ESTIMATING THE EXPECTED LATENCY TO FAILURE

DUE TO MANUFACTURING DEFECTS

A Thesis

by

DAVID MICHAEL DORSEY

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2003

Major Subject: Computer Engineering

ESTIMATING THE EXPECTED LATENCY TO FAILURE

DUE TO MANUFACTURING DEFECTS

A Thesis

by

DAVID MICHAEL DORSEY

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

M. Ray Mercer (Chair of Committee)	Hank D. Walker (Member)
Gwan Choi (Member)	Chanan Singh (Head of Department)

December 2003

Major Subject: Computer Engineering

ABSTRACT

Estimating the Expected Latency to Failure Due to Manufacturing Defects.

(December 2003)

David Michael Dorsey, B.S., Texas A&M University

Chair of Advisory Committee: Dr. M. Ray Mercer

Manufacturers of digital circuits test their products to find defective parts so they are not sold to customers. Despite extensive testing, some of their products that are defective pass the testing process. To combat this problem, manufacturers have developed a metric called defective part level. This metric measures the percentage of parts that passed the testing that are actually defective. While this is useful for the manufacturer, the customer would like to know how long it will take for a manufacturing defect to affect circuit operation. In order for a defect to be detected during circuit operation, it must be excited and observed at the same time. This research shows the correlation between defect detection during automatic test pattern generation (ATPG) testing and normal operation for both combinational and sequential circuits. This information is then used to formulate a mathematical model to predict the expected latency to failure due to manufacturing defects.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

TABLE OF FIGURES

INTRODUCTION

Manufacturers of integrated circuits (IC) spend a great deal of time testing their products to ensure that their products are defect free. Manufacturers would ideally test every possible input combination and compare the observed result to the expected result. Practically, this is impossible. A circuit with thirty inputs would have over one billion possible input combinations. Unfortunately, the tests patterns that are applied during testing do not detect every possible defect that could be in the circuit. Consequently, some defective parts will pass the testing process and be considered non-defective. The percentage of parts that contain defects but considered defect-free is called defective part level. Defective part level is an important industry metric because it gives the manufacturer and the consumer an estimate of the probability that the integrated circuit contains a manufacturing defect. Obviously, a lower defective part level is preferable to a high defective part level, but obtaining a defective part level of zero is impractical. The question that remains is how low a defective part level is low enough?

This work is the beginning of the answer to that question. We attempt to answer that question by finding the expected latency to failure due to manufacturing defects (ELF-MD) for a defective IC given that it was tested with a test pattern set that produced a defective part level. If the expected latency is very large – on the order of the life of the product or longer, then additional testing is probably unnecessary. Software errors or normal "wear and tear" will probably lead to errors before the

_____

The journal model is *IEEE Transactions on Automatic Control*.

manufacturing defect. Accordingly, if the expected latency is relatively small, then additional testing is necessary to filter out the defective parts.

We start the answer by determining what information is needed to estimate the ELF-MD. We test different circuits with automatic test pattern generation (ATPG) and then with random patterns. Random patterns are used to simulate the circuit in normal operation. First we determine how often each site in the circuit is observed during ATPG testing and during "normal operation." Then, under each operating condition, we determine how quickly different defects are detected. In essence, we determine how "easy" it is to detect a defect under each operating mode. We then use that information to develop an initial model that will predict ELF-MD.

The next section of this thesis describes the work that has been done in the area of defective part level prediction. Then the concept of scan based testing is discussed. Then the correlation between how "easy" it is to detect a defect in "normal operation" and with ATPG patterns is presented. Following that, a model that predicts ELF-MD is introduced. Finally, this thesis discusses future work to be done in this area and the conclusions that can be drawn from this research.

PREVIOUS WORK

At one time, the amount of logic contained in an integrated circuit was small enough to allow it to be tested with every possible input pattern. Using this test set gave the manufacturer a great deal of confidence that all parts that passed the tests were non-defective. As the industry advanced, this quickly became impractical because the number of possible test patterns grows exponentially with the number of inputs to the circuit. Consequently, the test patterns applied during testing are only a small subset of all possible test patterns. The question is how to determine in a cost-effective way whether the device in question was manufactured correctly [1].

In 1959, R. D. Eldred realized that because defects are physical entities that occur in the circuit's structure, test patterns can be created that attempt to detect the defects occurring at different circuit locations. This is done by targeting "faults", which predict the effect of modeled defects on the logic operation of the circuit. For this purpose he proposed the single stuck-at fault model [2]. In this model, a site in the circuit is either "stuck" at logic one, "stuck" at logic zero, or non-defective. If a site is "stuck" at logic one, that site will be considered a logic one even if the logic value at that site should be zero. The single stuck-at fault model also assumes that any defective circuit will only have one fault present. If all stuck-at faults were considered, there would be 3^N faults, where N is the number of sites in the circuit, because each site could be stuck-at one, stuck-at zero, or non-defective. Since simulating every defect would be impractical, the industry uses the single stuck-at fault model [1].

The number of possible stuck-at faults only grows linearly with circuit size unlike the number of possible input combinations, which grows exponentially. In

addition, a single test pattern that is generated to detect a specific stuck-at fault will often fortuitously detect many other single stuck-at faults as well. Therefore, the number of test patterns required to detect all stuck-at faults in the circuit is generally significantly smaller than the number of possible stuck-at faults.

The single stuck-at fault coverage of a test set has been used as a metric of test set quality. A test set that achieves 99% stuck-at fault coverage is considered better to one that achieves only 95% stuck-at fault coverage. While this is a useful metric, the actual goal of testing is to identify all of the defective parts and reduce the defective part level, not increase the stuck-at fault coverage. Therefore, an accurate defective part level predictor would be a better metric for comparing the effectiveness of test pattern sets. One of the most famous and widely used defective part level models is the Williams Brown model, which was published in 1981 [3].

The Williams Brown model uses fault coverage and the initial yield before test patterns have been applied to predict the final defective part level according to following formula.

$$DL = 1 - Y^{1-FC} \qquad\qquad (1)$$

Here, Y is the manufacturing yield and FC is the fault coverage of the test pattern set applied. This is usually the stuck-at fault coverage. In most defective part level models, the predicted value of the DL is a fraction which is then converted to parts per million by multiplying by 10^6.

