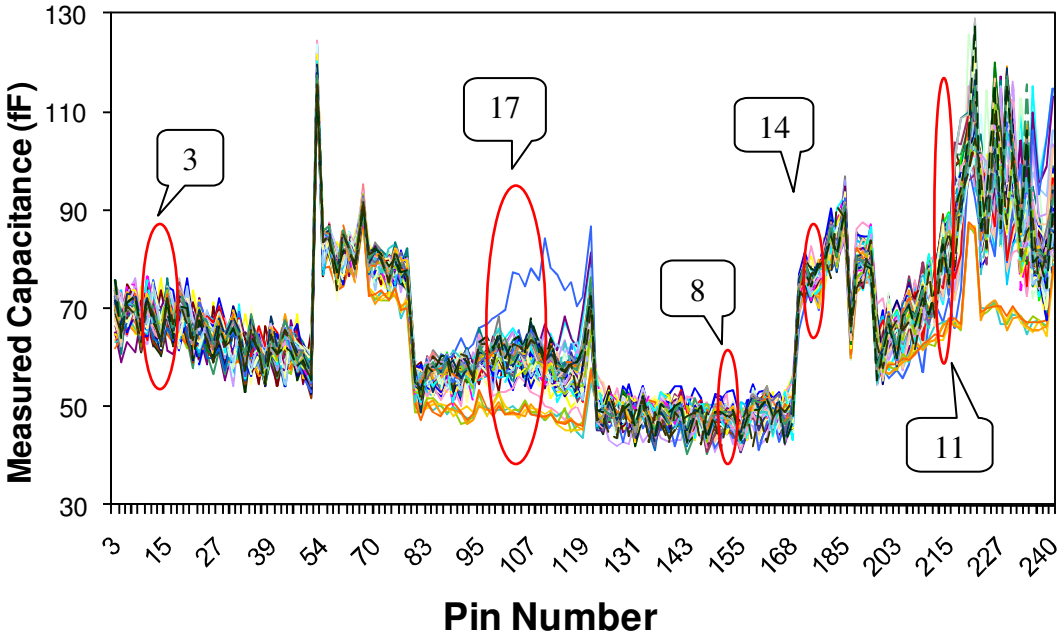


Figure 7- 1. Position of the outliers detected by STDev method in Data3_j24 plot

More importantly, the standard deviation method does not consider the correlation among the tested pins. In PCBs, an abnormal pin will cause the pins around it show abnormal test values too. The PCA implicitly captures the correlation between different pins for the final result.

Of course, the STDev method is not always incorrect. When it is applied to the Data3_j25, the test result is very close to the inspection and PCA method.

Chapter 8.
PCA-Based Analysis with Test Windows

When the number of measured pins per board is quite large, such as with Data_D1 evaluated in Chapter 5, it is difficult to identify defects that only affect the measurements of a small subset of pins by a relatively small magnitude. For such cases, a more sensitive scheme is needed to take into account local perturbations, i.e., those limited to a very few neighboring pins.

Since the open defect of a single pin can influences on the measurements of a few neighboring pins, the usage of test windows of neighboring pins for outlier detection is proposed. We term this as 'local analysis'. In contrast, the analyses with the whole dataset in previous chapters could be called 'global analysis'.

The local analysis requires test data to be sorted according to the physical order of pins on the connector. The union of outliers from different windows is considered to be the set of outliers. This section also considers the selection of the window size as well as use of overlapping windows to further enhance the scheme.

Local analysis may not be applicable to PCA based outlier detection schemes for other test problems such as IDDQ [4][5], when the effect of faults cannot be local to a set of measurements. With PCBs, the physical adjacency among the pins is known and the local analysis method produces a significant increase in sensitivity as illustrated below. Furthermore, it helps isolate the location of the outlier pins to within a small set, potentially resulting in a significant savings in the time required for subsequent board inspection and repair.

8.1. Location of Outliers

Figure 8-1 shows the schematic example of window division in DDR2 connector where the connector is partitioned into different test windows. Ten test windows of 24 pins each are selected equally dividing the 240 pins of connector j24.

If the untested power pins and ground pins are included, the total number of pins is the same in all windows. Those untested pins are colored in white in Figure 8-1.

The pins inside of one test window are formed by two different rows: one row with pin 1 to pin 120 and one row with pin 121 to 240 like that in Figure 8-1. With the local scheme, to locate the outlier in Data3_j24, all of the tested results are first sorted in the same order as the physical layout of the connector, which transfers Figure 5-1 to Figure 8-2. After being sorted, the pin number changes alternatively between two rows of the connector. If there are untested pins, they will be skipped.

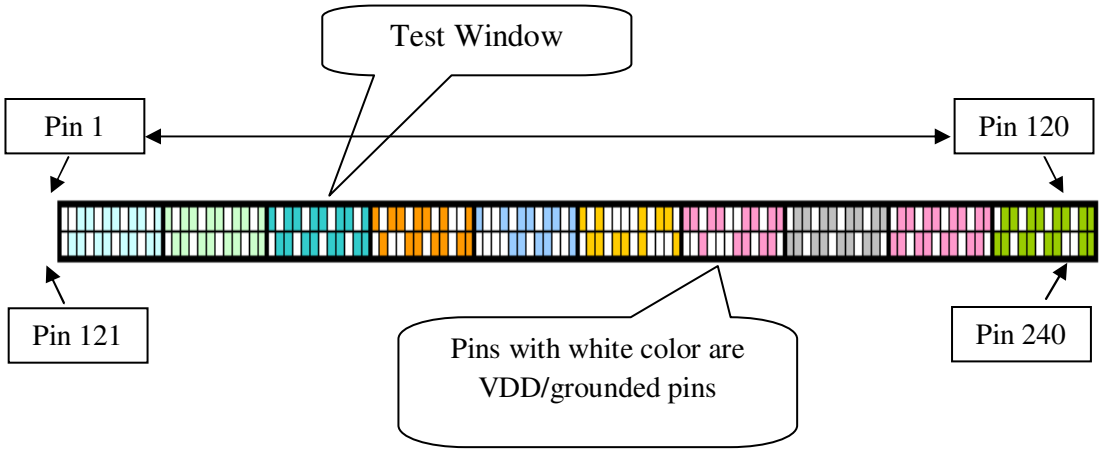


Figure 8- 1. Test window division on board connector j24

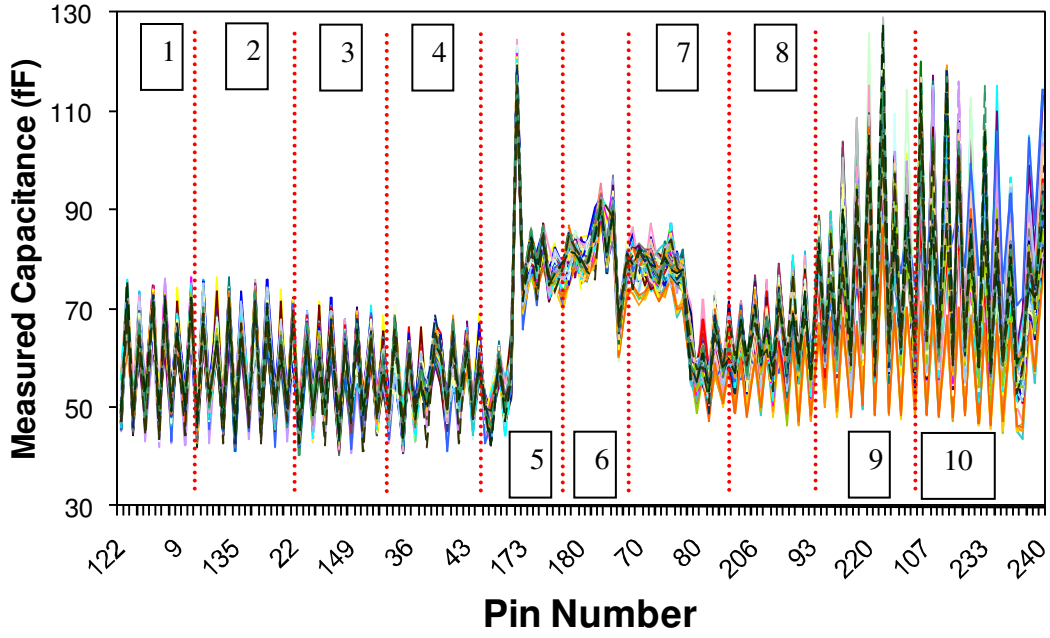
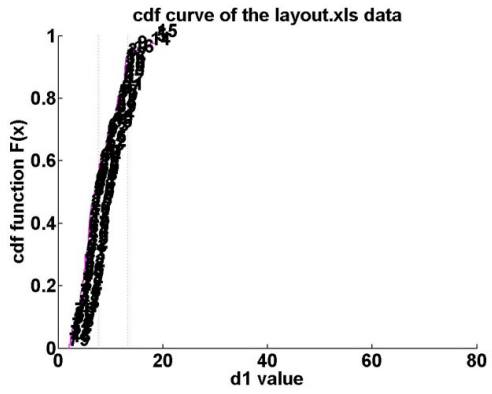


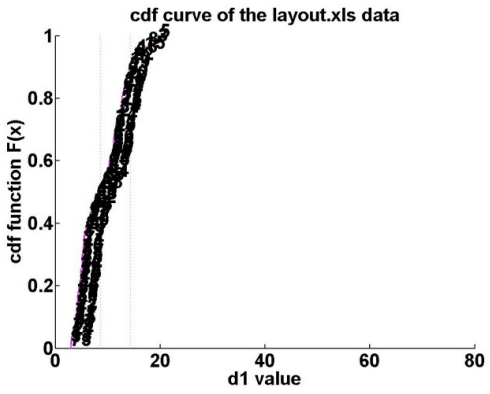
Figure 8- 2. Plot of 10 test windows in Data3_j24 according to the physical layout

Figure 8-2 is a plot of measured data sorted as described. Although the ten test windows equally divide the 240 pins, windows still show different sizes. The reason is that there are no measurements of the pins being supply or ground pins. For example, only tested pins 122, 3, 123, 4, 125, 6, 126, 7, 128, 9, 129, 10, 131, 12, 132 are included in test window one.

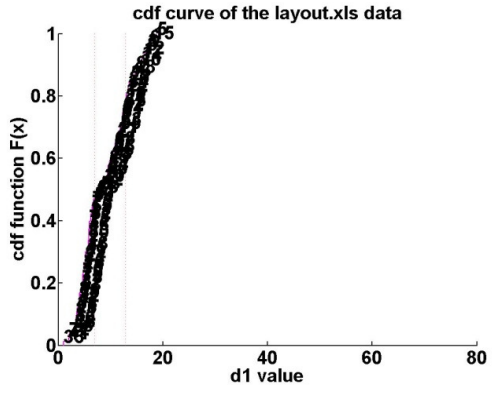
PCA based d_1 evaluation is applied to each of the 10 test windows individually. Under the same axis scale, ten d_1 values of the CDF plots are shown separately in Figure 8-3. Comparing the ten plots in Figure 8-3, we can see that the CDF curves in test windows 8, 9, 10 spans a longer d_1 range than those in other test windows. It implies that the measurements from different boards are very similar to each other based on window 1 2, 3, 4, 5, 6 and 7. In window 8, 9 and 10, there are clear outliers existing.


cdf curve of the layout.xls data
cdf function F(x)
d1 value
0
0.2
0.4
0.6
0.8
1
0
20
40
60
80

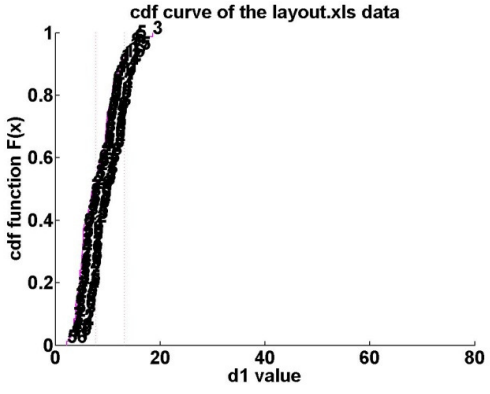
(a)


cdf curve of the layout.xls data
cdf function F(x)
d1 value
0
0.2
0.4
0.6
0.8
1
0
20
40
60
80

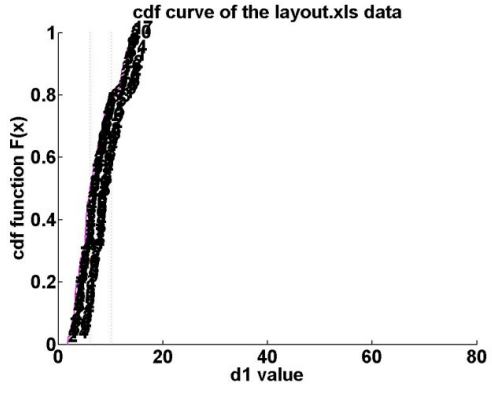
(b)


cdf curve of the layout.xls data
cdf function F(x)
d1 value
0
0.2
0.4
0.6
0.8
1
0
20
40
60
80

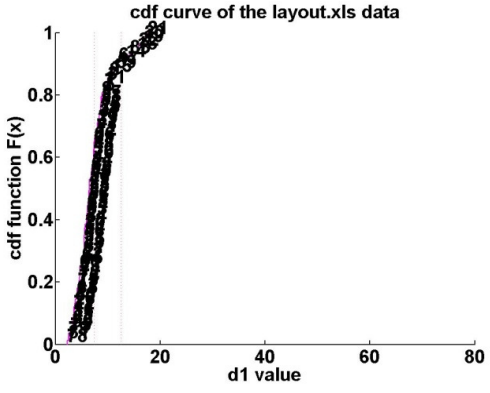
(c)


cdf curve of the layout.xls data
cdf function F(x)
d1 value
0
0.2
0.4
0.6
0.8
1
0
20
40
60
80

(d)


cdf curve of the layout.xls data
cdf function F(x)
d1 value
0
0.2
0.4
0.6
0.8
1
0
20
40
60
80

(e)


cdf curve of the layout.xls data
cdf function F(x)
d1 value
0
0.2
0.4
0.6
0.8
1
0
20
40
60
80

(f)

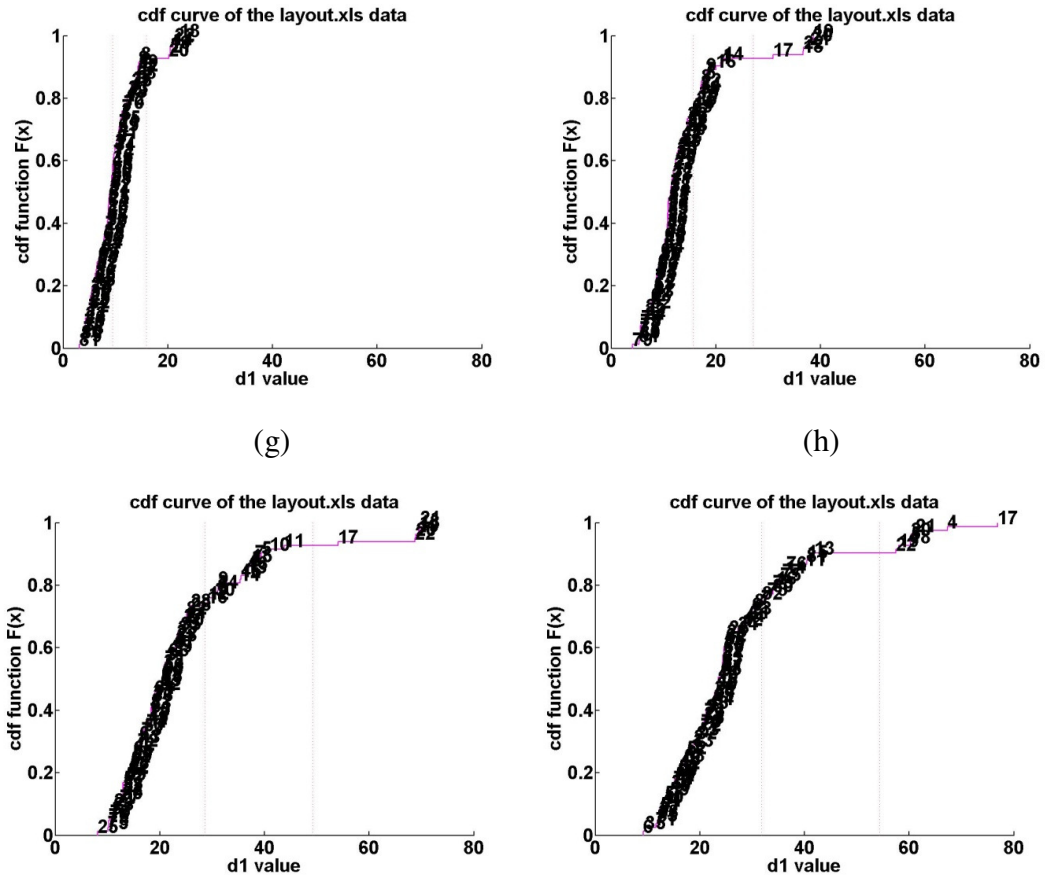


Figure 8- 3. CDF plot to Data3_j24 in (a) test windows 1, (b) test windows 2, (c) test windows 3, (d) test windows 4, (e) test windows 5, (f) test windows 6, (g) test windows 7, (h) test windows 8, (i) test windows 9, (j) test windows 10.

As discussed in Chapter 6, when clear outliers exist, we can expect the maximum d_1 value of that dataset be relatively high. Such experience can also be used in local analysis. The larger the d_1 value in a test window, the more likely there are outliers existing in that window. After finding the test windows containing outliers, we can then perform detailed outlier detection by identifying corresponding CDF plot.

Figure 8-4 plots the maximum d_1 value for different windows. Outliers are observed in the last three windows, which are clearly indicated by the wide range of d_1 values in

windows 8 to 10. When we examine the last three CDF plots in Figure 8-2, the variance of the last three test windows are much higher than those of other windows, which implies the possibility of abnormal measurements in the last three windows.

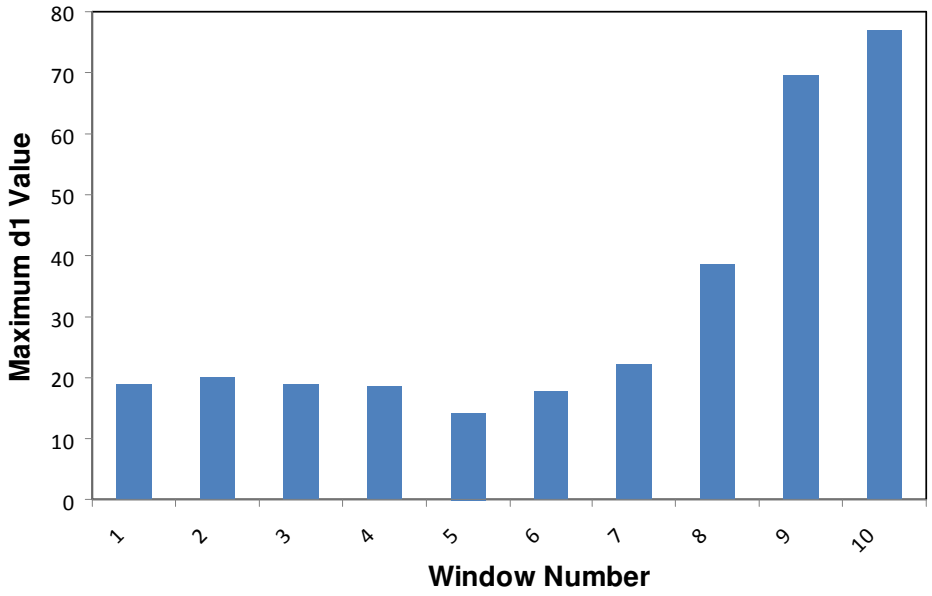


Figure 8- 4. Maximum d_1 value vs. test window number for Data3_j24

In physical view, the last three windows correspond to the 72 pins at the right end of the j24 connector shown in Figure 8-1. There is also a possibility that the tester introduced errors such as those due to misalignment of the sense plate. Retesting followed by repairing may be required for those related j24 connectors.

In Figure 8-3, even the last three test windows have the maximum d_1 much smaller than that in Figure 5-14. That is caused by smaller data size in each test window which leads to the low maximum d_1 value.

Since the total numbers of tested pins are not equal in the window analysis above, we can modify the window to the one containing equal number of tested pins, which will

make the comparison fairer. New Maximum d_1 value in each test window, after such a partition, is shown in Figure 8-5, which gives very similar trend to that in Figure 8-4.

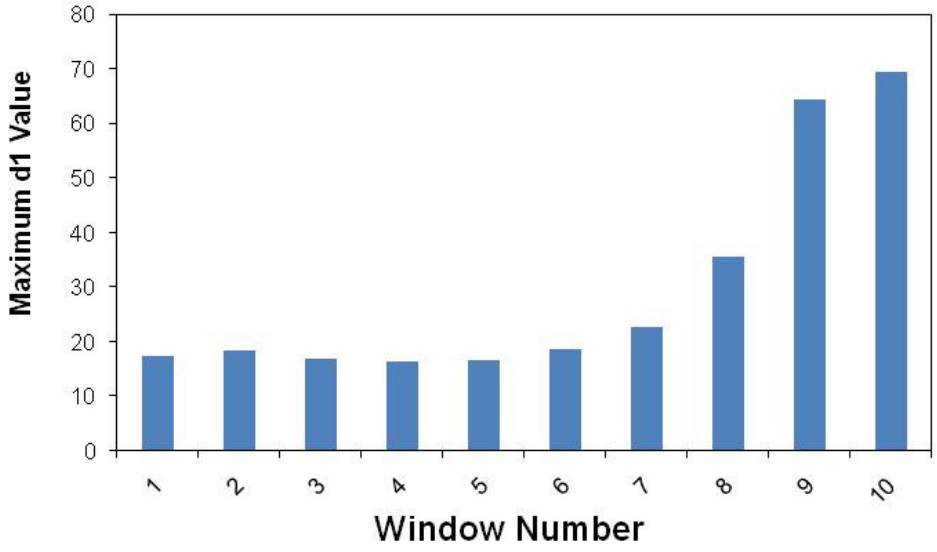


Figure 8- 5. Maximum d_1 value vs. test window number with equal number of tested pins per window for Data3_j24

8.2. Overlapped Test Window Analysis

Maximum d_1 values in the Figure 8-4 effectively show the locations of the outlier pins. However, in that case, test windows do not overlap. Thus it fails to take into account the impact of a fault on the boundary of one of the windows on the pins in the next window. Thus, another option is to use overlapping windows for the evaluation.

Figure 8-6 uses a set of test windows such that each of them includes the adjacent halves of the two test windows of Figure 8-4. For example, the test window 1 in Figure 8-4 includes pins {122, 3, 123, 4, 125, 6, 126, 7, 128, 9, 129, 10, 131, 12 and 132}; the test window 2 in Figure 8-5 includes pins {13, 134, 15, 135, 16, 137, 18, 138, 19, 140, 21, 141, 22, 143, 24 and 144}. The overlapping test windows in Figure 8-6 are named after the two test windows in Figure 8-5 they overlap. For example, window 1-2 in Figure 8-6

includes the second half of original window 1 and first half of original window 2, which thus contain {128, 9, 129, 10, 131, 12, 132, 13, 134, 15, 135, 16, 137, 18, 138}. Note that the unlisted pins are VDD or grounded pins.

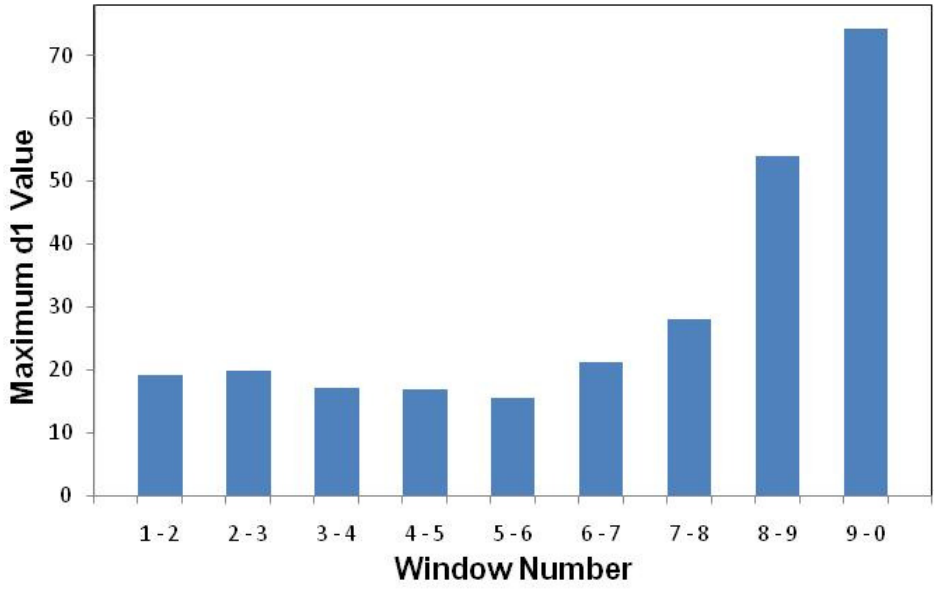


Figure 8- 6. Maximum d_1 value vs. overlapped test window number for Data3_j24

The plots of the maximum d_1 value in both Figure 8-4 and Figure 8-6 are similar. In this case, there is no fault influencing measurements on pins of adjacent non-overlapping window. However, in general, such cases cannot be ruled out. The overlapping test windows can be used to identify the location of outlier pins more finely by combining the analysis using both original and overlapped windows. Analysis combined Figure 8-4 and Figure 8-6 can increase test resolution.

For Data3_j24, similar phenomena are observed when test windows shift to the right along the connector the same window size. For each test window, it has nine different beginning locations.

Figure 8-7 shows the maximum d_1 value in each test window in partial overlapped window-based analysis, this kind of analysis is named as 'Moving Window Analysis' is this thesis. Compared with overlapped analysis mentioned previously, the Moving Window Analysis shifts the test window to right one pin at a time. As a result, the shifted test window is not the half-half overlapped to original windows any longer. For instance, a shifted test window can cover most pins of the fist window and least pins of the second one.. Simulation is performed in each step. Then, in Figure 8-7, there are 9 different test windows. The 13 bars in column 1 are the first windows of 13 different shifts. The 13 bars in column 2 are the second windows of 13 different shifts.

Figure 8-8 shows the change of device number that gives maximum d_1 value in each test window along with the shifts. With the small change of test window location, outlier devices detected in the test window don't change much. That can be detected clearly from the right part of Figure 8-8 (right circle area), where outliers exist. In the Figure 8-7 and Figure 8-8, each color represents the simulation result in the same shift.

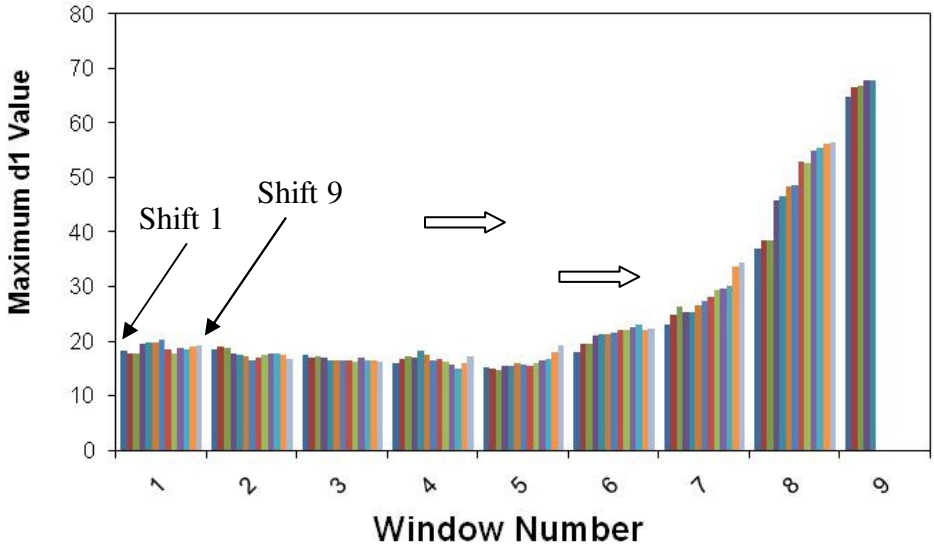


Figure 8- 7. Maximum d_1 value in Moving Windows Analysis for Data3_j24

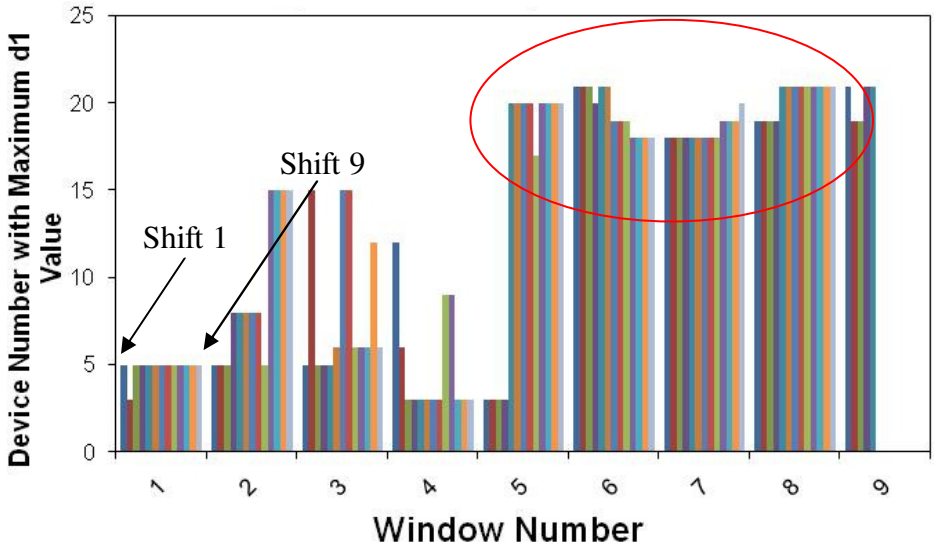


Figure 8- 8. Boardrun numbers with maximum d_1 value in Moving Window Analysis for Data3_j24

8.3. Comparison of Global and Local Methods

Local method detects outliers based on the evaluation of each test window, and as a result can effectively detect opens in that window. On the other hand, it fails to take into

account trends of more global nature, such as tester to tester variations, and alignment errors, which can be observed when measurements for different devices are compared over a larger window. The global method is more effective in capturing such characteristics.

Consider for example the TestJet results for j3_all as shown in Figure 8-9, boardrun 4 shows a clear spike at pin 33. With the local method in Figure 8-10, the window 11 with that pin indicates boardrun 4 as an extreme outlier, i.e., one that is on the right end of the CDF plot separated from others by a significant gap. With the global method, boardrun 4 doesn't appear as an extreme outlier since all other measures of the boardrun are within the range of variation of most of the boards. The local method can work with a relatively few devices, while the global method requires measurements for a relatively larger number of devices. Figure 8-10 shows the range of d_1 values observed for dataset j3_all. The larger the range of d_1, the more likely there are outliers in that window. The outliers can be clearly identified by using the corresponding CDF plot.

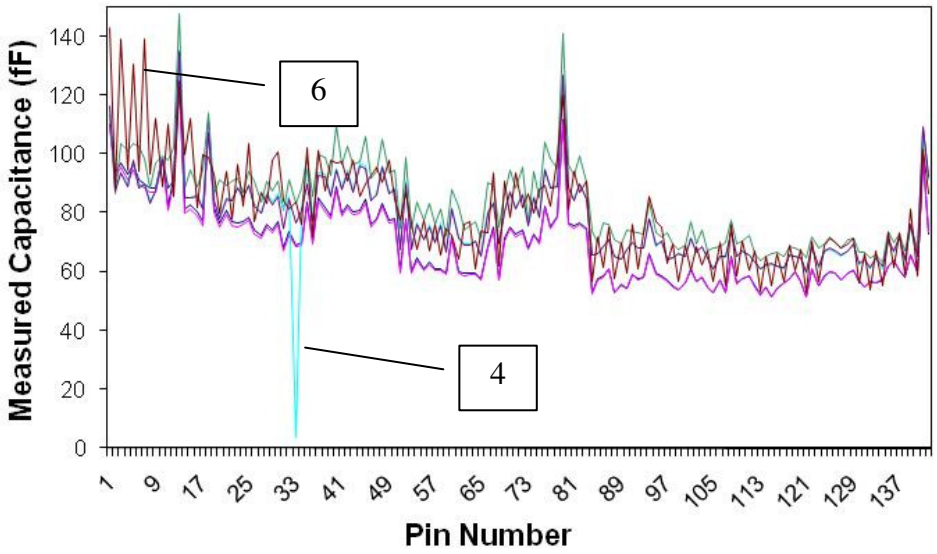


Figure 8- 9. Raw measurements plot to j3_all with boardruns 4 and 6 out of 6 boardruns

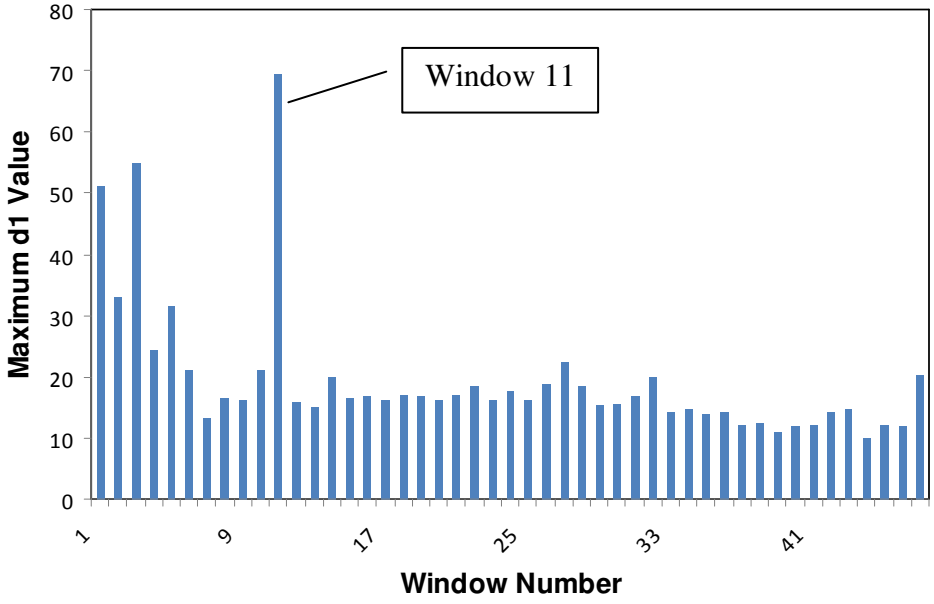


Figure 8- 10. Maximum d_1 value in each window for j3_all

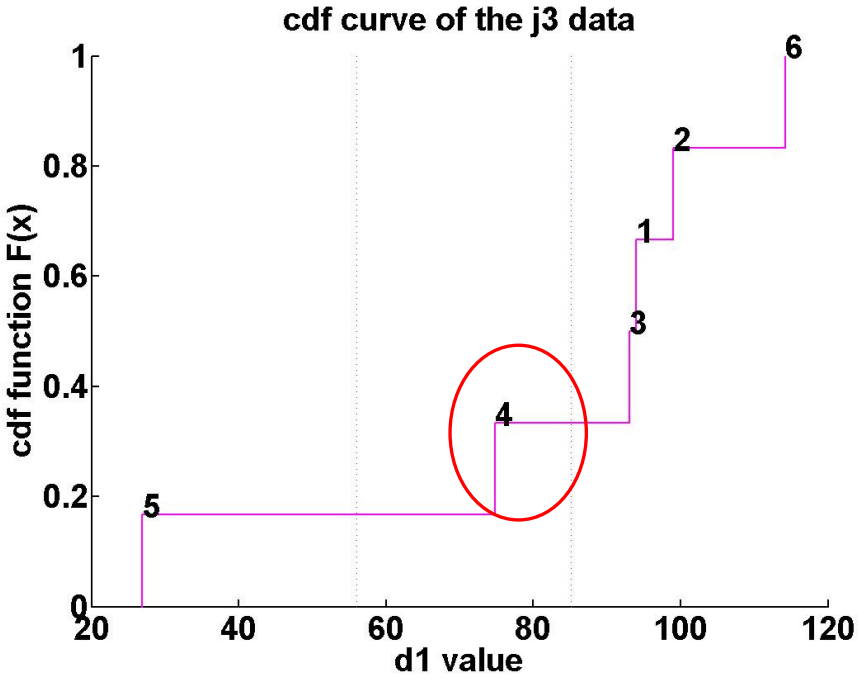


Figure 8- 11. CDF plot of j3_all with global analysis

Figure 8-11 shows the CDF plot of the j3_all obtained with the global method. As can be seen, boardrun 4 does not produce an extreme d_1 value. After the local analysis,

there is a total of five test windows that have a d_1 range higher than 30. Figure 8-12 identifies these windows, i.e., windows 1, 2, 3, 5, and 11.

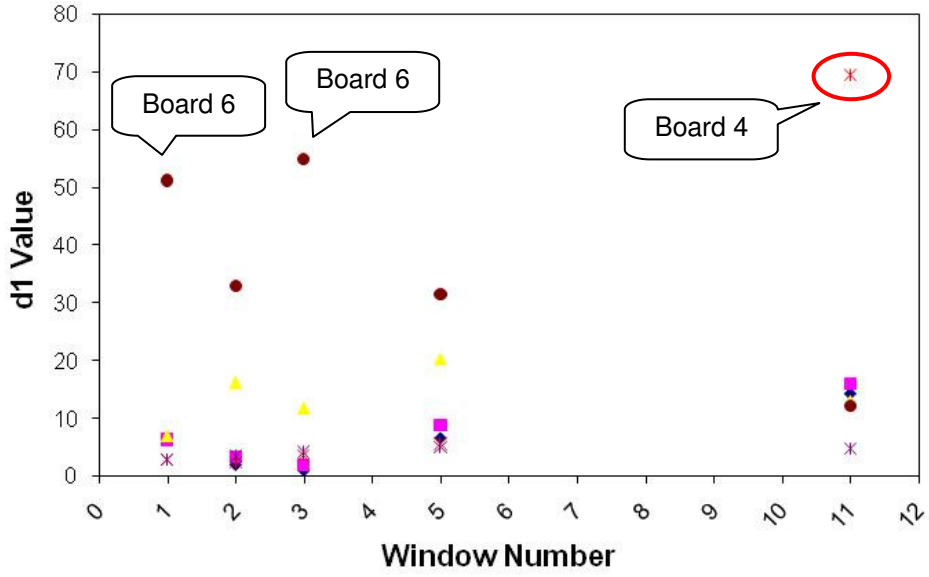


Figure 8- 12. d_1 values distribution in the five test windows

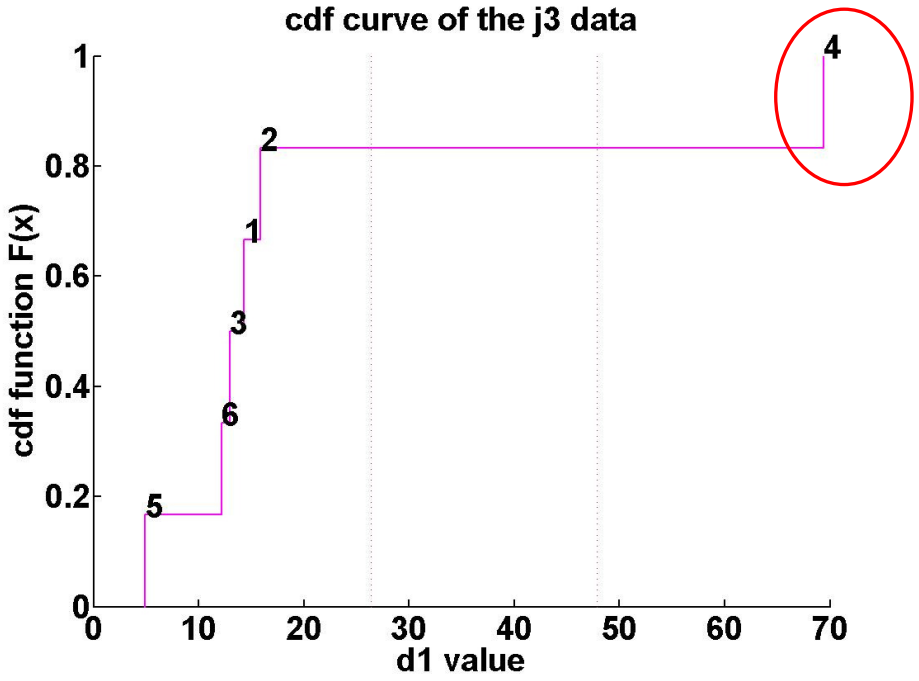


Figure 8- 13. CDF plot to j3_all with local analysis to window 11

From Figure 8-12, boardrun 4 shows an extreme value only in window 11; the other 4 test windows all contain very low d_1 values for it. Figure 8-13 shows the CDF plot of test window 11, which gives boardrun 4 very high d_1 value thus separating it clearly from the others.