However, the single stuck-at fault model, or any single fault model in fact, cannot represent the effects of all possible defects. Thus, stuck-at fault test sets may not detect an adequate number of the untargeted defects [4], [5], [6], [7], [8]. This is

true even though the Williams Brown model predicts that at 100% fault coverage the defective part level will be zero. In fact, as fault coverage approaches 100%, the test patterns targeted at the remaining stuck-at faults are biased in favor of detecting those faults at the expense of the remaining defects [9], [10], [11]. Furthermore, different test pattern sets with identical fault coverages may have very different defect coverages and thus produce very different defective part levels [12], [13], [14]. In fact, as fault coverage approaches 100%, the standard deviation of the defect coverage of the test sets increases—making fault coverage an inaccurate metric for predicting defect coverage and defect level [14].

However, while stuck-at faults do not capture the behavior of all possible defects, there is a common requirement for detecting any defect, regardless of type: the site where the defect occurs must be observed at an output. In other words, the incorrect value at the site where the defect occurs must be propagated through the circuit logic to a primary output. The excitation requirements vary from one defect type to another. Excitation refers to the need to create a difference in the expected logic value and the observed logic value at the defect site. Both excitation and observation must occur simultaneously for the defect to be detected.

Consider two circuits, each of which are composed of a single two-input OR gate as shown in Figure 1. However the second OR gate has a defect corresponding to a stuck-at 0 fault at site A'. If both circuits are assigned values as shown in Figure 1, the defect is not excited because there is no difference in the value caused by the fault and the expected value at that point. Also, the defect site is not observed because the logic one at sites B and B' cause the output of the OR gate to be a logic one regardless

of the logic value at sites A and A'. So this input pattern satisifies none of the requirements to detect a defect.
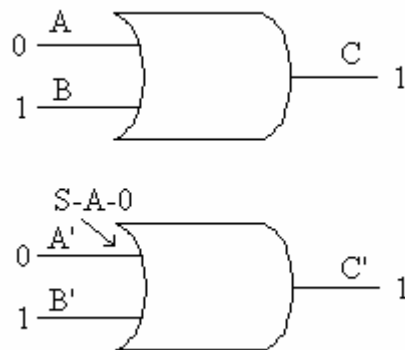
Figure 1. Test Pattern That Does Not Detect Defect

If the input pattern is as shown in Figure 2, the defect is not detected but it is excited. The logic value at site A' is remains zero despite inputting a logic one because of the stuck-at fault. However, the defect is not observed because the output is a logic one regardless of the values at A and A' because of the logic ones on B and B'.
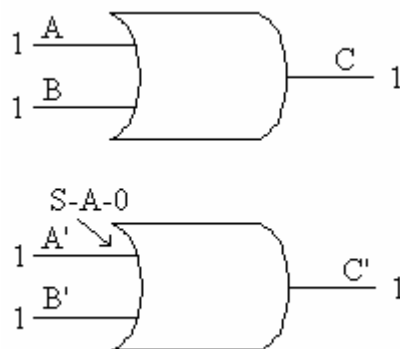
Figure 2. Test Pattern That Excites Defect But Does Not Observe the Defect

However, if the input pattern is the one shown in Figure 3, the defect is observed and excited and therefore detected. This is demonstrated by the difference in the values at C and C'. The logic values at sites B and B' do not control the output of their respective OR gates allowing the values at sites A and A' to influence the output. Since the defect at A' is also excited, the incorrect logic value is propagated to the output.
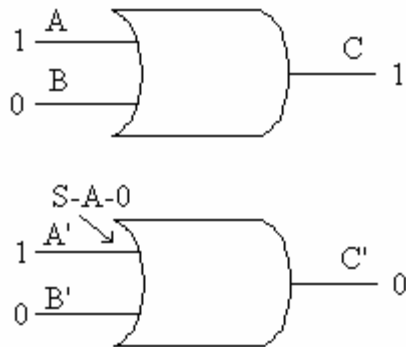
Figure 3.   Test Pattern That Detects Defect

Just as stuck-at faults may be fortuitously detected by test patterns targeting other stuck-at faults, a defect not well-modeled by a stuck-at fault may be fortuitously detected by a test pattern that targets that fault if the site where the defect occurs is observed and the defect happens to be excited. In fact, it has been found that as a site is observed more times, the probability of an undetected defect never having been simultaneously excited decreases significantly.

This analysis of the commonality and differences among excitation and observation requirements of defects lead to the Deterministic Observation, Random Excitation, and MPG-D Defective Part Level Estimation (DO-RE-ME) test pattern

generation method [15]. When the DO-RE-ME method is used, emphasis is placed on observing every circuit site, especially those that are difficult to observe, as many times as possible while randomly exciting whatever defects may occur at those sites. In addition, the MPG-D defective part level model is used to predict the defective part level of the resulting test set and to choose among possible subsets if too many vectors are initially generated to fit in the tester memory. Unlike defective part level models that predict the defective part level based upon simple fault coverage, the MPG-D model predicts the defective part level based upon the number of observations of different circuit sites or faults and has been shown to be more accurate, especially at very high fault coverages [16].

The observation data required for the MPG-D defective part level model can be obtained from a fault dictionary. Initially, all circuit sites are assigned a contribution to the overall defective part level. This contribution is usually equal for each circuit site. Thus, the defect level contribution of site i before any test patterns have been applied is shown in equation 2.

$$DL_i(0) = \frac{1 - Yield}{\# \, of \, sites} \tag{2}$$

The defect level contribution of each site then changes as patterns are applied based upon observation counts of those sites. The probability of exciting an undetected defect at a site given that that site is observed has been studied [17] and shown to be a decaying exponential function of the number of times that site has been observed previously and a time constant τ, as described in equation 3.

$$P_{excite|obs_i} = e^{-\frac{\#obs_i}{\tau}} \tag{3}$$

This makes intuitive sense. Consider Figure 4. Here the boxes represent all test patterns (input combinations) that will observe a given site i. Each oval within the boxes represents the test patterns that will detect a corresponding undetected defect. In other words, those test patterns will excite that defect while it is observed. The left rectangle represents the test spaces for undetected defects before a test pattern has been applied. A large portion of the box is covered, indicating that the simultaneous excitation of at least one undetected defect given that this site is observed is highly likely the first time it is observed.