Comparing the plots for both the global and local methods, following observation can be made: boardrun 4 does not show high d_1 value in the global analysis; however it shows one very high d_1 value in test window 11. From the overall data viewpoint, boardrun 4 is one that is very close to the normal range of measurement values for the rest set of tested pins; this can also be seen from the measurements plot of boardrun 4, where except for a single spike, other measurements are well within normal range of variation. A boardrun that does not show high d_1 value in the overall analysis but appears as an outlier in one or two of the test windows must contain spike signals.

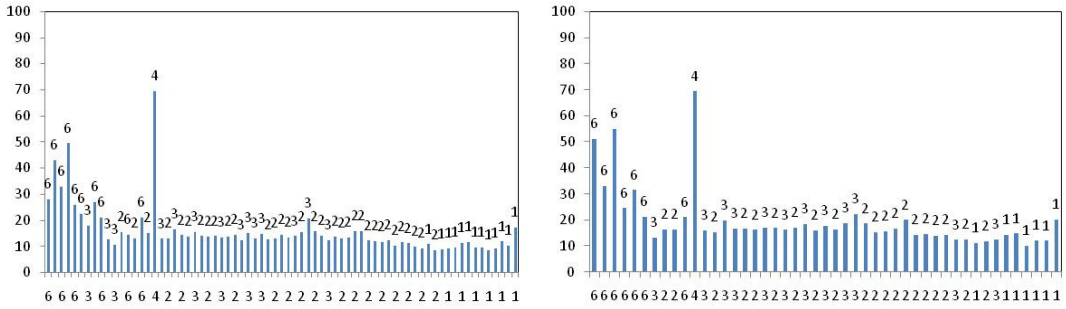
The global method can effectively catch the obtuse outliers like boardrun 6 in Figure 8-10. Here the measurements for a group of pins are outside the norm, but each one deviates only slightly from the norm. The local method can catch outliers with one or few sharp peaks, such as boardrun 4. The advantages of the two methods can be combined by a test technique using both global and local analysis and then taking the union of outliers.

8.4. Test Window Size in Local Analysis

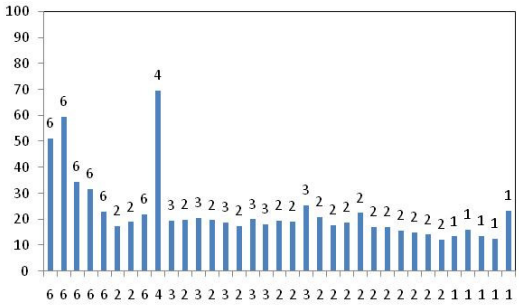
The global and local methods can be considered to be two instances of the same method, where the global method corresponds to a window that contains all the pins. Thus the notion of an optimal window size arises; however optimality is likely to depend on attributes of the dataset. A large window size will make the analysis less sensitive for

detection of the precise location of the outliers. It can also cause the problem of peak detection errors mentioned earlier in Chapter 6. A large window size can also miss outliers that manifiest only on one or very few pins, especially when the total number of pins is large. However, when a test window is too small, it may not be sensitive to the spatial correlation among the measurements. It would also take a lot of computation time to identify the outlier devices.

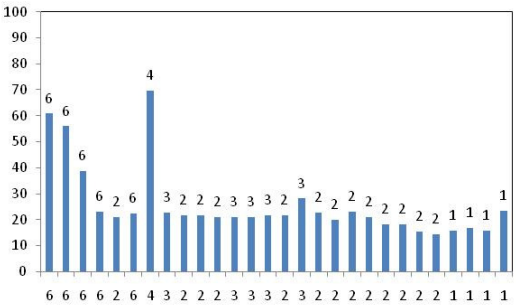
A preliminary analysis is done to see the impact of window size variation using test data. To identify a good criteria for the selection of window size, we use a metric, the maximum value minus the median value for different window sizes. A larger difference, indicates a higher variation in the test window. For example, in Figure 8-10 the highest d_1 value is the one corresponding to window 11, minus the midle maximum d_1 value of all test windows to set the criteria. The larger the difference the clearer the peaks in the test window. Figure 8-14 gives the maximum d_1 values in each test window for different window sizes. As can be seen, with the increase of test window size, total numbers of test window decreases.



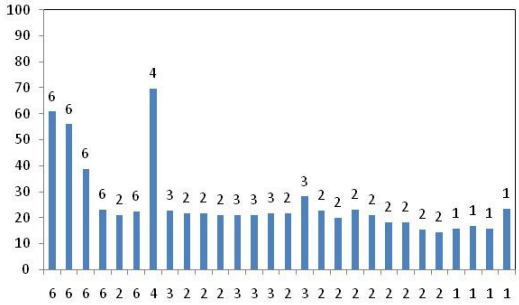
(c) (d)



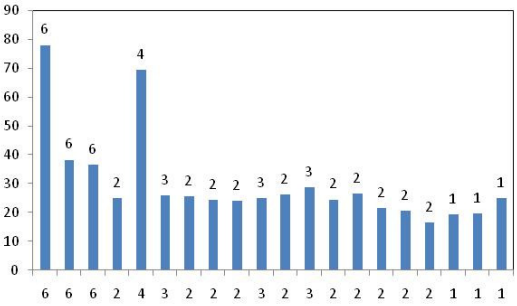
(c)



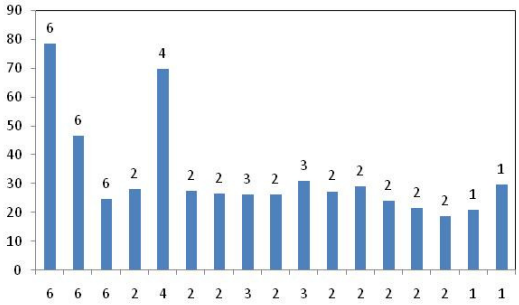
(d)



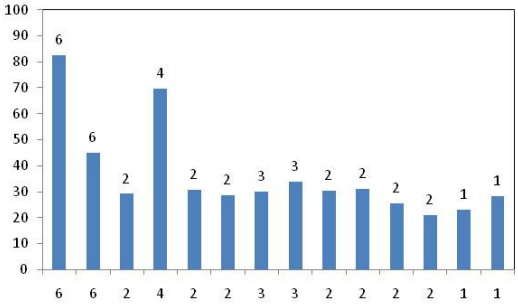
(e)



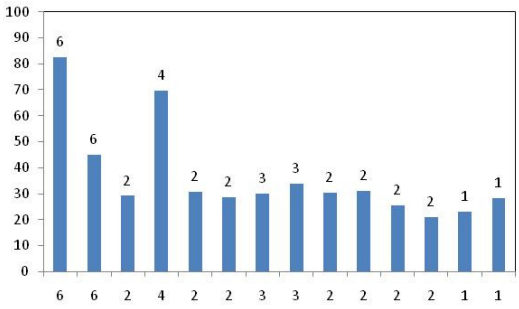
(f)



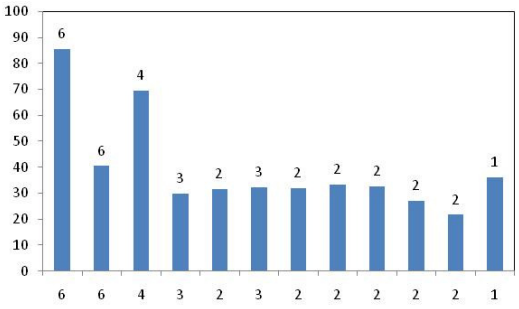
(g)



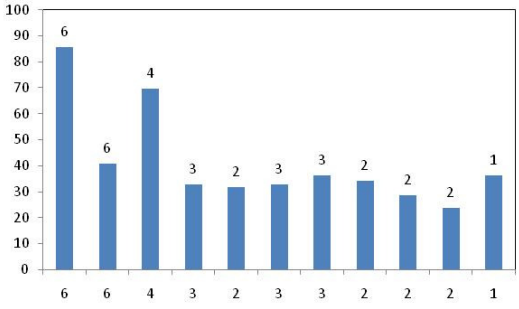
(h)


100 90 80 70 60 50 40 30 20 10 0
6 6 2 4 2 2 3 3 2 2 2 2 1 1
6 6 2 4 2 2 3 3 2 2 2 2 1 1

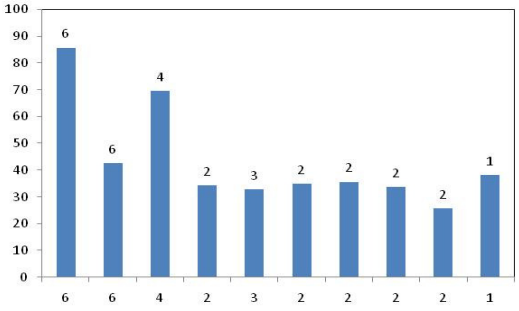
(i)


100 90 80 70 60 50 40 30 20 10 0
6 6 4 3 2 3 2 2 2 2 2 1
6 6 4 3 2 3 2 2 2 2 2 1

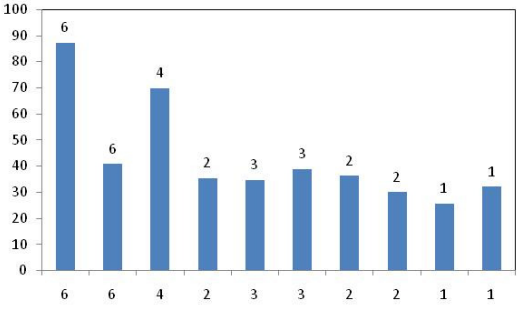
(j)


100 90 80 70 60 50 40 30 20 10 0
6 6 4 3 2 3 3 2 2 2 1
6 6 4 3 2 3 3 2 2 2 1

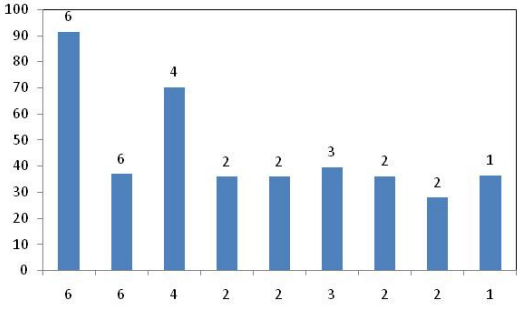
(k)


100 90 80 70 60 50 40 30 20 10 0
6 6 4 2 3 2 2 2 2 1
6 6 4 2 3 2 2 2 2 1

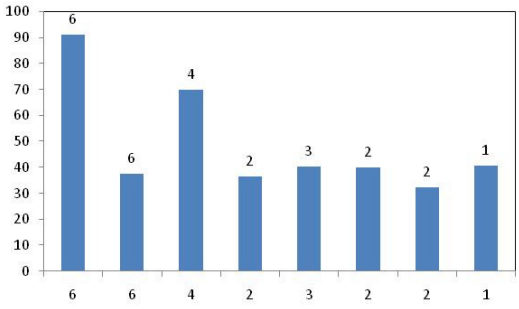
(l)


100 90 80 70 60 50 40 30 20 10 0
6 6 4 2 3 3 2 2 1 1
6 6 4 2 3 3 2 2 1 1

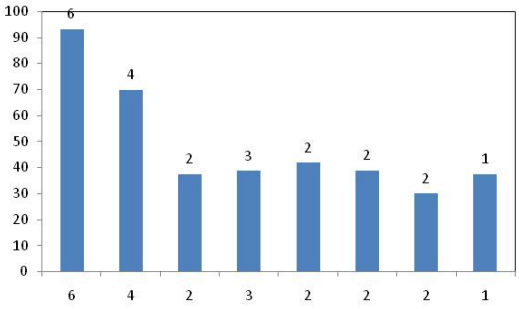
(m)


100 90 80 70 60 50 40 30 20 10 0
6 6 4 2 2 3 2 2 1
6 6 4 2 2 3 2 2 1

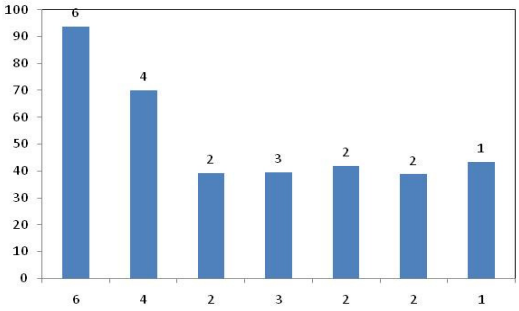
(n)


100
90
80
70
60
50
40
30
20
10
0
6 6 4 2 3 2 2 1
6 6 4 2 3 2 2 1

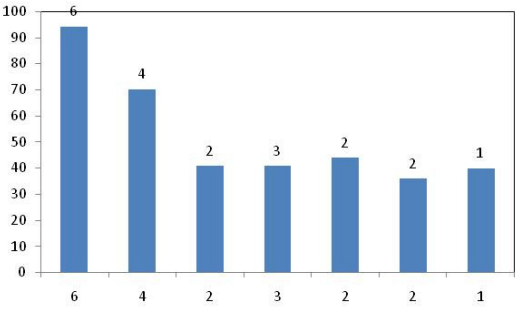
(o)


100
90
80
70
60
50
40
30
20
10
0
6 4 2 3 2 2 2 1
6 4 2 3 2 2 2 1

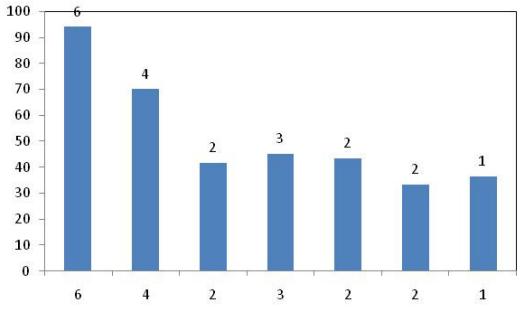
(p)


100
90
80
70
60
50
40
30
20
10
0
6 4 2 3 2 2 1
6 4 2 3 2 2 1

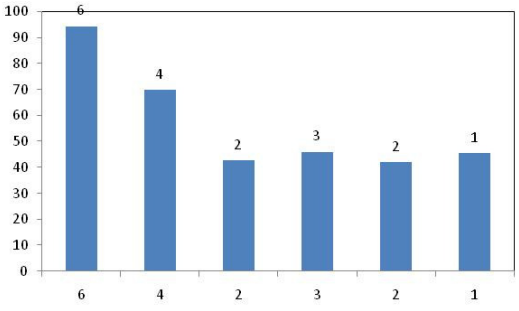
(q)


100
90
80
70
60
50
40
30
20
10
0
6 4 2 3 2 2 1
6 4 2 3 2 2 1

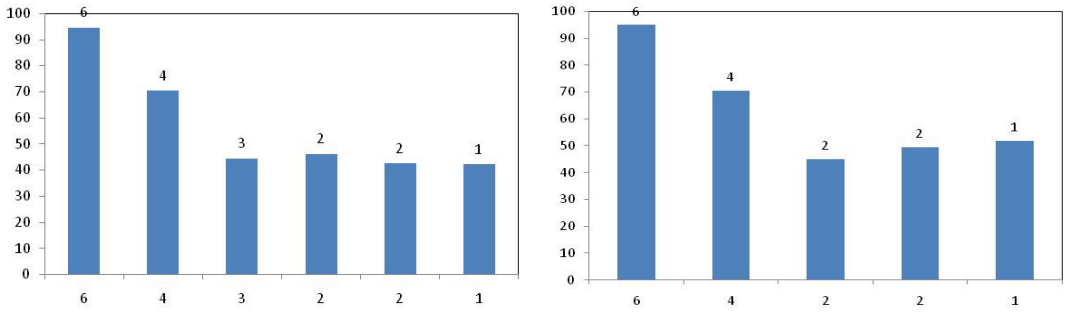
(r)


100
90
80
70
60
50
40
30
20
10
0
6 4 2 3 2 2 1
6 4 2 3 2 2 1

(s)


100
90
80
70
60
50
40
30
20
10
0
6 4 2 3 2 1
6 4 2 3 2 1

(t)

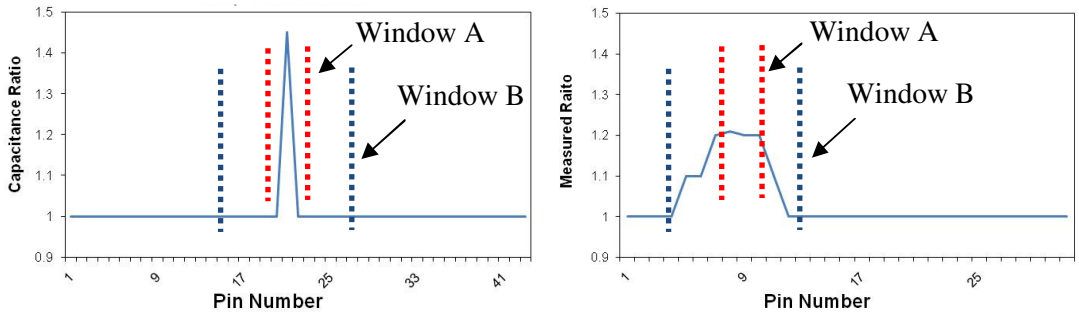


(u) (v)

……

Figure 8- 14. Maximum d_1 value showing boardrun number vs. test window number for Data3_j24 with (a) window size=2, (b) window size=3, (c) window size=4, (d) window size=5, (e) window size=6, (f) window size=7, (g) window size=8, (h) window size=9, (i) window size=10, (j) window size=11, (k) window size=12, (l) window size=13, (m) window size=14, (n) window size=15, (o) window size=16, (p) window size=17, (q) window size=18, (r) window size=19, (s) window size=20, (t) window size=21, (t) window size=22, (u) window size=23, (v) window size=24.

To detect sharp peak outliers like boardrun 4, smaller test window size is necessary, while to detect outliers like boardrun 6, which showed lot of shorter peaks, larger test window should be considered.



(a) (b)

Figure 8- 15. Plots comparison between measurements with (a) sharp outliers and (b) blunt outliers

Figure 8-15 illustrates situations with two different window sizes. When the test window is as small as Window A, correlations among different pins decreased to a small value. Then, the single peak or spike outlier will dominate maximum d_1 value easily. The outlier peak in Figure 8-15 (a) is higher than that in Figure 8-15 (b) resulting in its maximum d_1 value being higher.

However, with the increase of test window size such as the Window B in Figure 8-15, the outlier in Figure 8-15 (a) will be compensated for by its normal part leading to a smaller d_1 value in the larger test window. On the other hand, the outlier in Figure 8-15 (b) will add many new outliers inside of the large window making its d_1 value higher and higher.

In some datasets, different kinds of outliers may exist at the same time. For a fast and accurate analysis, an optimal test window size is needed. One criterion to find the window size is the most stable maximum-d_1-value difference observed with the windows size sweep simulation. The maximum-d_1-value is the maximum d_1 value in a window analysis, for example, window 11 in Figure 8-10 shows the maximum-d_1-value, the highest d_1 value in the plot. Similarly, median-d_1-value is the median d_1 value showed in window analysis.

Figure 8-16 shows the maximum maximum-d_1-value and the median maximum-d_1-value with different window size test in dataset j3. For example, when the window size is 22, corresponding point on blue curve of Figure 8-16 shows the maximum maximum-d_1 value in the last plot of Figure 8-15. Corresponding point on red curve of Figure 8-16 shows the median maximum-d_1 value among the five maximum d_1 values in the last plot

in Figure 8-14. Using median value instead of mean value helps avoid the cluster data effect.