Test pattern applied
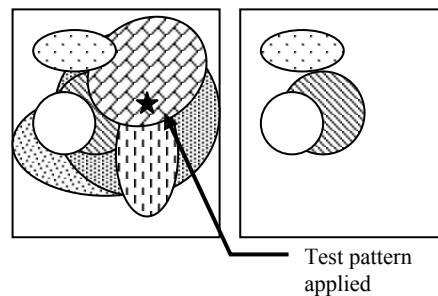
Figure 4. Test Spaces of Undetected Defects Given Before and After the Test Pattern Is Applied

Now assume that the first test pattern to observe this site is located at the point indicated by the star. If this test pattern is chosen, then several of the defects will be detected and therefore do not appear in the box on the right. The probability of exciting at least one undetected defect given that the site is observed is now considerably lower.

This information is used to calculate the change in defective part level contribution of each site as a result of whether or not it has been observed by a pattern according to the following formula:

$$\Delta site_i = \begin{cases} DL_i(n-1)*(A*P_{excite|obs_i}), \\ \quad \text{if site } i \text{ was observed by pattern } n \\ \\ 0, \text{otherwise} \end{cases} \quad (4)$$

Here, the constant A represents the fraction of defective part level contribution that will be removed from the site given that at least one undetected defect is excited and observed.

Other equations are used to calculate additional changes in defective part level contribution due to the sharing of defects among circuit sites, giving a resulting value of $\Delta share_i$. Then a new value for each site's defective part level contribution after pattern n has been applied is calculated according to the following equation:

$$DL_i(n) = DL_i(n-1) - \Delta site_i - \Delta share_i \quad (5)$$

The overall defective part level is calculated by summing the defective part level contributions of every site.

$$Total\_DL(n) = \sum_{i=1}^{\# \, of \, sites} DL_i(n) \quad (6)$$

However, while the defective part level obtained by a given test pattern set is very valuable information for both the integrated circuit manufacturer and the customer, another valuable metric to consider would be the expected latency to failure due to manufacturing defects. This is important because it gives an estimate of how soon the defect will affect the operation of a circuit. It allows for a quantitative analysis of

whether defective part levels obtained with current test sets are low enough for a given application. In addition, it allows the probability that the first error that occurs during circuit operation will be due to a manufacturing defect to be compared to the probability that the first error will be due to either a software error or early-life failure.

SCAN BASED TESTING

Scan based testing is used to test sequential circuits. Sequential circuits are circuits that contain memory. Their outputs depend on the current input and previous inputs. Therefore, a defect that is excited during one clock cycle may be observed in the current clock cycle or future clock cycles. Figure 5 shows a schematic diagram of a sequential circuit.
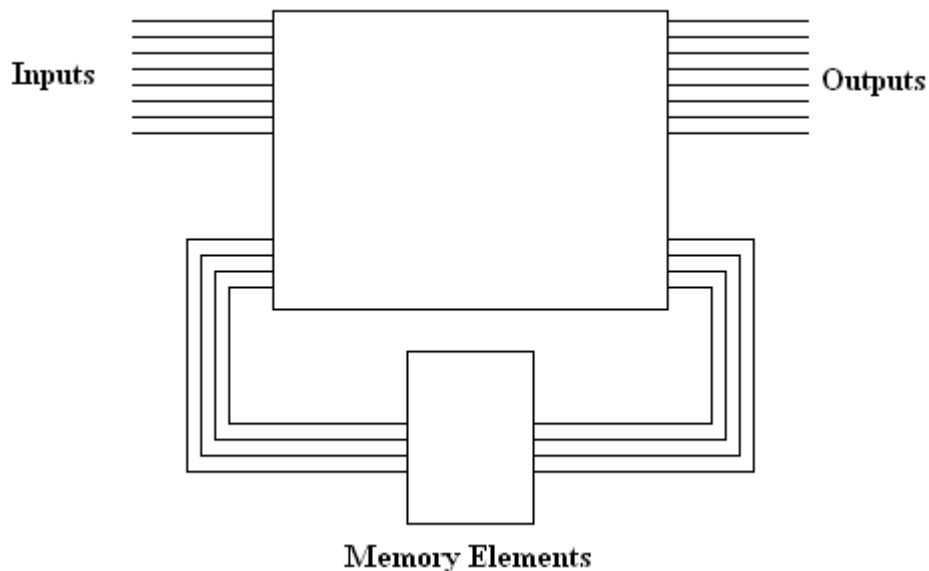
Figure 5. Huffman Model of a Sequential Circuit

Testing sequential circuits is difficult because controlling the internal state of the memory elements is a difficult task. Controlling the outputs of the memory elements is important because of the need to excite defects that might exist near the memory elements.

To overcome this difficulty, many companies developed scan based testing. Transforming a sequential circuit into a combinational circuit is important because combinational circuits are easier to test than sequential circuits. Also, the testing of combinational circuits is a well understood problem and many techniques have been developed to test combinational circuits. In scan based testing, the D flip-flops in the circuit can operate in two modes. In the normal mode, the D flip-flop acts like it would normally – at the specified clock edge, the value at the input is transferred to the output. In scan mode, instead of the input being from the circuit, the input is from an input into the circuit. The D flip-flops are hooked in series connected to the next D flip-flop's scan input. In essence, the D flip-flops in the scan chain behave like a shift register [1]. Figure 6 shows a scan chain that has three D flip flops.
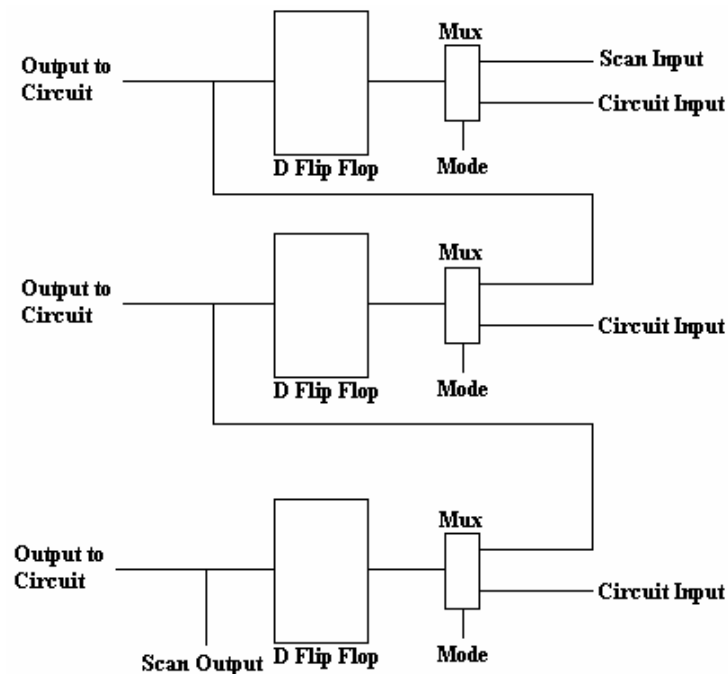
Figure 6. A Three D Flip-Flop Scan Chain