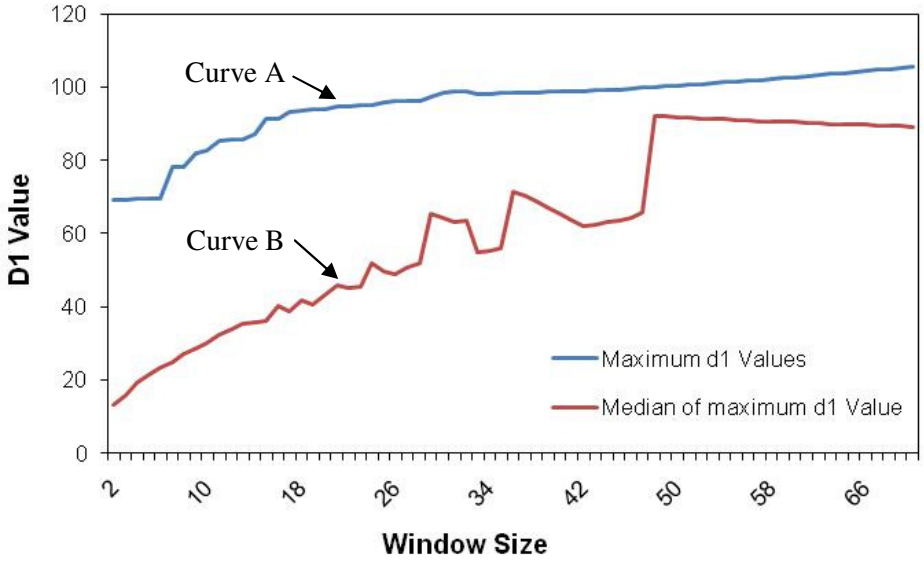


Figure 8- 16. Maximum and median maximum-d1 values with different window size in j3_all

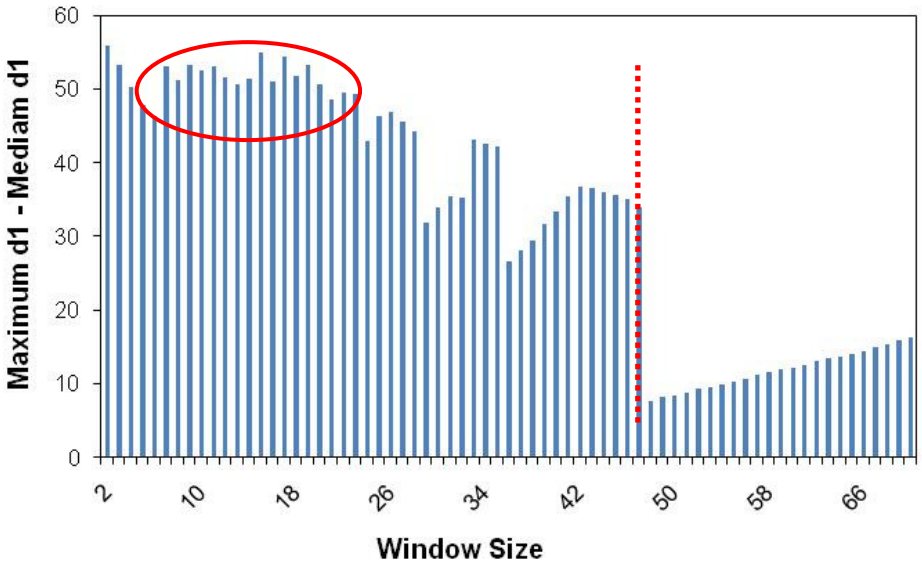


Figure 8- 17. Largest Median d_1 values for j3_norm with different window size

The difference values between the maximum maximum-d_1-value and median maximum-d_1-value for different window sizes are plotted in Figure 8-17, which is the difference value between curve A and curve B in Figure 8-16. We can tell that the high difference values occur with the lower window size.

As can be seen, the difference values at the end of Figure 8-17 are much smaller than others, but they change approximately linearly. The window size in that part is so large that only two test windows exist simultaneously. The separation between the two windows changes one by one to right side.

Similar simulations are applied to data3_j24 and data3_j27, and the results are shown in Figure 8-18 and Figure 8-19 respectively. In Figure 8-18, the maximum difference value no long exists at the low window size. While, the window sizes between 10 and 26 still hold a high enough value of difference. Some other high difference values appear in the later curve, but those test window sizes are too large to give a fine analysis resolution. In Figure 8-19, the difference of maximum and median maximum-d_1-values increases with the increase of test window size till a window size of 40.

According to Figure 8-17, 8-18 and 8-19, high values for the difference don't always exist with the small window size. When the window size is very large, analysis resolution will decrease. Based on the comparison, test window sizes between ten and twenty five can be a good selection for different PCB measurement datasets. When further tests are conducted to detect the peak or blunt outliers, the test window can be shrunk or expanded after the primary test.

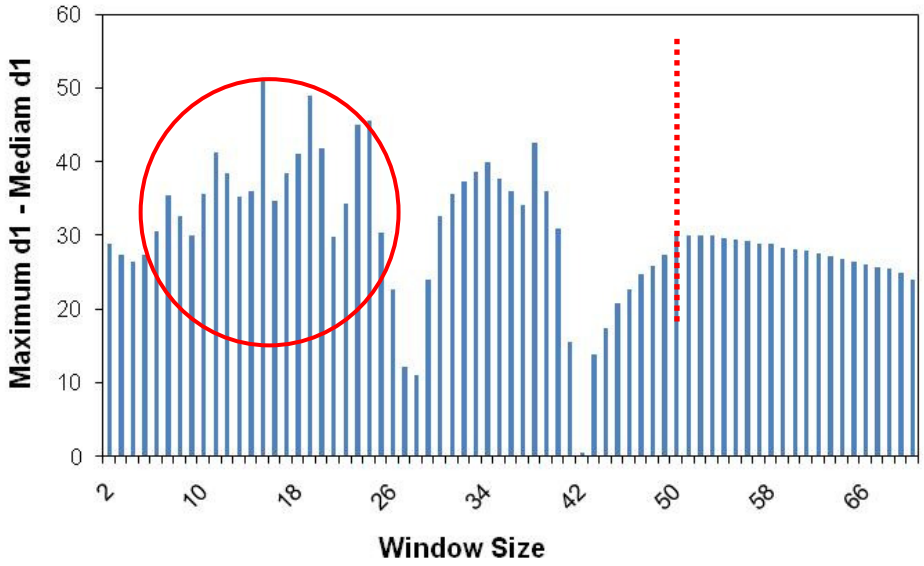


Figure 8- 18. Largest median d_1 values for Data3_j27 with different window size

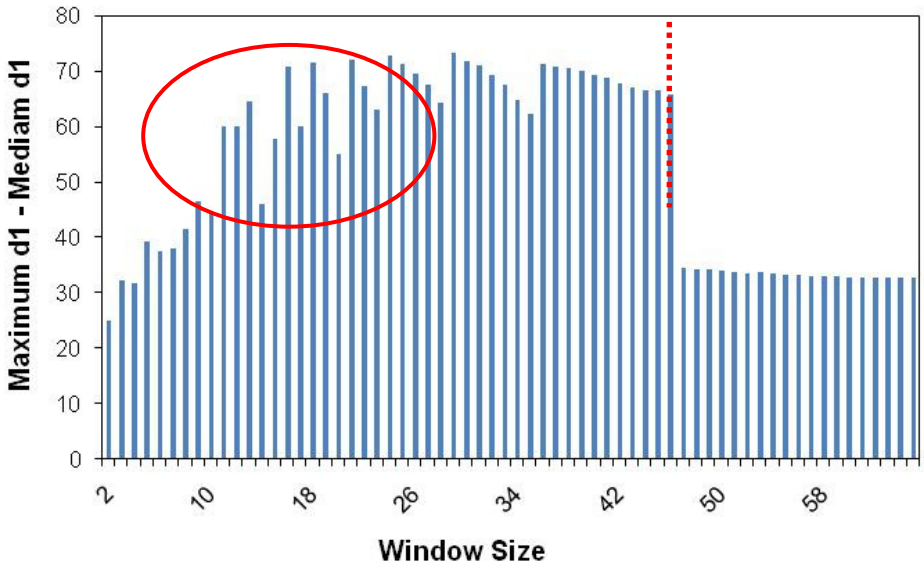


Figure 8- 19. Largest median d_1 values for Data3_j24 for different window sizes

8.5. Discussion of Local and Global Methods

The global and local analysis have their own advantages and disadvantages in detecting outliers. Thus we can benefit by combining the two methods.

Global / Local	Pass (P)	Fail (F)
Pass (P)	(i) Not Outlier	(ii) Possible Outlier
Fail (F)	(iii) Possible Outlier	(iv) Outlier

Table 8- 1. Analysis based on combination of global and local analysis

In Table 8-1, we listed combinations between global and local analysis results. When a boardrun is diagnosed as Pass (P) with both methods (i), it is not outlier. Similarly, when a boardrun is diagnosed as Fail (F) with both methods (iv), it is an outlier.

For the situation in (iii), where the boardrun fails in local analysis but passes the global one, we can conclude that some defects giving a sharp peak signal is very likely to be in the dataset, for example, a solder open on the connector. For situation (ii) the result may be caused by variation that shifts the multiple pin measurements to a less obvious level. The variation can be board to board variation or multiple defects. Then, the boardruns in situation (ii) and (iii) can termed as possible outliers.

A flow chart for combined PCA based global and local analysis for boardrun outlier detection is given in Figure 8-20 below:

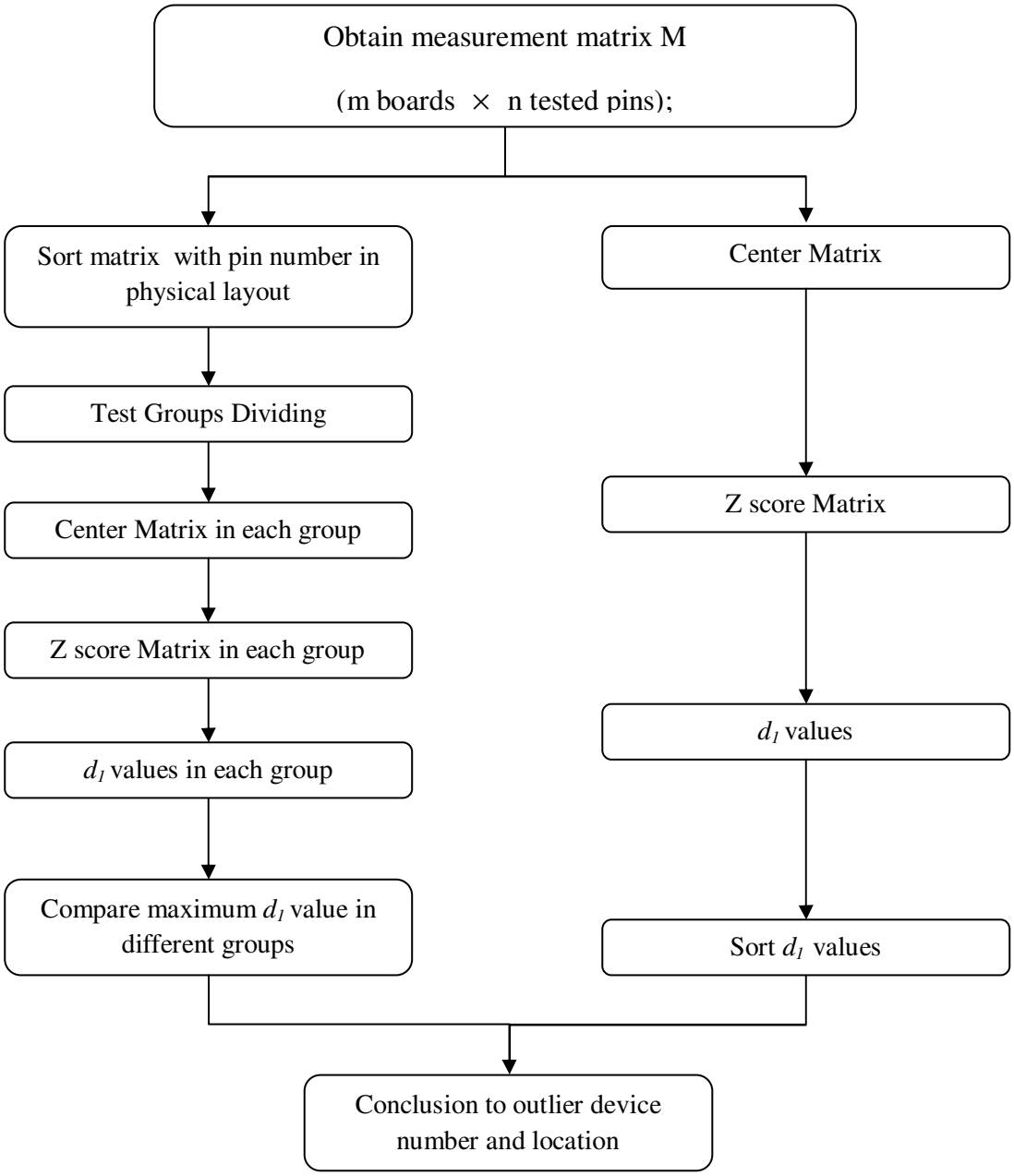


Figure 8- 20. PCB outlier detection steps with PCA

Chapter 9.
Analysis Based on Fixture Variation

As mentioned in previous chapters, PCA-based approach treats multi-variable measurement from connectors in a holistic manner to overcome the measurement and component parameter variations inherent in test data. Sensitivity of the approach can be adjusted by selecting windows of pins.

However, when good boards are tested under different fixtures, which are subject to some mechanical misalignments of the sense plate, the measured values vary depending on the degree of misalignment. This chapter considers approaches to compensate for the measurement differences for different fixtures that will help avoid incorrect classification of outliers. Using different datasets, compensation methods are evaluated to see if PCA is able to detect outliers correctly. PCA based analysis is presented for data sets containing different modes of misalignment illustrating its ability to identify the degree of misalignment.

9.1. Causes of Mechanical Variation

Mechanical variations in the connector or sense plate cause changes in the measurement values although the connector is not defective. Such situations may occur when there is variation in connector height with respect to pin height within the connector, or coplanarity of the connector ball connections is lopsided causing the sense plate to be tilted or vertically shifted.

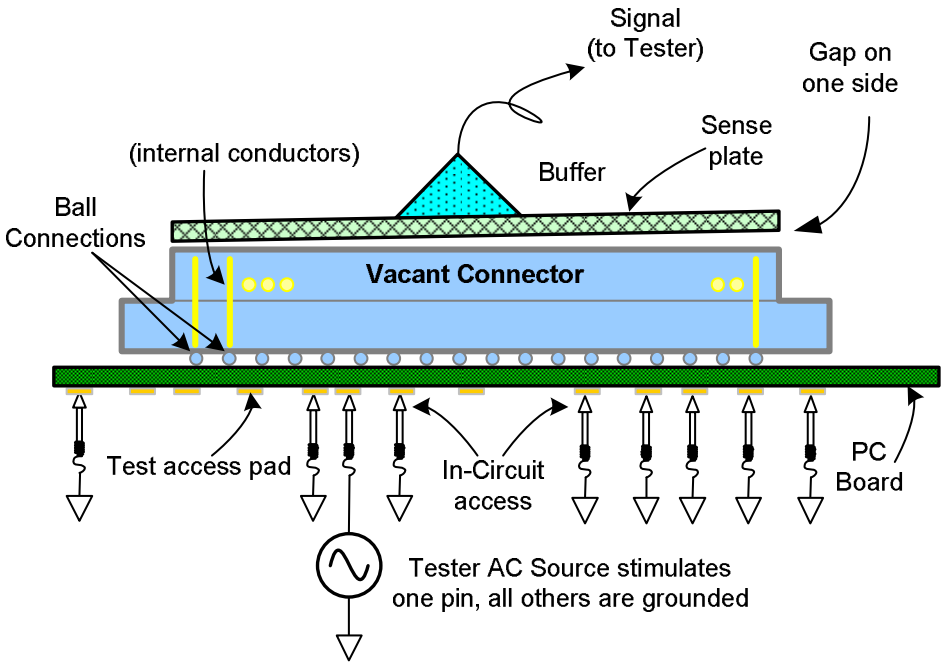


Figure 9- 1 Testjet with mechanical 'tilt' on the sense plate [Courtesey of K. Parker, Agilent Technologies]

Figure 9-1 shows a normal connector under test with mechanical variation on a Testjet sense plate, where the fixture has a variation in its planarity relative to the connector – that is, it is tilted with the left side slightly lower than the right. We refer to this as a left-tilted sense plate. This in turn opens a gap between the right side of the sense plate and the connector. Figure 9-2 shows a connector tested with a vertically shifted sense plate. The sense plate is raised a bit higher with respect to the internal connector pins being sensed. A simple variation in the height of the plastic connector housing can cause this. Indeed, typical specifications on connector housing height are given in 'max' form with no tolerance. The actual tolerance is likely to be in the +/- 0.004 inch (0.1 millimeter) range.

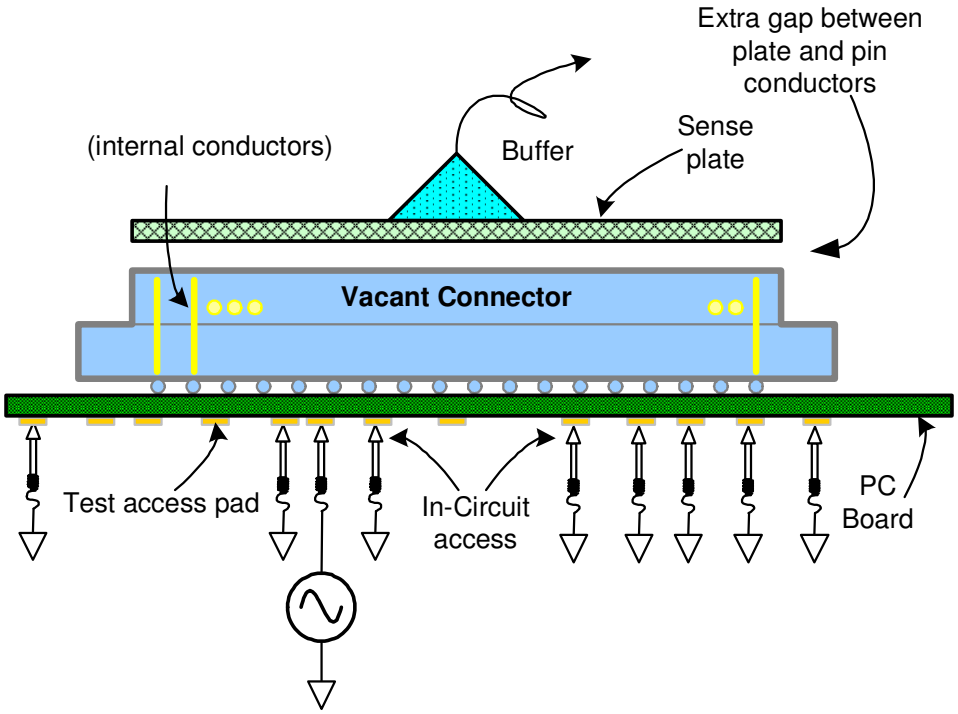


Figure 9- 2 Testjet with mechanical "shift" on the sense plate [Courtesey of K. Parker, Agilent Technologies]

Mechanical variations can cause systematic variations in measurement data that then appear in outlier detection processing. In a laboratory setting, the tester in Agilent Technologies injected specific mechanical variations of both "shift" and "tilt" types, and then collected data for a variety of connectors on a small set of boards. The amounts of shift or tilt injected were relatively small offsets from a "zero" reference. They were 8, 16 and 24 thousandths of an inch (mils) or, 0.20, 0.41 and 0.61 millimeters. The respective setups (see Table 1) are named as Tilt_1, Tilt_2, Tilt_3 and Shift_1, Shift_2, and Shift_3 in the following text. The shift and tilt experiments were independent; we did not combine their effects. Tilt_0 and Shift_0 both refer to the reference case, which does not contain any tilts or shifts.

Setup	Left or right End Height (mils)	Setup	Double Ends Height (mils)
Ref	0	Ref	0
Tilt_1	8	Shift _1	8
Tilt_2	16	Shift_2	16
Tilt_3	24	Shift_3	24

Table 9-1. Experimental data

9.2. Fixture Misalignments

In a real test environment, DUTs tested under abnormal sense plate positions, such as tilted or vertically shifted sense plates, may result in measurements significantly different from those tested with normal sense plates. Consequently, even a non-defective board tested with a fixture containing mechanical variations will have measurement values different from those from the reference fixture. We may draw incorrect conclusions when comparing the test results from the two fixtures. In this section, we will present analytical expressions representing such variations and discuss methods to compensate for the effects caused by DUT and fixture mechanical misalignments. However, to avoid having to use regression separately to identify the compensation equation coefficients for an individual device, the last section of this chapter presents a PCA based scheme that can be used to estimate the tilt or shift, and thereby generate coefficients for the compensation equations.

The goal of this section therefore is to identify and compensate for the measurement variation caused by abnormal sense plate positions so that it alone does not cause a DUT to be flagged as an outlier.