The values for the D flip-flops are shifted into the scan chain serially. Once all the values are inputted, the circuit is put into normal operation. Then the circuit is clocked to get the result of the circuit. Then the circuit is put back into scan mode and the results are outputted serially through the scan output. The output results can then be compared the expected values at that point [1].

Scan based testing is an important industry tool to test sequential circuits. It increases the controllability and observability of points in the circuit. This research uses scan based testing to test sequential circuits during ATPG testing.

ELF-MD: WHAT INFORMATION DO WE NEED?

Before we can extend MPG-D to determine ELF-MD, we need to know what we need to know in order to determine ELF-MD. To that end, we ran tests on four of the ISCAS85 [18] benchmark circuits and two of the ISCAS89 [19] benchmark circuits. Specifically, we tested circuits C432, C499, C880, and C2670 from the ISCAS85 benchmark circuits and circuits S27 and S344 from the ISCAS89 benchmark circuits. Using Verilog, we inserted high impedance faults, stuck-at 1 faults, stuck-at 0 faults into each input wire and each internal wire of the circuit such that only a single surrogate was inserted for any instantiation of the circuit. We also identified the non-feedback AND bridges in circuit C432 and modeled a subset of those surrogates. We applied ATPG patterns to the circuits to determine when the surrogates would be detected during the testing process. To simulate "normal operation", we applied random patterns to the circuit. Since random patterns are not predictable, we tested the circuit one thousands times and took the average to find the cycle number the circuit failed on. Taking the inverse of this number gave us the probability of detecting the surrogate during "normal operation". We then compared this probability to the probability of find the surrogate during ATPG testing.

We discovered that the correlation between detecting a defect during ATPG testing and during "normal operation" depends greatly on how often you observe the defect site of the circuit. Another important factor we discovered is whether the circuit is combinational or sequential.

The first piece of information we wanted was how often each site is observable for each circuit. We used high impedance faults to determine the observability of the

sites in each circuit. In our simulations, high impedance points are always excited because they simulate a disconnected wire in the circuit. The value at that point is lost and an unknown value propagates towards the output instead. If the unknown value affects a primary output, a don't care value will appear on that output during simulation. It is possible to detect don't care values during simulation, so whenever the site is observed the surrogate will be detected.

Observability of Sites for Circuit C432
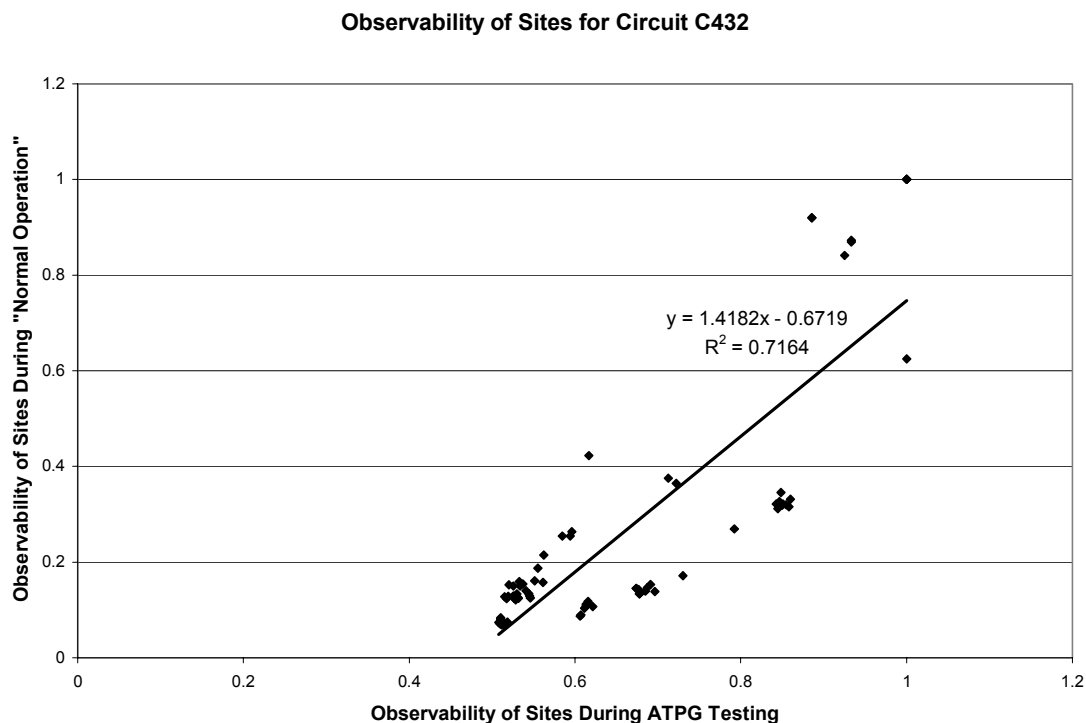
Figure 7. Observability of Sites during ATPG Testing and "Normal Operation" for C432

Figure 7 shows that there is a correlation between observing sites during ATPG testing and during normal operation for circuit C432. It also shows that, in general, it less likely to observe a site during "normal operation" than during ATPG testing. This

makes intuitive sense because the DO-RE-ME method attempts to observe each site as many times as possible while "normal operation" does not.