9.3. Effect on Measurements Due to a Tilted Sense Plate

The capacitance between sense plate and the connector pin is given by:

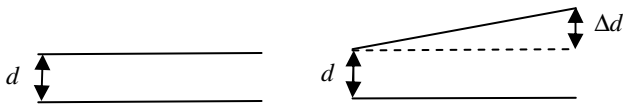
$$C = \frac{\varepsilon \times A}{d} \tag{9.1}$$

where d is the distance between connector and sense plate. Note that this relationship holds most accurately when A >> d^2, conditions that are stretched a bit in this case. However this relationship is a good first approximation for understanding what is happening.

When a tilted sense plate exists, the Δd value, which is the difference from original distance d, would change from minimum to maximum value linearly depending on the pin position. Here we consider only the variation along the length, although the analysis may be extended when such variation exists both along the length and width. The ratio CR of the capacitance values with and without tilt is approximately linear along the length as given below:

$$CR = (\frac{\varepsilon * A}{d}) / (\frac{\varepsilon * A}{d + \Delta d}) = \frac{d + \Delta d}{d} = 1 + \frac{\Delta d}{d} \tag{9.2}$$

With distance along the normal sense plate, Δd changes linearly with pin physical position on connector from left to right, see Figure 9-3. Equation (9.2) results in Equation (9.3) below where x is the normalized position along the connector:



(a) (b)

Figure 9- 3 Fixture with (a) normal sense plate (b) left tilted sense plate

$$CR(x) = A + Bx \tag{9.3}$$

For measurements with left tilted sense plate, where the right side of plate (near pins 120 and 240 as there are two rows) is lifted, the ratio, relative to the normally positioned measurement, should be a straight line which intercept at A ($\simeq$ 1) at y-axis, shown in Equation (9-3). For the right tilted measurement, the ratios would decrease from left to right.

The datasets used in this section includes measurement for boards with connectors termed *j3* and *j10*. Four boards of each type were involved, they were termed boards B1, B2, B3 and B5. Besides the normal test (identified as reference), each board was also tested under sense plate for 3 different right and left tilt angles and 3 different vertical shift heights. Thus there are a total of 10 different measurements for one connector on a single board. We analyze below in detail the left tilt and shift cases.

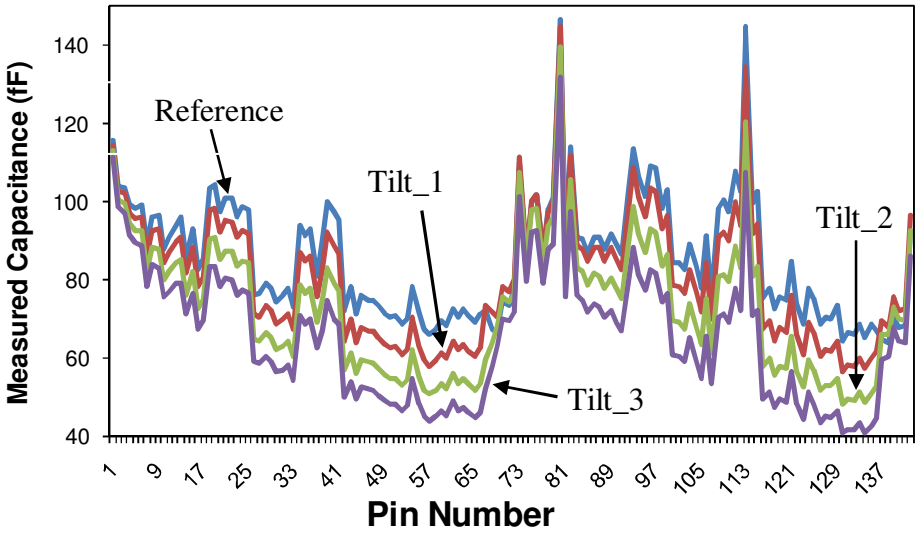
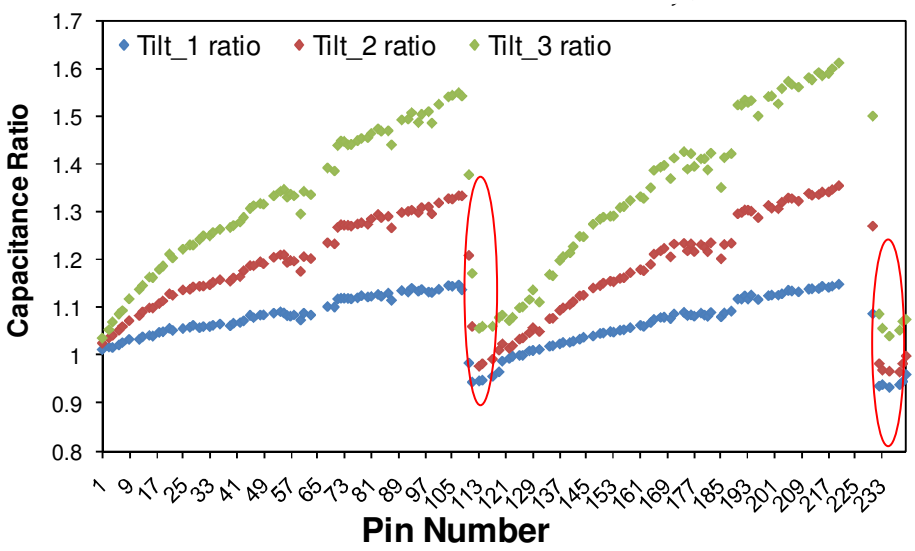
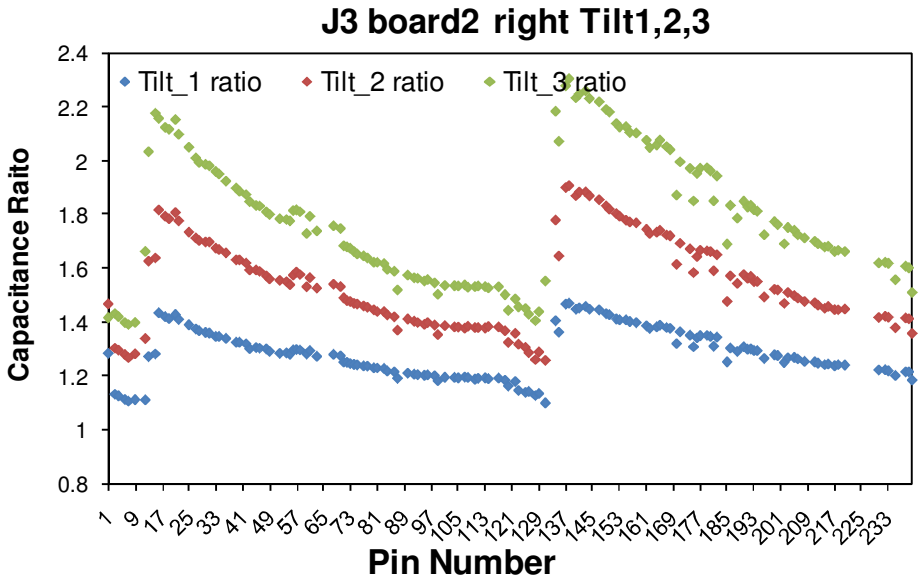


Figure 9- 4 Raw measurements plot for data B2 with one normal and three left tilted sense plates

Figure 9-4 above shows one example of j3 measurements, corresponding to the board B2. The original reference measurement is plotted with a blue curve; the others are plots of measurements Tilt_1, Tilt_2 and Tilt_3 with the tilted sense plate.



(a)



(b)

Figure 9- 5 Capacitance ratios in Data_j3 with (a) left tilted sense plate (b) right tilted sense plate

The relationship between normal measurement and tilted plate measurement, i.e., the Capacitance Ratios CR obtained by dividing the normal measurement by its respective tilted one, is shown in Figure 9-5. As indicated in Table 1, the slope of sense plate changes gradually resulting in Capacitance Ratios plots from Tilt_1 (blue) to Tilt_3 (green) in Figure 9-5. The amount of additional maximum Δd between each adjacent pair of curves is 8 mils (thousandths of an inch).

All the plots in Figure 9-5 follow Equation (9-3) fairly closely. The abnormal readings identified by the red ovals are caused by the paper spacers used in the experiment, which change the dielectric constant from that of air. To simplify analysis and to make it a more representative, these measurements are excluded in the following analysis.

9.3.1 Compensation with Trend Line for Tilted Plate Measurement

To find a method to compensate for tilt variation, measurements with highest added tilt (and hence the highest CR) were selected for analysis. This corresponds to Tilt_3 in Figure 9-4. The measurement ratios for the two connector rows are slightly different, which is likely to have been caused by the specific physical layout in the two rows. Thus, we need to plot the measurement ratios for two rows, row with pin 1 to pin 120 and row with pin 121 to pin 240, separately.

Figure 9-6 gives the CR plots for connector j3 in dataset B2 where the x-axis now gives the pin position. The data points are fitted using regression to the expression in Equation (9.3) above. The fit appears reasonable, not quite as well as expected. In reality, the complex distribution of the electromagnetic field makes determination of the actual capacitance value much more complicated than the simple plate capacitance model assumed. Note that the fit can be significantly improved if a third order polynomial is used for regression as shown in Figure 9-7. Data for the other boards also indicate similar results, i.e., the first order model is reasonably accurate, but third order model may be used when higher accuracy is needed. We plan to evaluate additional measurements to come up with the order of curves to be used to meet the test requirements.

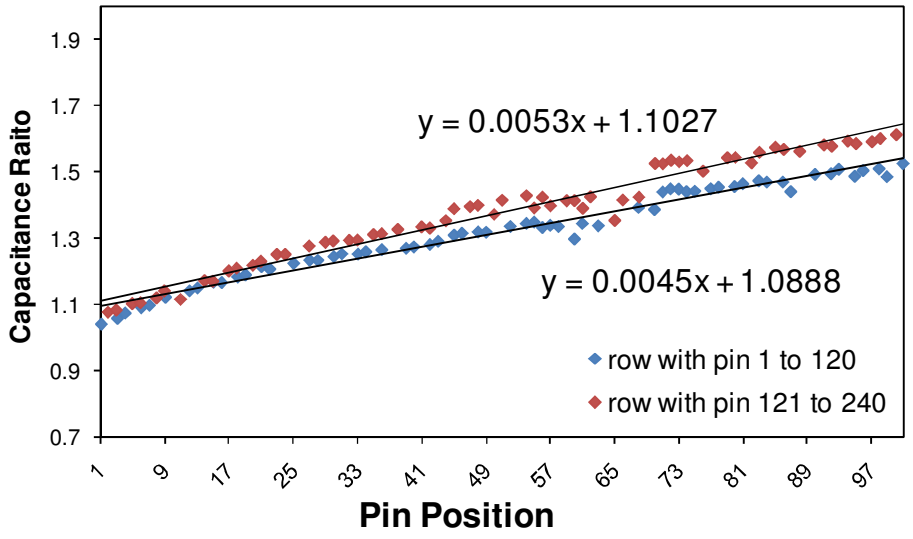


Figure 9- 6 Measurement ratio curves fitted with 1st order trend lines for connector j3 in dataset B2.

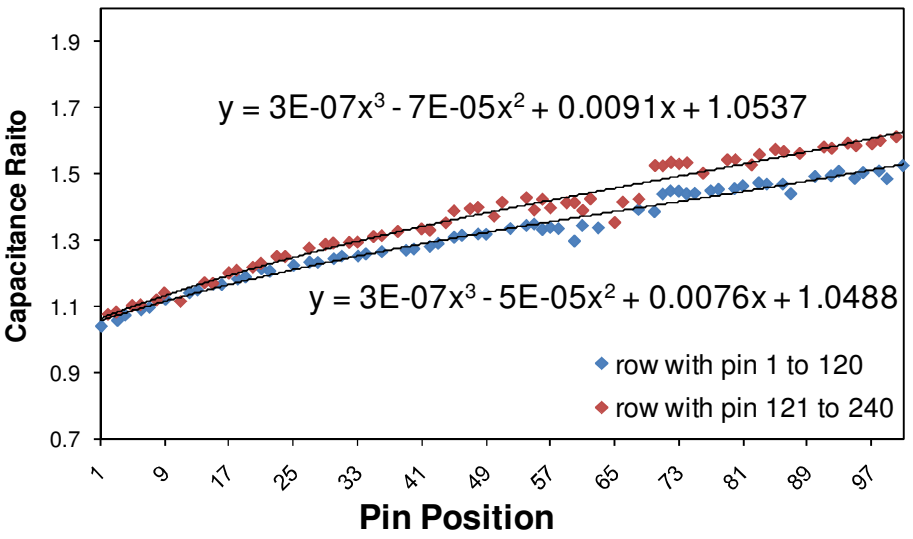


Figure 9- 7 Measurement ratio curves fitted with 3rd trend lines for connector j3 in dataset B2

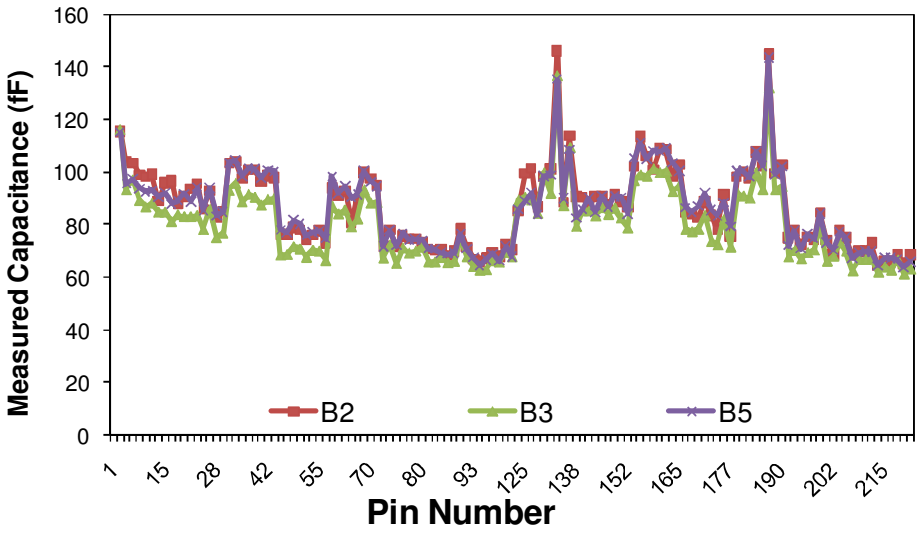
Figure 9-7, which fit measurement ratios for the two rows in connector are:

$$y = 3E\text{-}07x^3 - 7E\text{-}05x^2 + 0.0091x + 1.0537 \quad (9.4)$$

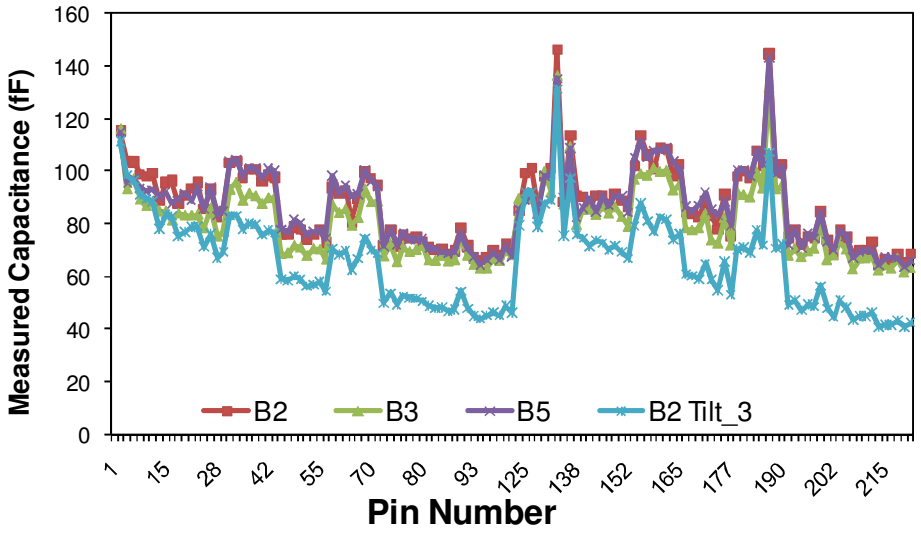
and

$$y = 3E\text{-}07x^3 - 5E\text{-}05x^2 + 0.0076x + 1.0488 \quad (9.5)$$

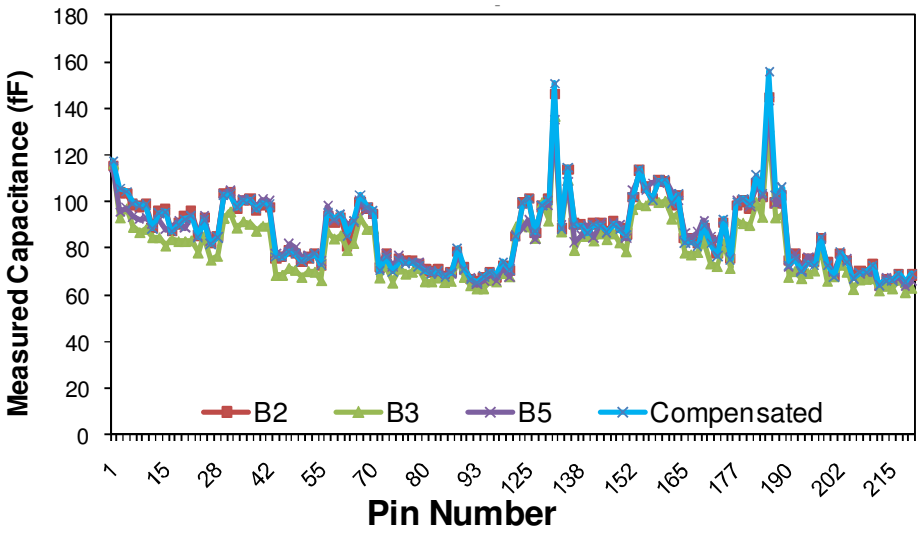
Our results indicate that Equations (9.4) and (9.5) are accurate for other boards of the same type for the corresponding amount of tilt [33]. Thus the regression analysis needs to be carried out only for one board with a particular tilt to be able to obtain the CR. Thus if the tilt for a board is known, we can apply the appropriate correction. Estimation of this tilt is addressed in Section 9.4. Figure 9-8 shows steps in compensating. The data set for evaluation consists of three boards as well as B2 measured with a tilted sense plate. With the Equations (9.4) and (9.5), the measurement at each pin of the abnormal measurement is multiplied by its ratio. The results for the three boards after compensation are shown in Figure 9-8 (c).



(a)



(b)



(c)

Figure 9- 8 Compensation flow to Data J3 B3 (a) raw measurement with part of pin measurements truncated (b) 3 truncated measurements and another one with tilted sense plate (c) plot of measurements after compensation

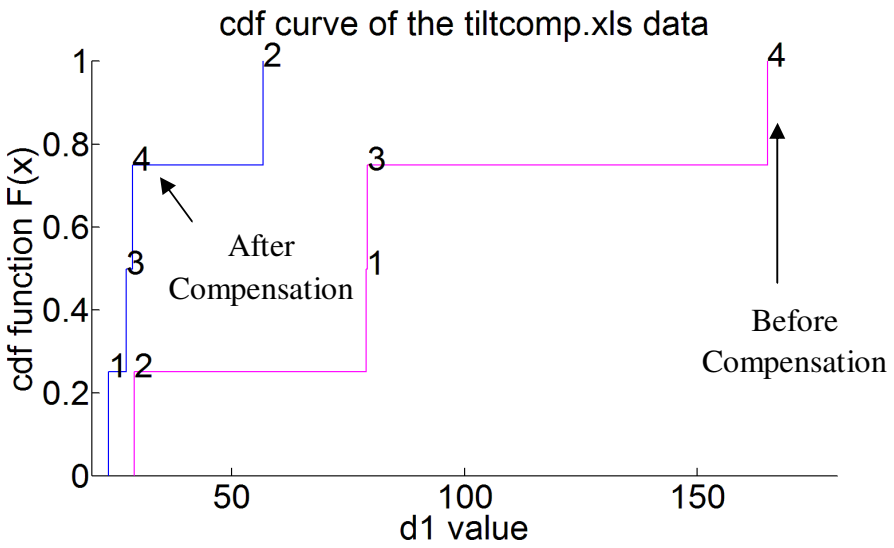


Figure 9- 9 CDF plot of Data J3 B3 before and after compensation

Figure 9-9 illustrates the results of PCA based outlier detection with and without compensation for tilt B3 for Data J3. Comparing the two CDF plots in Figure 9-9 we can clearly see the effect of compensation on the measurements. The d_1 value of boardrun 4 is decreased from 170 to less than 30 implying it is no longer an outlier.