Figure 8 shows that there is also a correlation between observing sites during ATPG testing and during normal operation for circuit C499. The correlation in this circuit is much stronger than in circuit C432. Observing a point during "normal operation" is just as likely as during ATPG testing.

Observability of Sites for Circuit C499

Figure 8. Observability of Sites during ATPG Testing and "Normal Operation" for C499

Figure 9 shows that there is also a correlation between observing sites during ATPG testing and during normal operation for circuit C880. The correlation in this circuit is stronger than in circuit C432, but not as strong as it is in circuit C499.

Observing a point during "normal operation" is less likely than observing a point during ATPG testing.

Observability of Sites for Circuit C880

The chart shows data points with the regression equation:

$$y = 1.0313x - 0.0894$$
$$R^2 = 0.7818$$

The x-axis is labeled "Observability of Sites During ATPG Testing" and the y-axis is labeled "Observability of Sites During 'Normal Operation'".
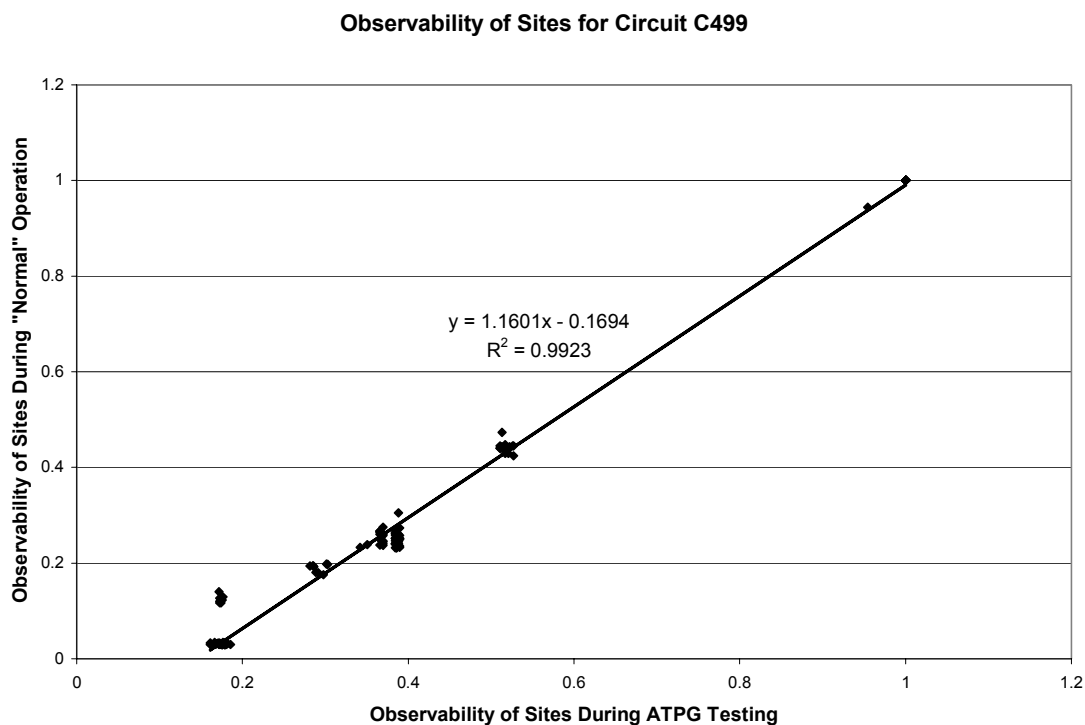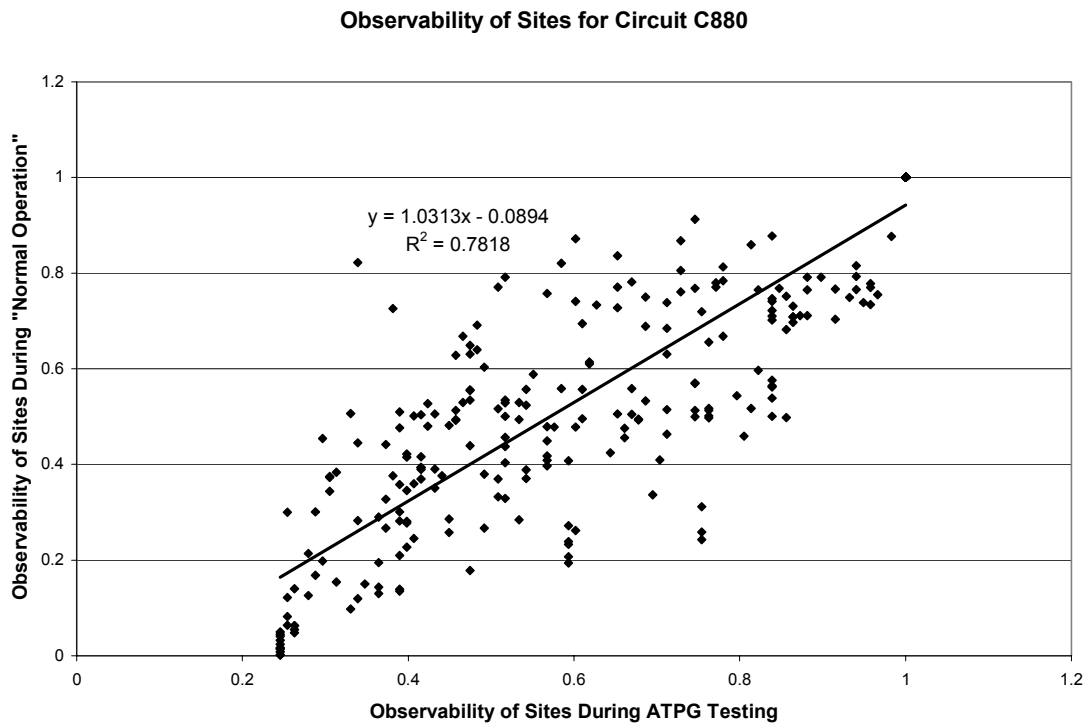
Figure 9. Observability of Sites during ATPG Testing and "Normal Operation" for C880

Figure 10 shows that there is also a correlation between observing sites during ATPG testing and during normal operation for circuit C2670. The correlation in this circuit is weaker than the correlation in the other circuits. However, unlike the other circuits, observing a point during "normal operation" is far less likely than observing a point during ATPG testing. In fact, some points that are observable over 68 percent of the time during ATPG testing are observable less than 0.001 percent of the times during "normal operation".