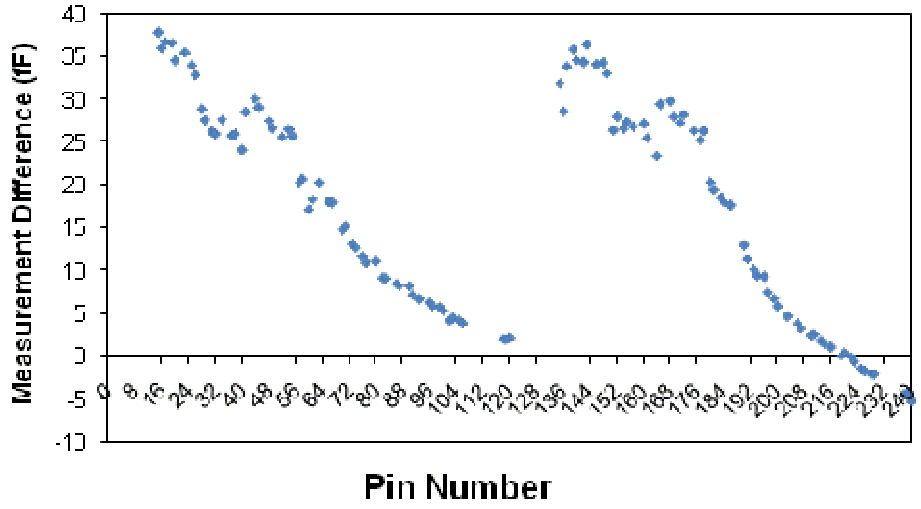
This indicates that the approach works. But we need to be able to estimate the tilt in order to be able to compensate for it without having to run regression analysis for each device.

9.3.2 Compensation with Difference Measurement Values

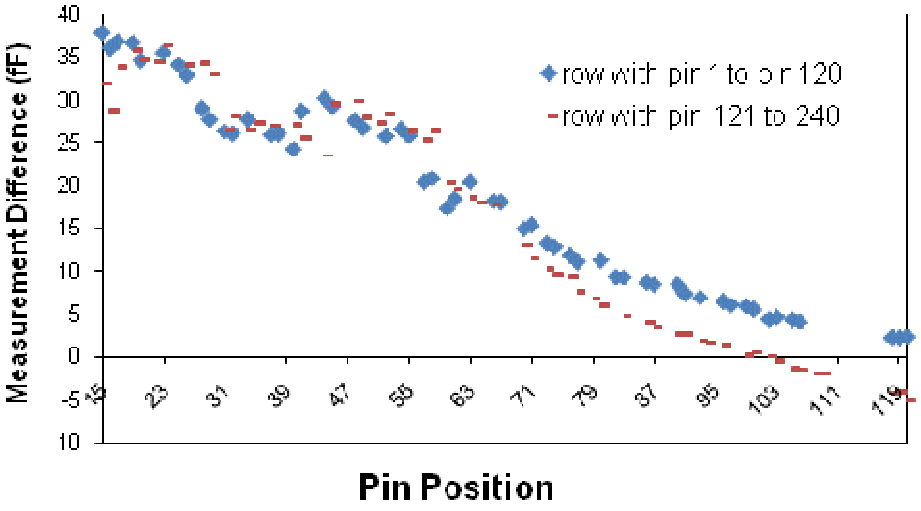
Method used in Section 9.3.1 utilizes ratio regression line (trend line) formula as the compensation. The method needs to get the regression line before getting the formula. For the abnormal ratio part, its accuracy would be decreased significantly.

The relationship between the tilted-sense-plate measurement and the normal measurement can also be obtained via the measurements difference. Figure 9-10 shows the difference value at each tested pin for measurements of connector j10 in dataset B2.

When the difference values are plotted row by row Figure 9-10 (a) is transformed into 9-10 (b), where the blue and red plots are the measurements to row 1 and 2 respectively.

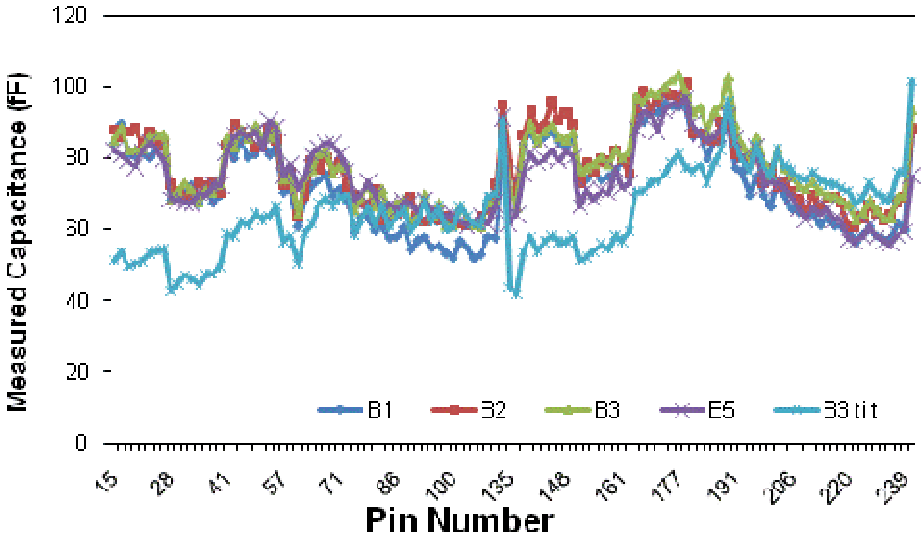


(a)

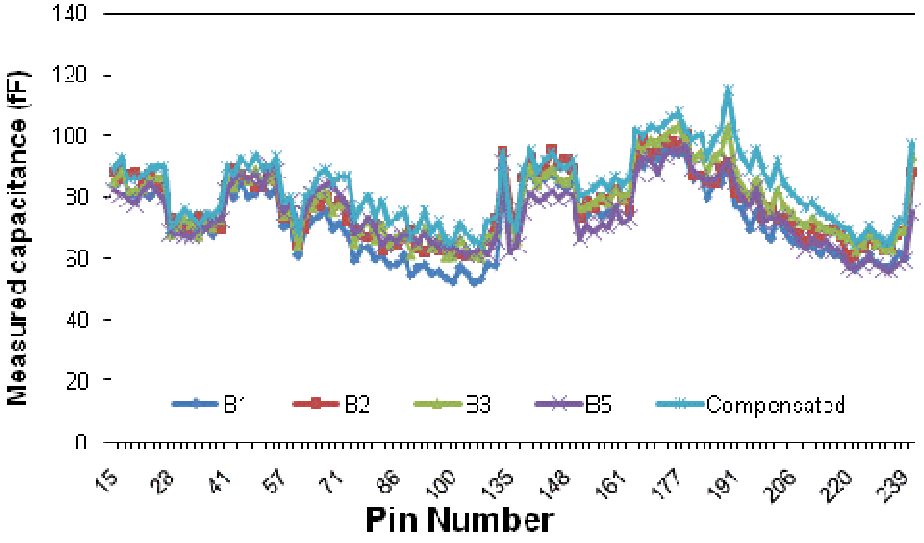


(b)

Figure 9- 10 Difference plot between tilt-measurement and normal measurement viewed from (a) pin 1 to pin 240 (b) row by row



(a)



(b)

Figure 9- 11 Measurement plot of Data tilted-B3 (a) before and (b) after compensation

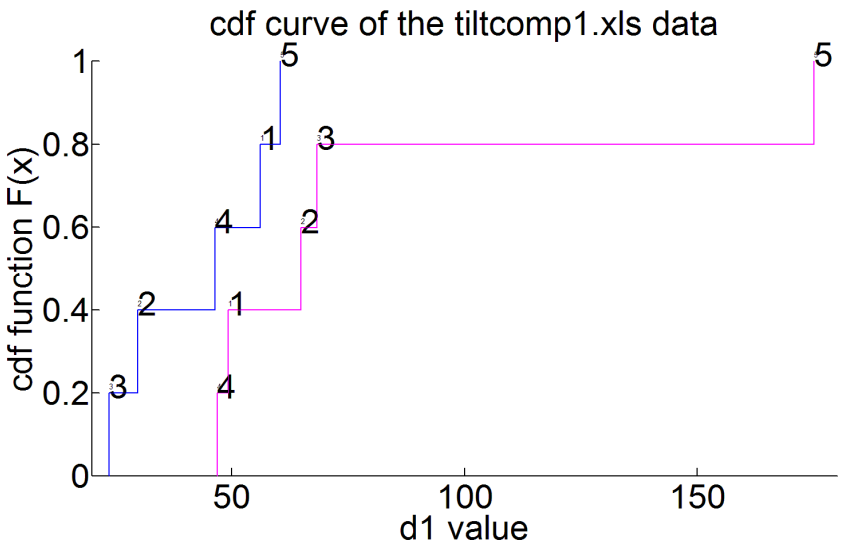


Figure 9- 12 CDF plot after difference value compensation for tilted B3

Figure 9-11 shows the compensation steps with the difference value. Each tilt_3 measurement in Figure 9-12 (a) add its own difference value plotted in Figure 10. Then the tilt_3 measurements are compensated. To show the compensation effect, Figure 9-12 gives the PCA results in CDF plots for measurements without compensation and with compensation. As can be seen, after the compensation, d_1 value of boardrun 5 is shrunk from 180 to less than 65 which imply that it is no longer an outlier.

Comparing the two methods shown in 9.3.1 and 9.3.2, we can see the second method decreases maximum d_1 value more effectively and efficiently.

9.4. Effect on Measurements Due to a Vertical Shift of Sense Plate

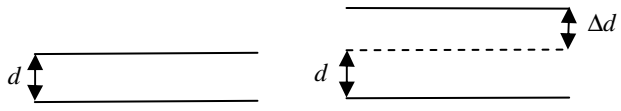


Figure 9- 13 Connector with (a) normal positioned sense plate (b) vertical shifted sense plate

Normal and vertically shifted sense plate diagrams are plotted in Figure 9-13 (a) and (b). The vertical mechanical variation of sense plate is caused by variability in the height of the plate above the pin tips in the connector. Variations of 8, 16, and 24 mils were inserted.

Figure 9-14 shows the measurements of Data B5 in the LVLD3 for connector j3 with 4 different sense plate heights. The sense plate shifts vertically gradually. The height introduced in sense plate and connector are 0, 8, 16, 24 mils. Purple curve is the measurement with largest vertical height between connector and sense plate.

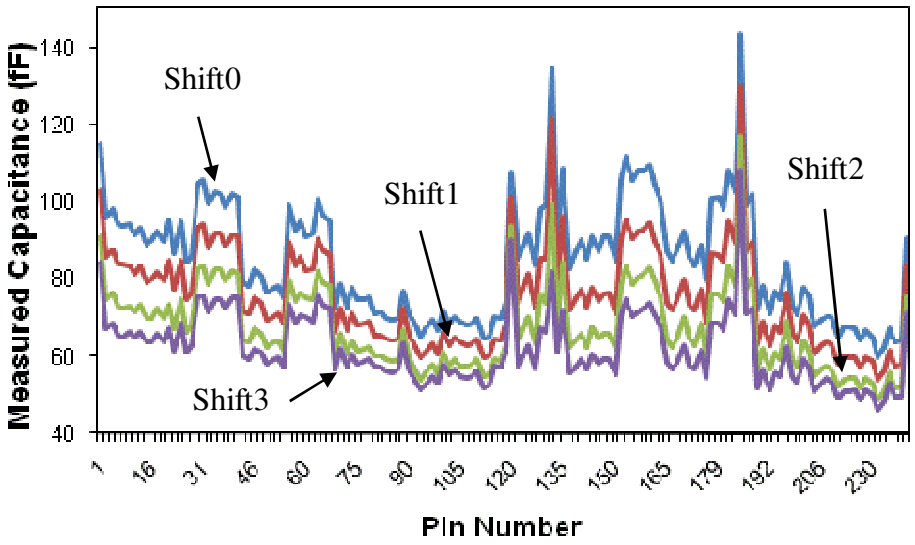


Figure 9- 14 Plot of measurements with normal and vertical shifted sense plate in dataset B5 for connector j3

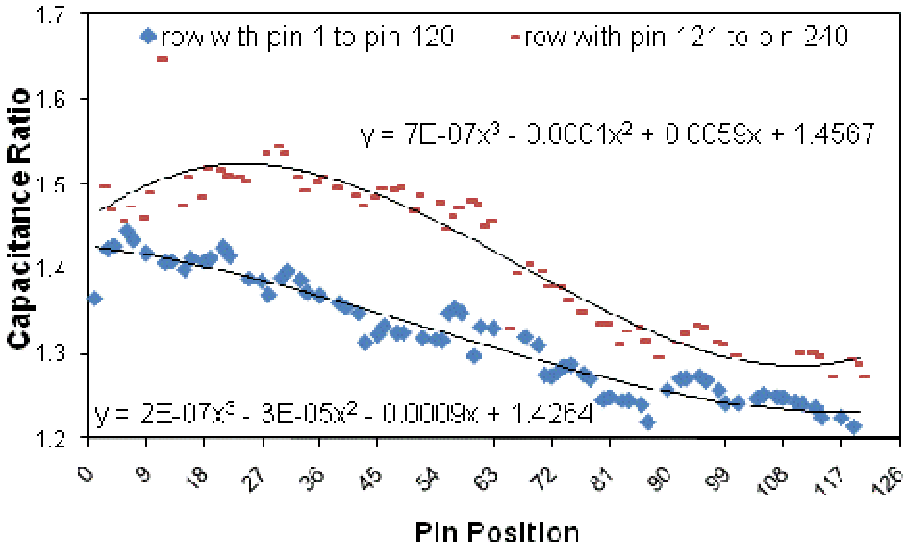


Figure 9- 15 Measurement ratio curves fitted with 3rd order trend lines for connector j3 in dataset B5.

Figure 9-15 shows the capacitance ratio plot to the two row of connector j3. As can be seen the 1st order regression will be far from accurate in this case. The compensation and analysis steps are similar to those in Figure 9-8 and Figure 9-9 which will be ignored here.

9.4.1 Compensation for Vertical Shifted Sense Plate with Difference Values

Based on formula (9.1) we can derive measurement difference between normal and abnormal measurement at each pin shown in Equation (9.6).

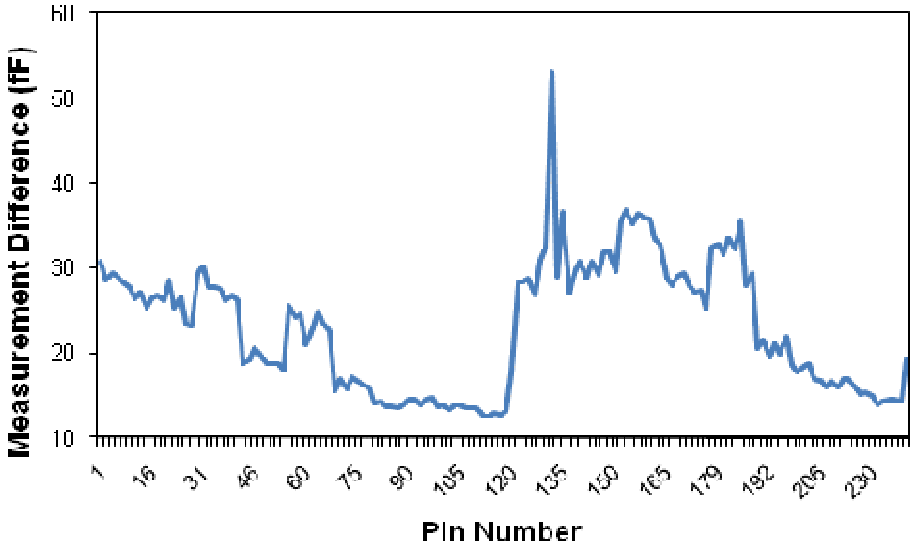
$$Difference = (\frac{\varepsilon * A}{d}) - (\frac{\varepsilon * A}{d + \Delta d}) = \frac{\varepsilon * A * \Delta d}{(d + \Delta d) * d} \quad (9.6)$$

For each case, Δd is constant for each pin on the connector.

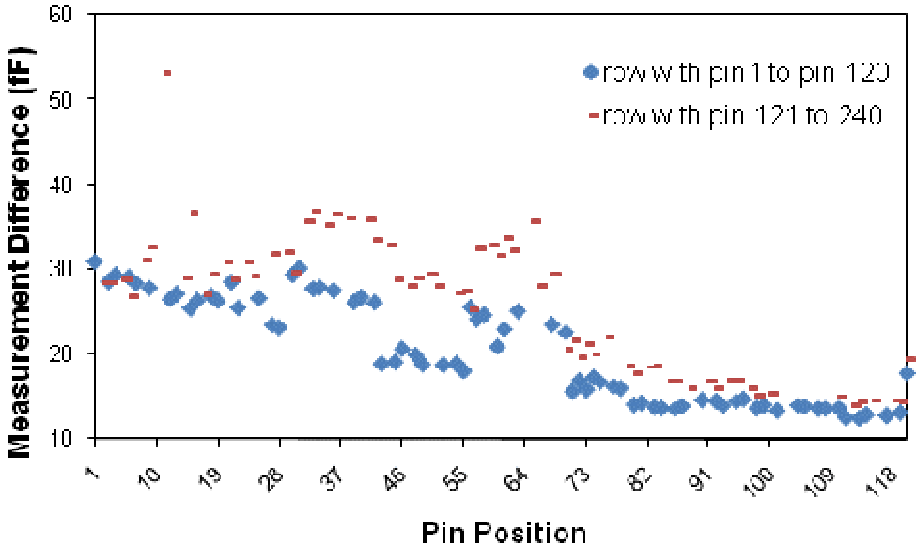
Since d and Δd value won't change along the pin location, difference value should be a constant. Then Equation (9.6) can be simplified as Equation (9.7)

$$CD(x) = K \tag{9.7}$$

where K is a constant.



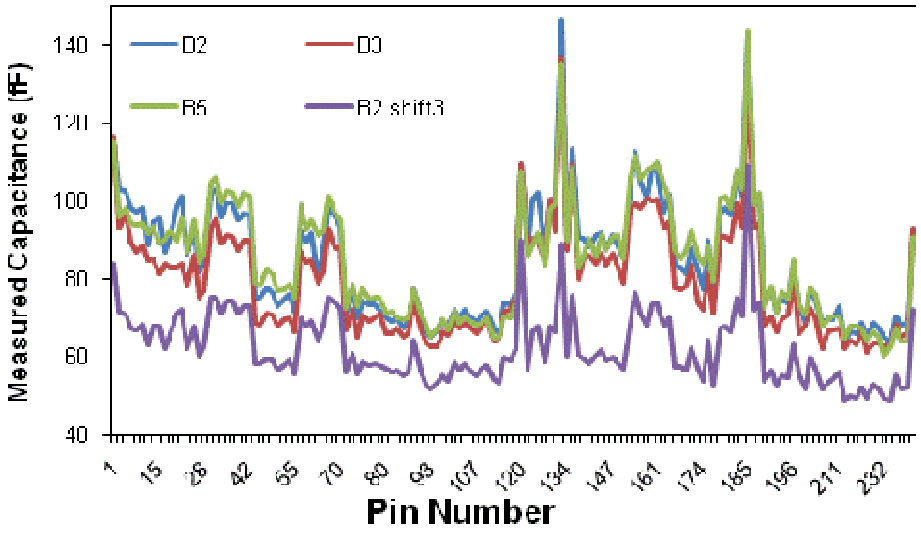
(a)



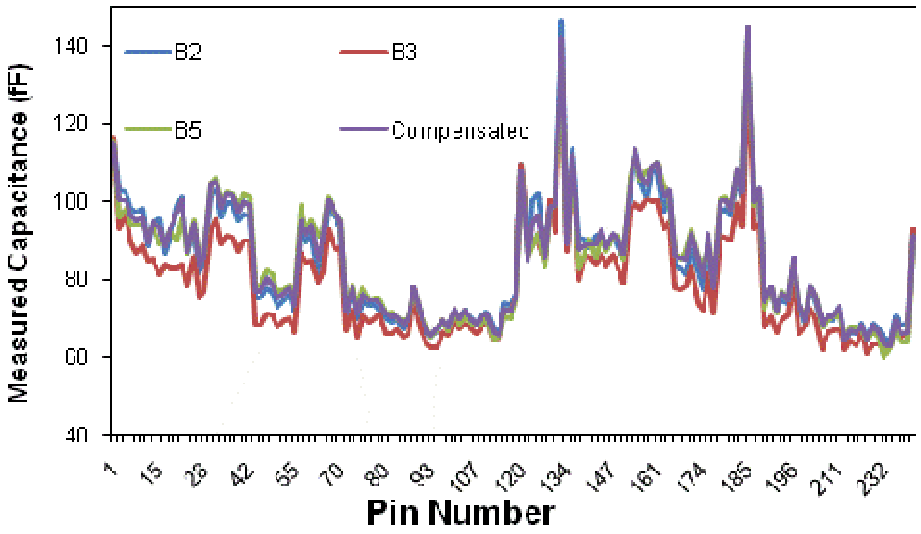
(b)

Figure 9- 16 Plot of difference value between measurements with normal and vertical shifted sense (a) in whole view (b) row by row

Figure 9-16 shows measurement difference between those with normal and vertical shifted sense plate. With the difference values at each pin, we will be able to apply compensation to the dataset B2. Figure 9-17 (a) shows three normal measurements and one with a shifted plate. Each pin measurement in the shifted-plate-measurement (the purple curve) adds its own measurement difference value in Figure 9-16. Then compensating for measurement difference, we got the Figure 9-17 (b), which is much more compact than that in Figure 9-17 (a).