Observability of Sites for Circuit C2670
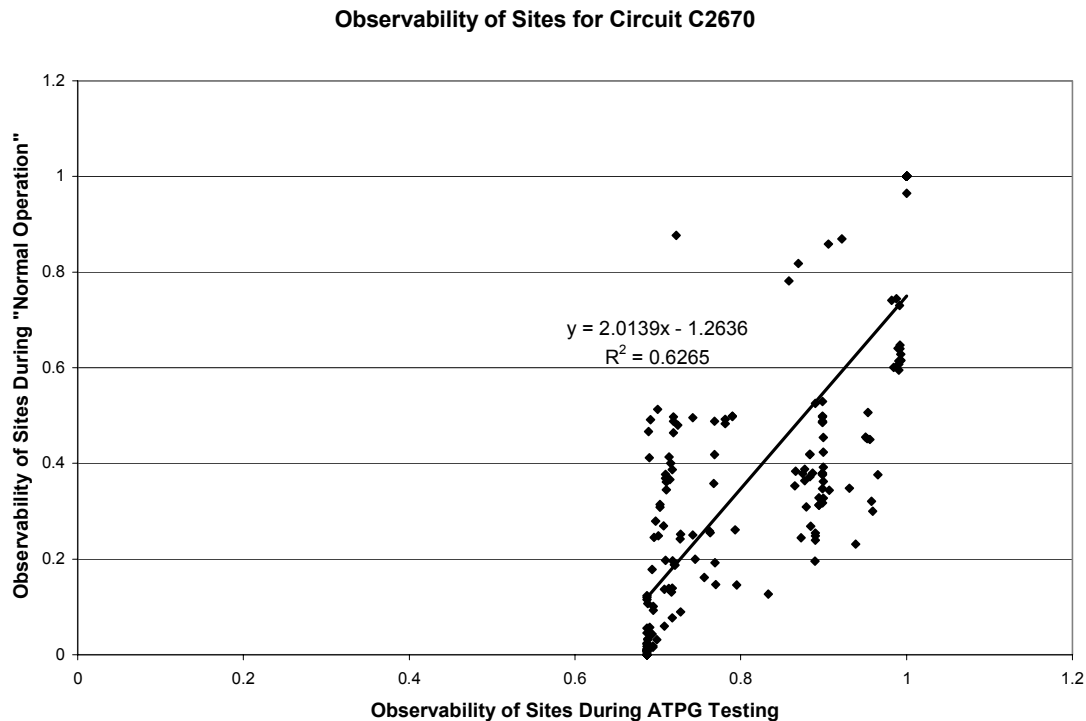
Figure 10. Observability of Sites during ATPG Testing and "Normal Operation" for C2670

Based solely on combinational circuits, there would seem to be a decent correlation between observing points during ATPG testing and during "normal operation". However, when we tested sequential circuits we found that this was not true at all.

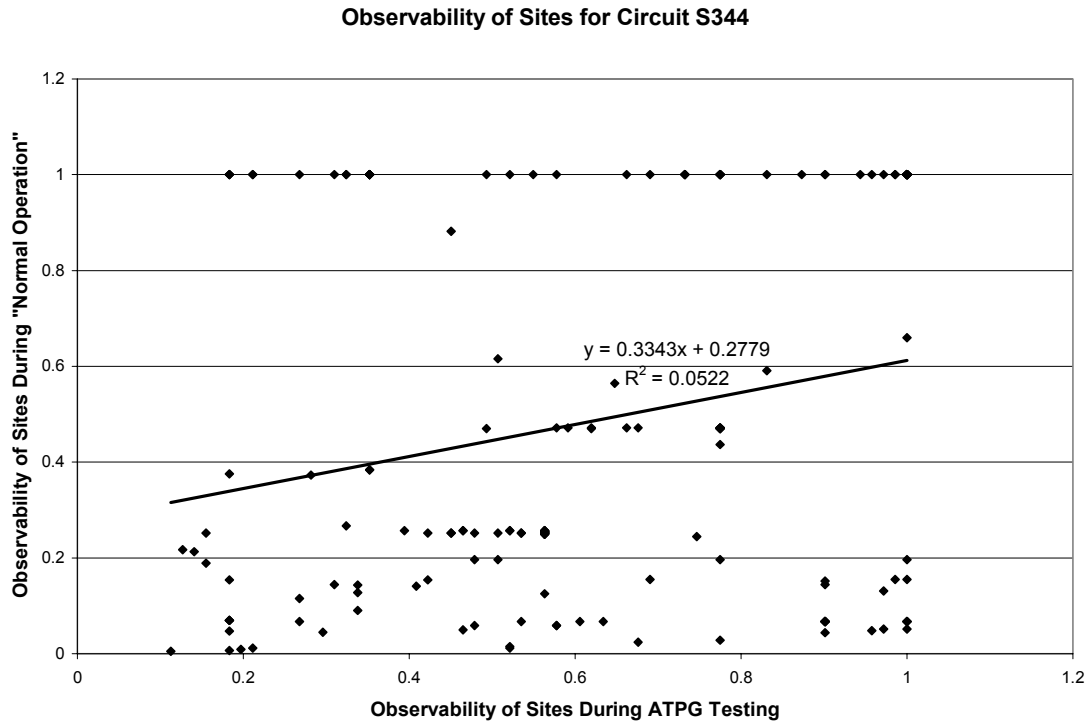Figure 11 shows the observability of every site for the sequential circuit S27. As you can see, there is no apparent correlation between observing a site during ATPG testing and during "normal operation". In fact, two of the points that were observed 100 percent of the time during ATPG testing were only observed 15 percent of the time during "normal operation".