(a)



(b)

Figure 9- 17 Measurement plot before and after compensation (a) plot before compensation (b) plot after compensation

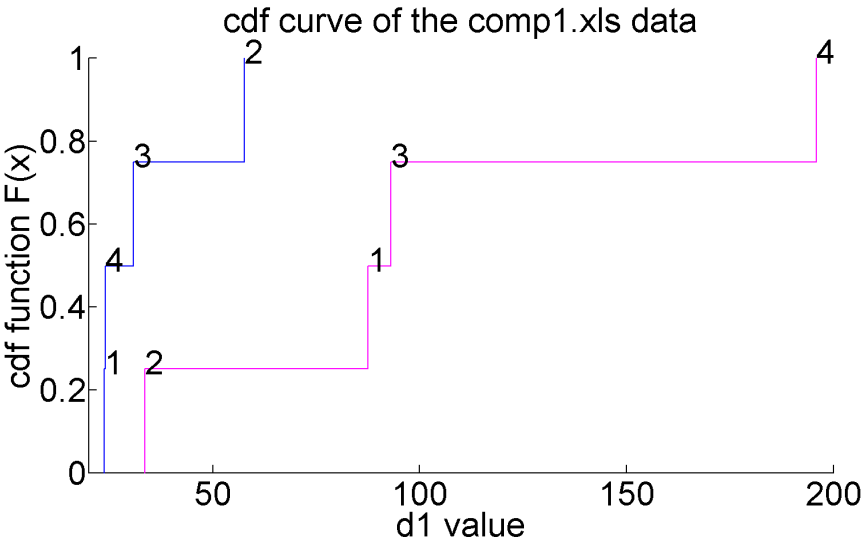


Figure 9- 18 CDF plot before and after compensation

The comparison between two CDF plots in Figure 9-18 clearly shows the validity of distance compensation method. Maximum d_1 value is shrunk from 198 to 58.

9.5. PCA Based Tilt and Shift Evaluation

The objective of this section is to investigate how to identify the amount of tilt in a sense plate so that parameters for the compensation curve can be derived. For this discussion, consider four different sense plates, each with a different tilt. By measuring the same board using these different sense plates, we show that we can identify the degree of tilt of different plates.

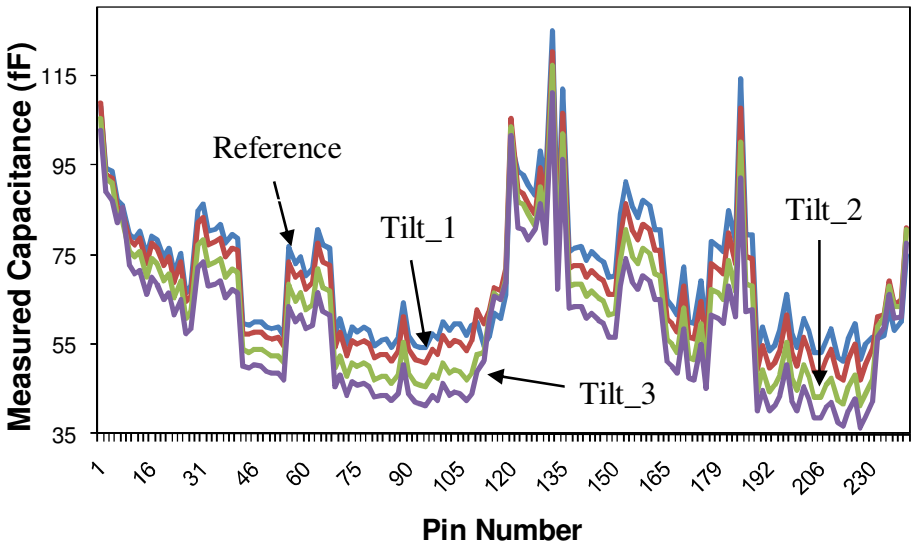


Figure 9- 19 Plot of raw measurement of data B1 with three left tilted plated measurements

Figure 9-19 shows measurements plot corresponding to board B2 including reference case and for three values of tilt. We next evaluate the Principal Components corresponding to this measurement set. The PC plot in Figure 9-20 shows the values of different PCs for the four cases. Only 4 of them are visible as the remaining PCs are negligible in this case. Note that the value of 1st PC, in Figure 9-20, is almost linearly proportional to the tilt corresponding to the board.

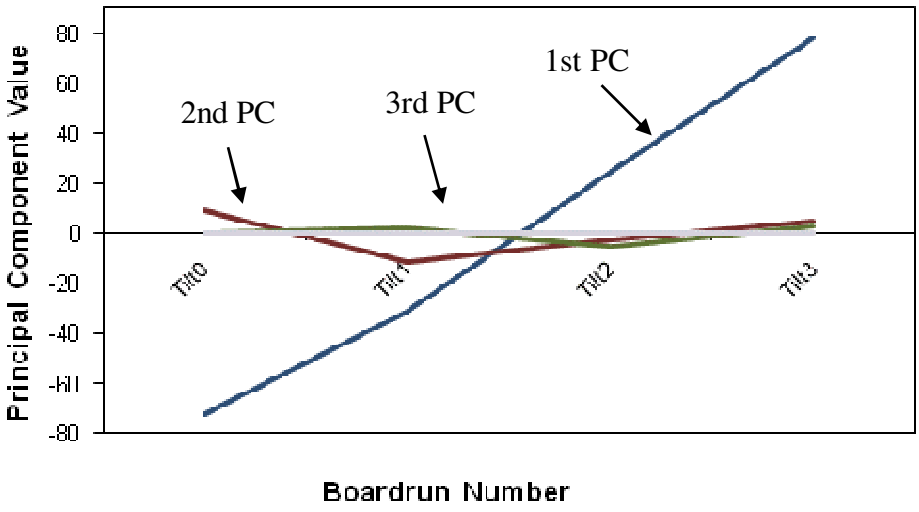


Figure 9- 20 Plot to principal component values to data B1

The same phenomenon is observed on measurements with vertical shifted sense plate. Figure 9-21 is the measurements plot to data B2 with vertical shifted sense plates. The height introduced in sense plate and connector are 0, 8, 16, 24 mils. Again, it is clear that the first PC is able to identify the amount of shift in the sense plate as shown in Figure 9-22. Note that the PC coefficients are different in the case of tilt from those for shift, as the PC coefficients are data dependent. Also it is interesting to realize that the PC calculation ignores the order of the pins, yet is able to identify the amount of shift or tilt from the first component. Human observations, on the other hand, depend on the order of the pins to recognize whether shift or tilt is present. For example, if you shuffle pin numbers randomly for data shown in Figure 9-19, a human observer is unlikely to detect the pattern corresponding to a shift or tilt.

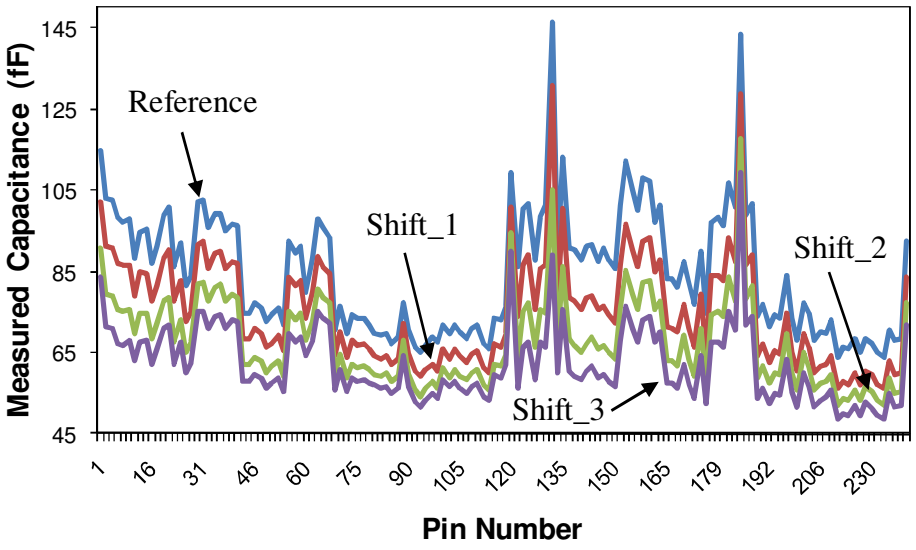


Figure 9- 21 Raw measurement of data B2 with three vertical shifted plated measurements

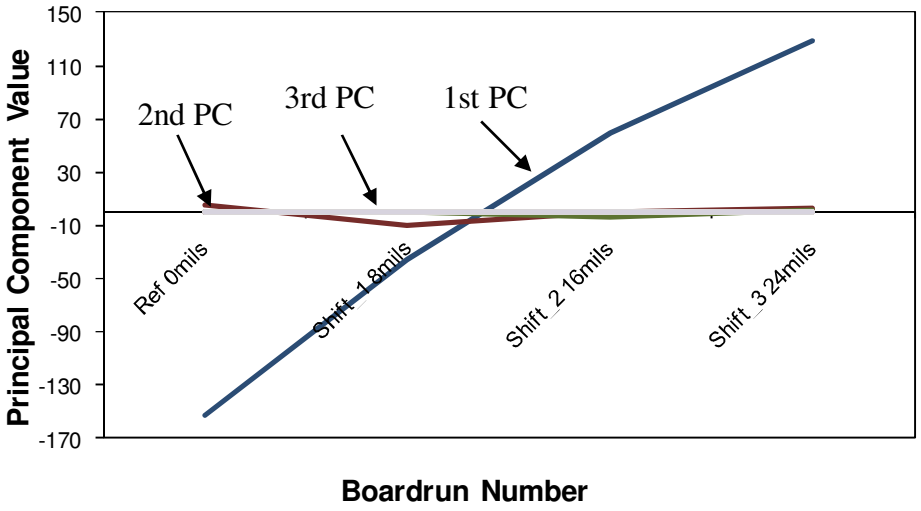


Figure 9- 22 Plot to principal component values to data B2

The misalignment of the sense plate can diminish the accuracy of outlier identification using PCA for TestJet-based board testing. Approaches for compensating the misalignment are examined in this study. Test data was presented to indicate the

impact of tilt and shift type misalignments. Evaluation of data obtained using tilted sense plates was analyzed in detail to show that compensation involves multiplication of the data by a polynomial. First order polynomial is likely to be adequate, but a higher order polynomial can be used when higher accuracy is needed.

A PCA based method that can be used to identify the degree of tilt (or shift) present in a sense plate is presented. The matrices needed for the calculation can be evaluated based on data obtained by testing the same board using sense plates with different tilts. Estimation of accurate tilt information will allow the calculation of coefficients for compensation polynomials. This scheme needs to be further developed and evaluated using additional test data.

Chapter 10.
Measurement Error Distribution

In this Chapter, we plot and analyze the distribution of the test measurements from single board and multiple boards. The purpose is to fit these distributions with specific types, which could be applied for predictions for future analysis.

10.1. Repeated Measurements Distribution

Figure 10-1 shows 50 repeated measurements to board 1 in LVLD2 J3 dataset. Although the difference among the measurements is small, error and percentage error values at each tested pin are plotted in Figure 10-2 (a) and (b).

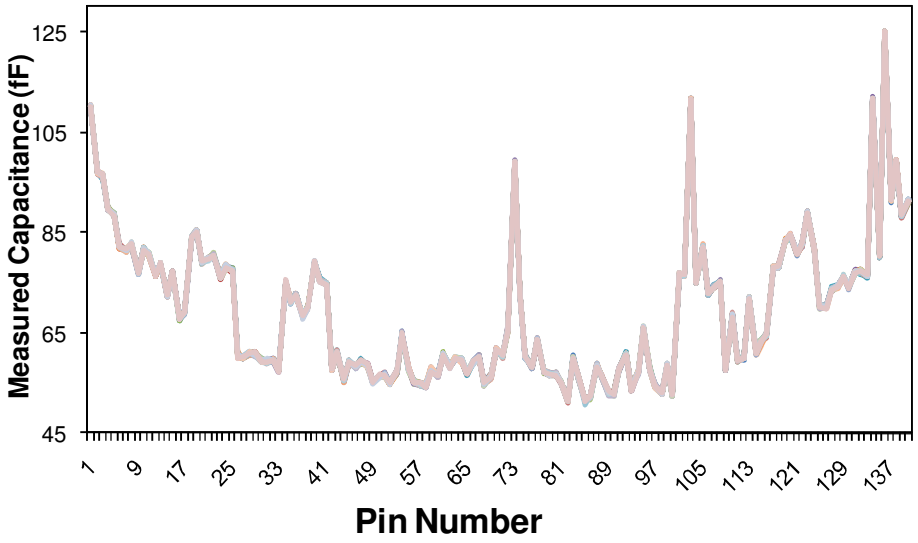
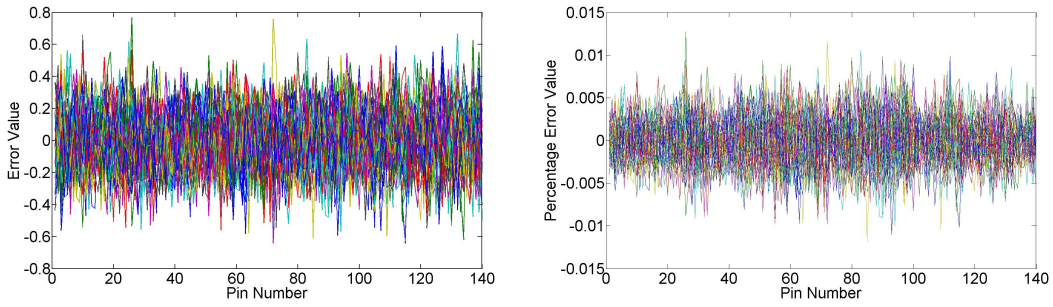


Figure 10- 1. Raw measurement plot of the 50 repeated measurements in LVLD2 j3



(a) (b)

Figure 10- 2. Plot of the 50 repeated measurements with (a) measurement errors (b) percentage measurement errors

The error value is the distance of the measurement value from the average measurement value at each pin. When divided by the average value, the measurement error changes to percentage error value, as shown in Figure 10-2 (b).

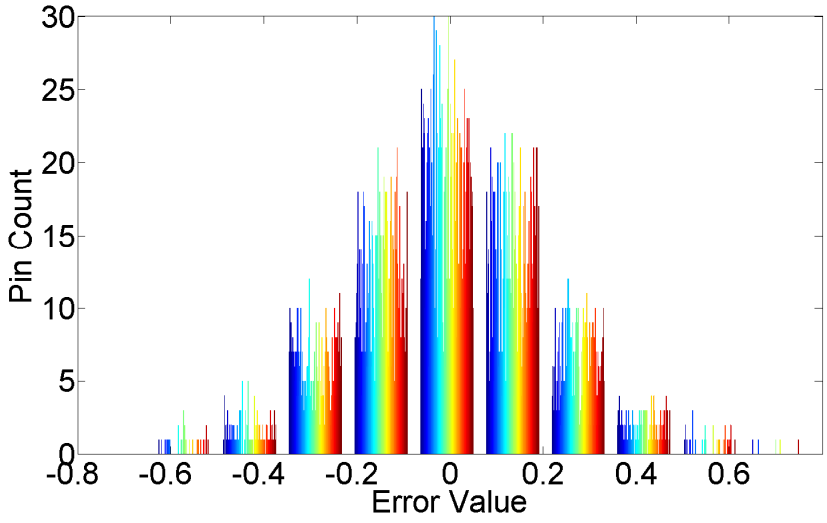


Figure 10- 3. Histogram of error values from the 50 repeated measurements

Figure 10-3 and Figure 10-4 plot the distribution of the error values and percentage error values to the 50 repeated measurements (boardruns). Each color in Figure 10-3 represents a distribution of a boardrun. Then in Figure 10-3 and Figure 10-4, the 50 boardruns leads to 50 different colors.

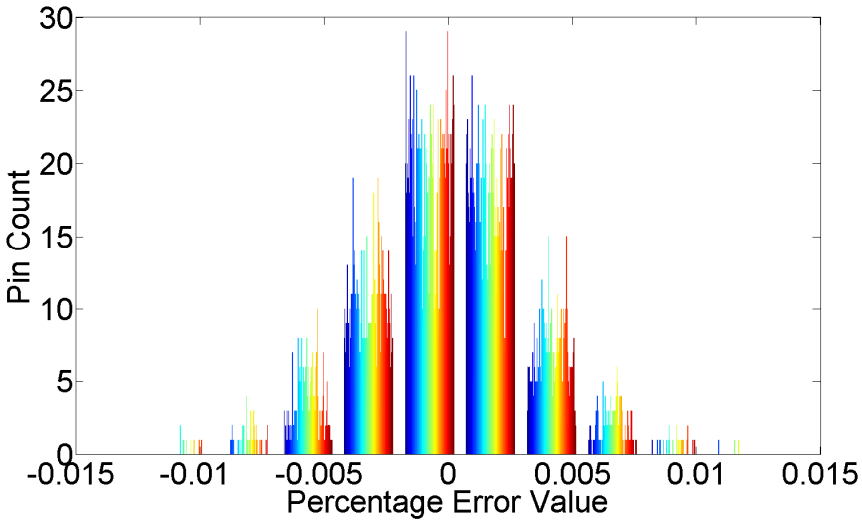
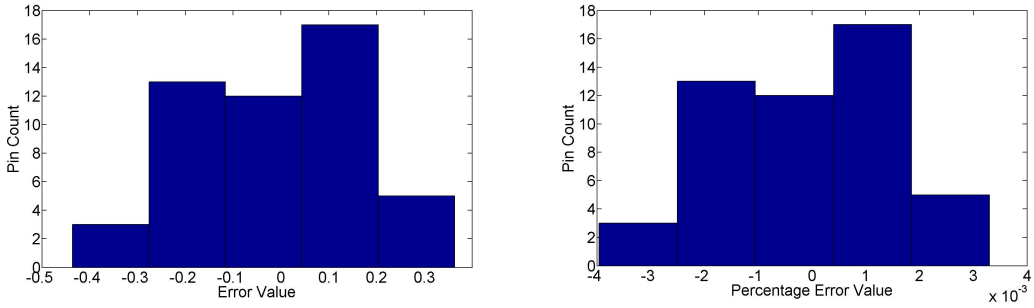


Figure 10- 4. Histogram of percentage error values for the 50 repeated measurements

From what is illustrated in Figure 10-3 and Figure 10-4, the error and percentage error distribution at each boardrun is close to normal, although the distribution bell shapes are not exactly same. At the left and right tail of the two figures, only a few boardruns show values there, which means the distributions show lower height and larger width.

Figure 10-5 and Figure 10-6 give error and percentage error value distributions for pin 1 and pin 140 respectively.

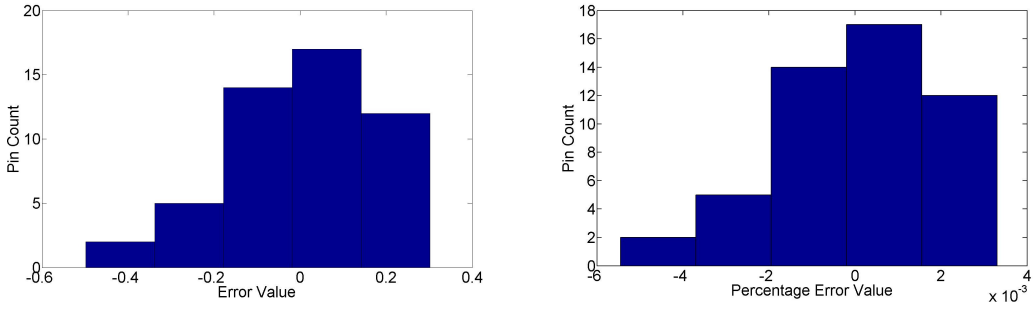
Pin 1



(a) (b)

Figure 10- 5. Plot of the pin 1 measurements with (a) measurement errors (b) percentage measurement errors

Pin140



(a) (b)

Figure 10- 6. Plot of the pin 140 measurements with (a) measurement errors (b) percentage measurement errors

The histogram plots in both Figure 10-5 and 10-6 show the similarity to normal distribution, although the distributions of pin 140 measurements are skewed to the right side.

10.2. Non-repeated Measurements Distribution

Figure 10-7 plots 41 measurements from 41 different boards. Figure 10-8 (a) and (b) show the distribution of error and percentage error values at each boardrun.

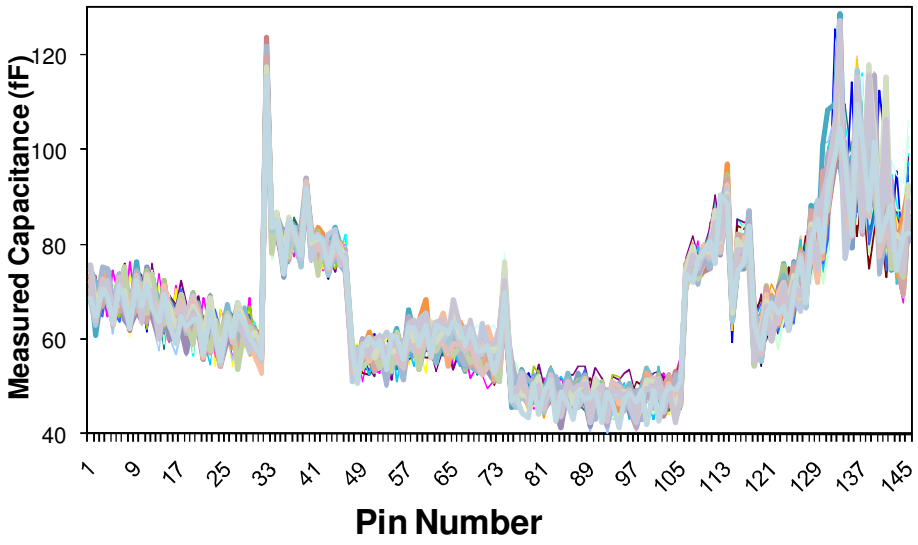
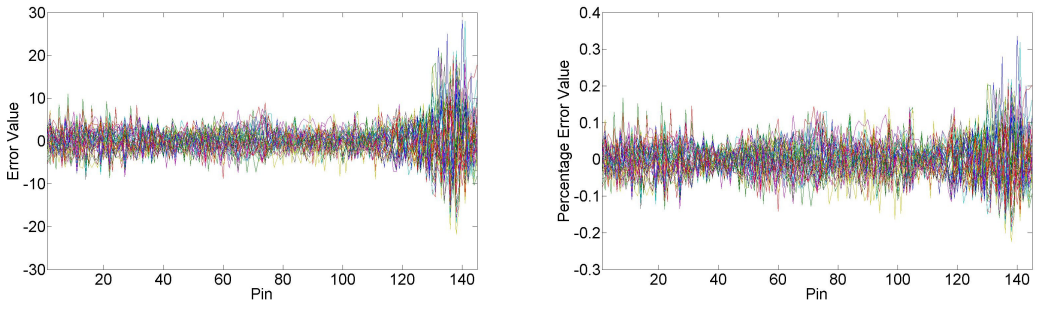


Figure 10- 7 Raw measurement plot of different connectors in Data3_j24



(a) (b)

Figure 10- 8. Plot of the pin 140 measurements with (a) measurement errors (b) percentage measurement errors

In Figure 10-9 and Figure 10-10, each colored distribution is based on measurement values from a single boardrun. The error distribution and percentage error distribution are still close to normal.

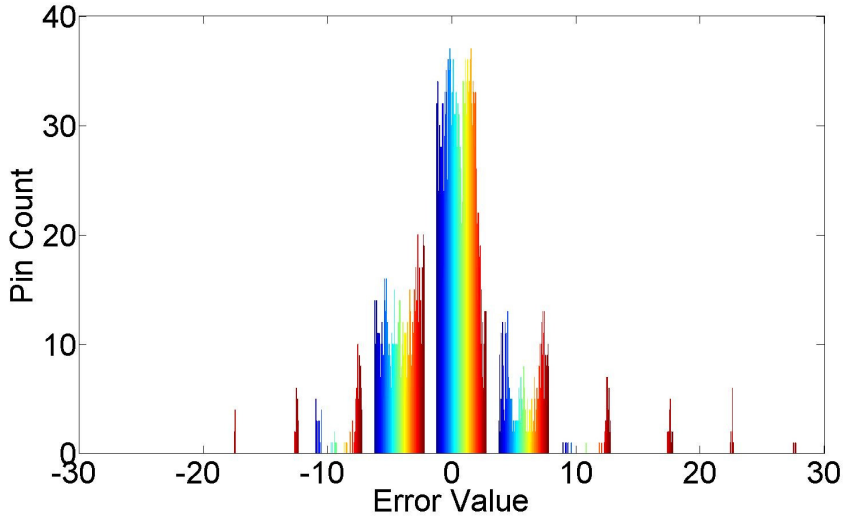


Figure 10- 9. Histogram of error values from different connector measurements in Data_j24

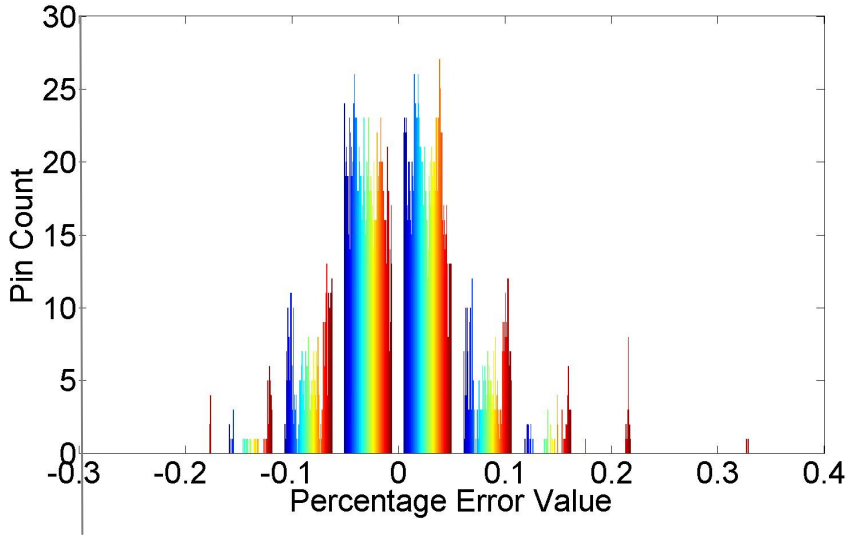
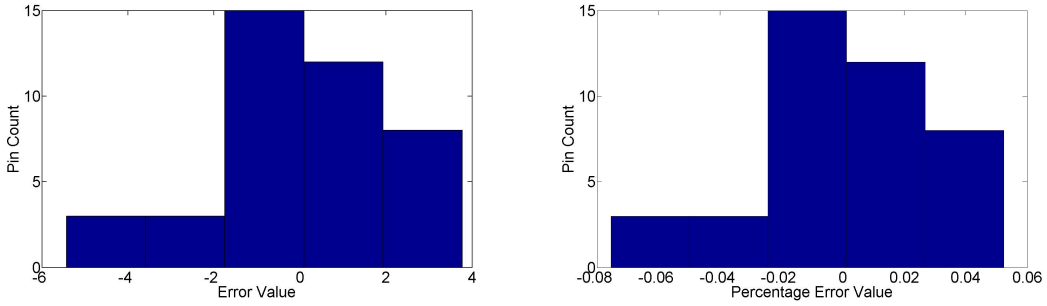


Figure 10- 10. Histogram of percentage error values from different connector measurements in Data_j24

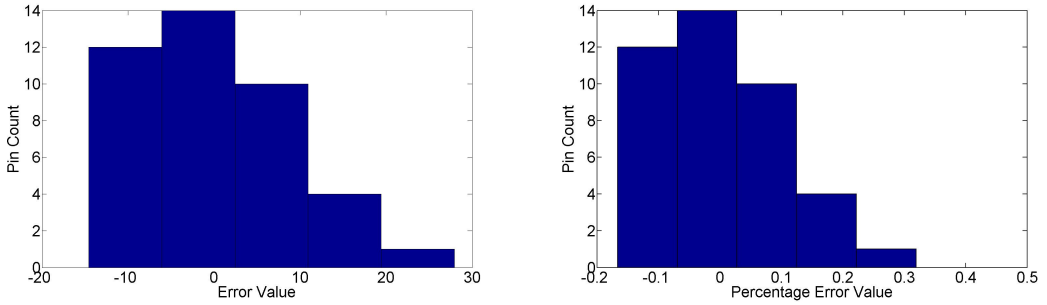
Pin1



(a) (b)

Figure 10- 11. Plot of the pin 1 measurements in Data3_24 with (a) measurement errors (b) percentage measurement errors

Pin 141



(a) (b)

Figure 10- 12 Plot of the pin 1 measurements in Data3_24 with (a) measurement errors (b) percentage measurement errors

The distributions of error value and percentage value at pin 1 and pin141 show much difference, which can be detected from Figure 10-11 and Figure 10-12. Pin 1 measurements distribution in Figure 10-11 are skewed to the right side, while those in Figure 10-12 are skewed to the left side.

The non-repeated measurement error and percentage error distribution of non-repeated measurements don't show much similarity to those of the repeated

measurements. This is because of the variation among different boards is much larger than that among measurement variation.

Chapter 11.
Conclusions and Future Work

This thesis presented a novel PCA-based technique for automated identification of outlier printed circuit boards which can potentially be used to identify faulty and/or marginal boards [20]. We have identified the test statistic d_1 with the most significant principal components to be the most robust for detecting the outliers in PCBs. The technique treats the data in a holistic manner to identify the outliers. We also presented a localized analysis technique, which applies PCA to a smaller window of pins. It can increase the test resolution, and can assist in the location of some outliers. Since both the global and localized analysis have some trade-offs in the face of different types of outliers, they can be combined to effectively classify and filter the outliers. Global Analysis is able to identify the trends affecting a large fraction of measurements. The proposed PCA-based approach has been compared with the traditional standard deviation-based approach for identifying outliers. The PCA-based approach significantly outperforms this traditional approach. Approaches for compensating for mechanical misalignment are proposed and evaluated as well.

Techniques are investigated to enhance the effectiveness of the outlier detection, and related computational issues for on-line testing are addressed.

We are also conducting further research on the effects of common types of mechanical variations due to sense plate mechanism and connector alignment to see what these effects may be and whether PCA can identify these variations systematically. These variations (such as shifted plates) cause measurement variations that are not indicative of actual defects. The results presented indicate that PCA based analysis can be used to

predict the amount of tilt or the amount of shift. Further experimentation however is necessary to deal with more practical cases where both tilt and shift are present. Future work will thus involve to the investigation of effects and methods to overcome the variations caused by the use of different testers, and measurement errors introduced by mechanical and electrical tolerances. Such techniques would be required by the future denser and faster PCB technologies to minimize the probability of a good board being identified as marginal or bad even when higher board densities reduce the available resolution.

References

[1] D. Gizopoulos, “Advances in Electronic Testing: Challenges and Methodologies,” Springer-Verlag, 2006, Page 389-391.

[2] K. P. Parker and S. Hird, “Finding Power / Ground Defects on Connectors – A New Approach,” Proceedings, International Test Conference, Santa Clara, CA, Oct. 2007.

[3] S. Hird and R. Weng, “Finding Power/Ground Defects on Connectors – Case Study,” Proceedings, International Test Conference, Santa Clara, CA, Oct. 2008.

[4] A. Sharma, A. P. Jayasumana, Y. K. Malaiya, “X-IDDQ: a novel defect detection technique using IDDQ data,” Proceedings, IEEE VLSI Test Symposium, Berkeley CA, May 2006.

[5] P. M. O’Neill, “Production Multivariate Outlier Detection Using Principal Components”, Proceedings, IEEE International Test Conference, Santa Clara CA, Oct. 2008.

[6] A. S. Banthia, A.P.Jayasumana and Y.K. Malaiya, "Data Size Reduction for Clustering-Based Binning of ICs using Principal Component Analysis," Proc. IEEE International Workshop on Defect Based Testing (DBT 2005), CA, May. 2005

[7] A. Nahar, R. Daasch, “Burn-in reduction using principal component analysis,” Proceedings. IEEE International Test Conference, Austin, TX, Nov 2005.

[8] T. Turner, “Capacitive Lead Frame Testing,”Proceedings, International Test Conference, Washington DC, Oct 1996.

[9] J. E. Jackson, " A User's Guild to Principal Components", Wiley, 2003

[10] L. Elden, "Matrix Methods in Data Mining and Pattern Recognition (Fundamentals of Algorithms," Society for Industrial and Applied, 2007.

[11] I. T. Jolliffe, "Principal Component Analysis," Springer-Verlag, 2002, page 233-248.

[12] T. Hastie, R. Tibshiraani, J. Friedman "The Elements of Statistical Learning" , Springer,2001,page 62-64.

[13] http://img.hexus.net/v2/motherboards/intel/intel/DX48BT2/DSCF6366-copy-big.jpg

[14] P. Praus, "SVD-based Principal Component Analysis of Geochemical Data," Central European Journal of Chemistry, 2005, page 731-741.

[15] K.P. Parker and N.S. Devnani, "Methods and Apparatus for Diagnosing Defect Locations for Electrical Paths of Connectors of Electrical Assembles" US Patent Application Publication, PUB No. US2005/0099186 A1, May 12, 2005

[16] D. Gizopoulos, "Advances in Electronic Testing: Challenges and Methodologies," Springer-Verlag, 2006, Page 308-332.

[17] D. Hand, H. Mannila, P. Smyth, "Principle of data mining," The MIT. Press, Cambridge, London, 2001

[18] E.R. Malinowski "Factor Analysis in Chemistry", 2nd ed., John Wiley & Sons, New York, 1991.

[19] A. Sharma, "X-IDDQ: a novel defect detection technique using IDDQ data," Master Thesis, Colorado State University, 2006.

[20] X. He, Y. K. Malaiya, A. P. Jayasumana, K. P. Parker and S. Hird, "An Outlier Detection Based Approach for PCB Testing," Proc. IEEE international Test Conference, Austin, Texas, November 2009.

[21] N.S. Tlale. "Mechatronic Modular Design of an Apparatus for PC board Inspection" Proc. Australasian conference on Robotics and Automation, Auckland , Nov.2002

[22] **"**IEEE Std 1149.1 (JTAG) Testability Primer" Chapter 3

[23] L. E. Frenzel, " The Embedded Plan For JTAG Boundary Scan" Electronic Design, Nov. 2008

[24] R. Oshana, "Introduction to JTAG". *Embedded Systems Design,* October 29, 2008

[25] **"**Eichelberger, Edward, "Method of Level Sensitive Testing a Functional Logic System" , US 3761695, issued 9/25/1973

[26] D. Sumit, "Level sensitive scan design (LSSD) system", US 4293919, issued 10/06/1991

[27] S. Kannan, B. K. Kim, "Automatic Diagnostic Tool for Analog-Mixed Signal and RF Load Boards," Proc. IEEE International Test Conference, Austin, pp 1-1. Nov 2009.

[28] K. P. Parker , "Implications of 3-D Integrated Circuits at Board Test Position Paper," Proc. IEEE International Test Conference, pp 1-2 Austin, Nov 2009.

[29] C. L. Tee, T. H. Tan and C. C. Ng, "Augmenting Board Test Coverage with New Intel Powered Opens Boundary Scan," Proc. IEEE International Test Conference, pp 1-10 Austin, Nov 2009.

[30] B. F.P.M. “Macro Testing Unifying IC and Board Test”, IEEE Design & Test of Computers 1986

[31] C. Lotz. “Functional Board Test-Coverage Analysis: what does it mean when a functional test passes?” , European Board Test Workshop, Southampton, UK,2006

[32] “White Paper: Comparing Board Test Strategies, ” International Test Technologies.

[33] X. He, Y. K. Malaiya, A. P. Jayasumana, K. P. Parker and S. Hird, " Principal Component Analysis-Based Compensation for Measurement Errors Due to Mechanical Misalignments in PCB Testing" Proc. IEEE international Test Conference, Austin, Texas, Oct. 2010.

[34] R. A. Horn and C. R. Johnson, Matrix Analysis, Cambridge University Press 1985.

[35] D. Hand, H. Mannila, P. Smyth, Principles of Data Mining, The MIT Press, 2001.

[36] S. Hyv¨onen, Linear Algebra Methods for Data Mining - Lecture 5, University of Helsinki, Spring 2007.

Appendix

```
function singval_d(p,q,number,data)
clc;
p
q
subm=2;
subn=2;
pos=subm*subn;
AA = xlsread(data,1,'A2:E121')%%%%change with different data
Ar=AA';
[r,c] = size(Ar)
B=[1:r];
An=Ar-repmat(mean(Ar),r,1);
A=An;
qq=c+2+number;  %
[U,S,V] = svd(A,0); % "economy size" decomposition
newdata = A*V
Loca=strcat(pwd,'\','score_matrix.xls');
xlswrite(Loca,newdata);
newdata(:,1)=0
Loca=strcat(pwd,'\','New_score_matrix.xls');
xlswrite(Loca,newdata);
M_new=newdata*inv(V);
Loca1=strcat(pwd,'\','new_raw_matrix.xls');
xlswrite(Loca1,M_new);
variances = diag(cov(newdata));

%% CALCULATION OF D1
j = p - q + 1;
n = r;
%generating the matrix of q pcs
for i = 1:n
   sum = 0;
   sum1 = 0;
   for k = j:p
      temp = newdata(i,k)*newdata(i,k);
      sum = sum + temp;

   end
      d1_sq(i,1) = sum;
end
```

```
d1=sqrt(d1_sq);

Range=strcat('C',int2str(qq+0*pos),':','Q',int2str(qq+0*pos));
Location=strcat(pwd,'\','d1.xls');  %write dvalue to this location in the Excel
xlswrite(Location,d1',Range);
d1;
%% data sort
C=[d1';B]';

for i=1:(r-1) % didn't use sort command here
for j=1:(r-1)

      k1=C(j,1);
      k2=C(j+1,1);
      k3=C(j,2);
      k4=C(j+1,2);

      if k1>k2

      C(j,1)=k2;
      C(j+1,1)=k1;
      C(j,2)=k4;
      C(j+1,2)=k3;

      end

end
end
disp('d1 value after sorting ')
C
%% cdfplot
BB=C(:,1);
xBB=log10(BB);
hold on;
cB=cdfplot(BB);
hold on;
set(cB,'color','m');
hold on;
x1=min(BB)+((max(BB)-min(BB))/3);
x2=min(BB)+((max(BB)-min(BB))*2/3);
y=[0:0.01:1];
plot(x1,y,'r--')
```

```
hold on;
plot(x2,y,'r--');
hold on
grid off
titl=sprintf('cdf curve of the %s data',data)
title(titl)
xlabel('d1 value')
ylabel('cdf function F(x)')

%%
for k=1:r
    m=k/r;
   text(BB(k,1),m+0.015,int2str(C(k,2)));
hold on;
end
pq=sprintf('p=%d,q=%d',p,q);
lot=sprintf('Lot: %s',data);
text(x2+4*(max(BB)-min(BB))/30,0.3,pq)
text(x2+4*(max(BB)-min(BB))/30,0.38,lot)
xlim=([min(BB) max(BB)]);
Range1=strcat('B',int2str(qq+0*pos),':','B',int2str(qq+0*pos))
Range2=strcat('A',int2str(qq+0*pos),':','A',int2str(qq+0*pos))
```