Observability of Sites for Circuit S27
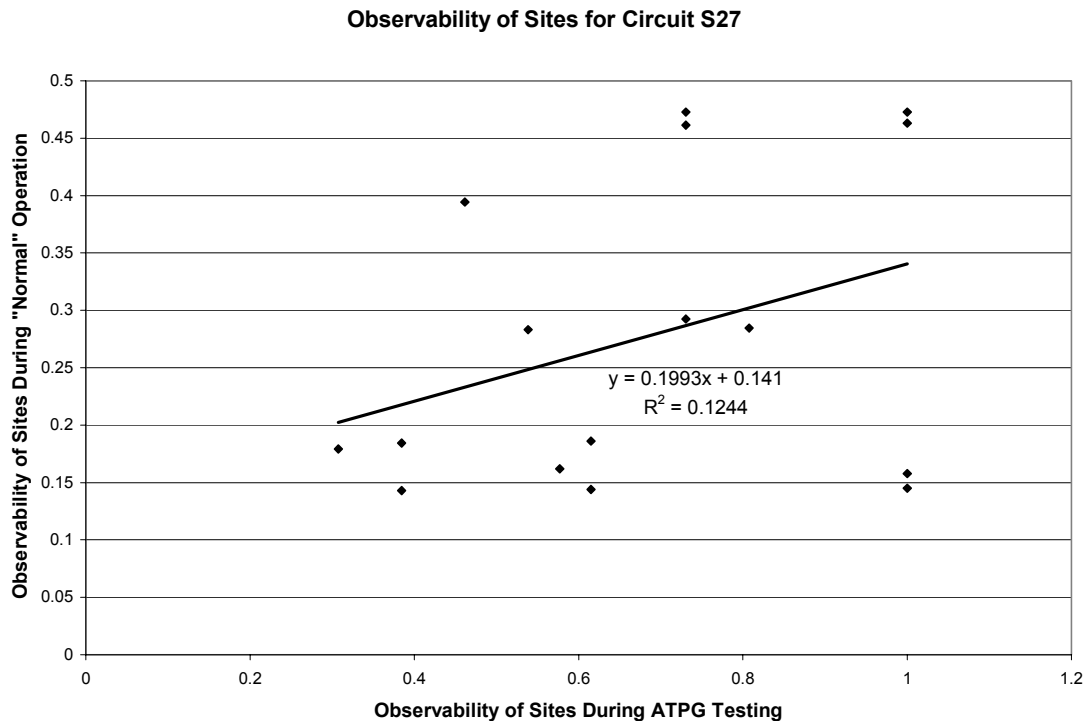
Figure 11. Observability of Sites during ATPG Testing and "Normal Operation" for S27

These two points were inputs to two of the D flip-flops in the circuit. They were observed every time during ATPG testing because they were tested using scan based testing techniques describe in the previous section. In scan based testing, the inputs to D flip-flops become pseudo-outputs of the circuit. In sequential operation, these points lead back into the circuit, giving the logic value there a chance to be blocked.

We tested another sequential circuit, S344, to verify that this result was not limited to circuit S27. Figure 12 shows the observability of every point in S344. As expected, there is no apparent correlation between observing a site during ATPG testing and during "normal operation". In fact most, points were observed less than 40 percent of the time during "normal operation".

Observability of Sites for Circuit S344

Figure 12. Observability of Sites during ATPG Testing and "Normal Operation" for S344

Based on these results, we can conclude that there is a correlation between observing a site during ATPG testing and during "normal operation" for combinational circuits, but not for sequential circuits. Any predictor of ELF-MD should contain information about "normal" circuit operation.

Another useful piece of information to know is how likely it is to detect a defect during ATPG testing compared to "normal operation." Determining that the likelihood of detection during "normal operation" is lower than ATPG testing would be ideal because that would extend the value of ELF-MD. Unfortunately, this may not always be the case because it is possible for the detection requirements for many of the particular faults targeted during ATPG to conflict with the detection requirements for a specific defect.

Figure 13 shows stuck-at 0 surrogate detection in each operating mode for circuit C432. Figure 14 shows stuck-at 1 surrogate detection in each operating mode for circuit C432. Neither simulation shows a strong correlation between defect detection during ATPG testing and "normal operation." However, both figures show that is more likely to detect a surrogate during ATPG testing than in "normal operation."

Stuck-at 0 Defect Detection for Circuit C432
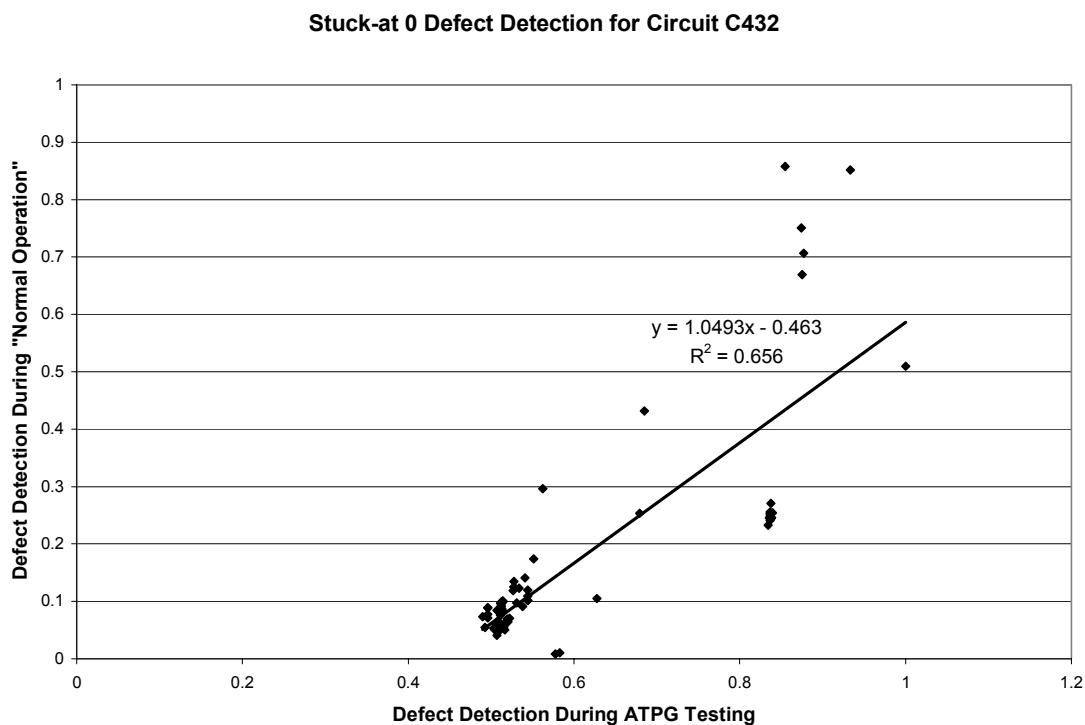
Figure 13. Stuck-at 0 Defect Detection during ATPG Testing and "Normal Operation" for C432

Stuck-at 1 Defect Detection for Circuit C432
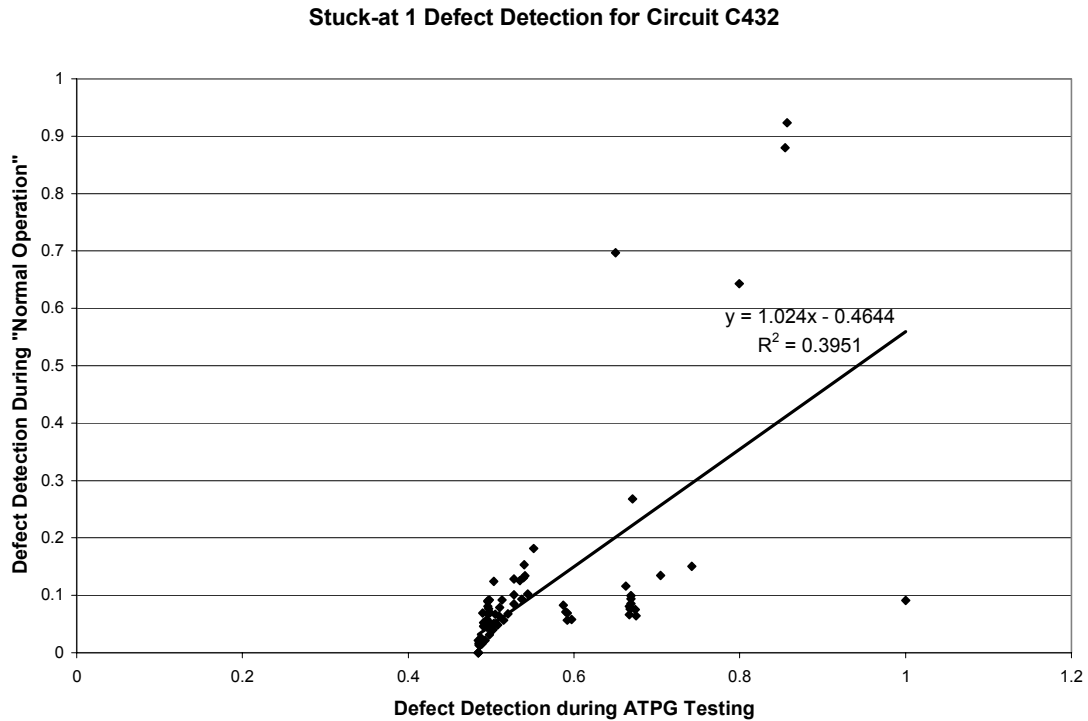
Figure 14. Stuck-at 1 Defect Detection during ATPG Testing and "Normal Operation" for C432

Figure 15 shows stuck-at 0 surrogate detection during ATPG testing and "normal operation" for circuit C499. Figure 16 shows stuck-at 1 surrogate detection in each operating mode for circuit C499. When C499 was modeled with stuck-at 0 faults, there was some correlation between defect detection during ATPG testing and "normal operation." When C499 was modeled with stuck-at 1 faults, there was a strong correlation between the operating modes. In addition, both figures show that, in general, it is more likely to detect a surrogate during ATPG testing than in "normal operation."
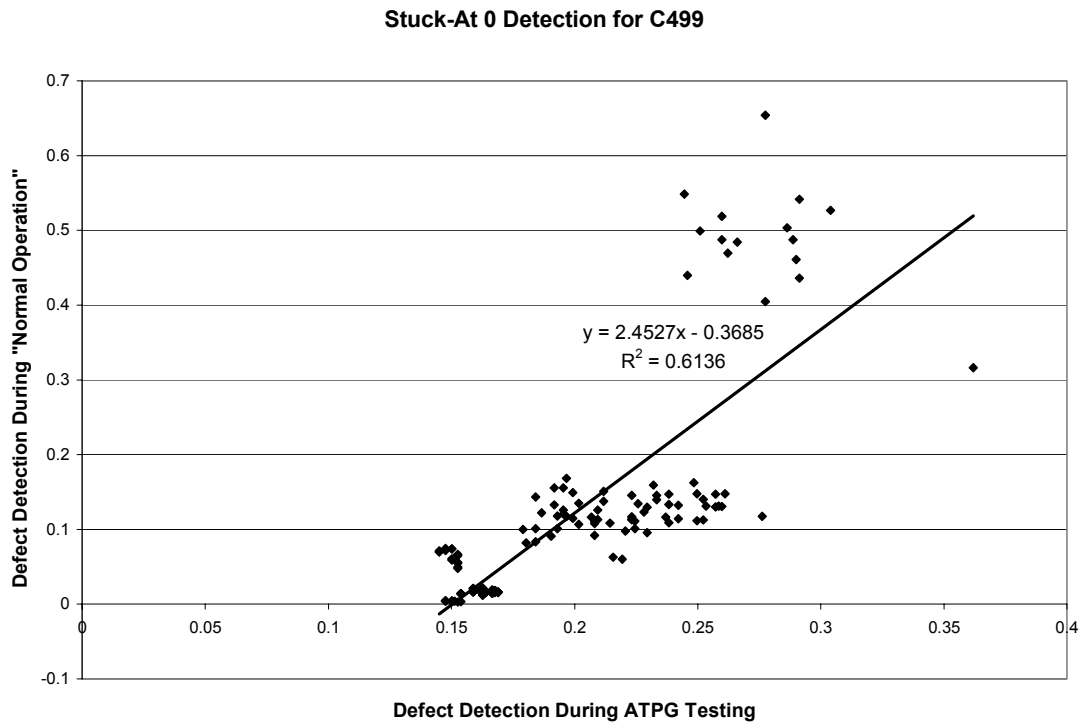
Stuck-At 0 Detection for C499

Figure 15. Stuck-at 0 Defect Detection during ATPG Testing and "Normal Operation" for C499

Stuck-at 1 Detection for C499
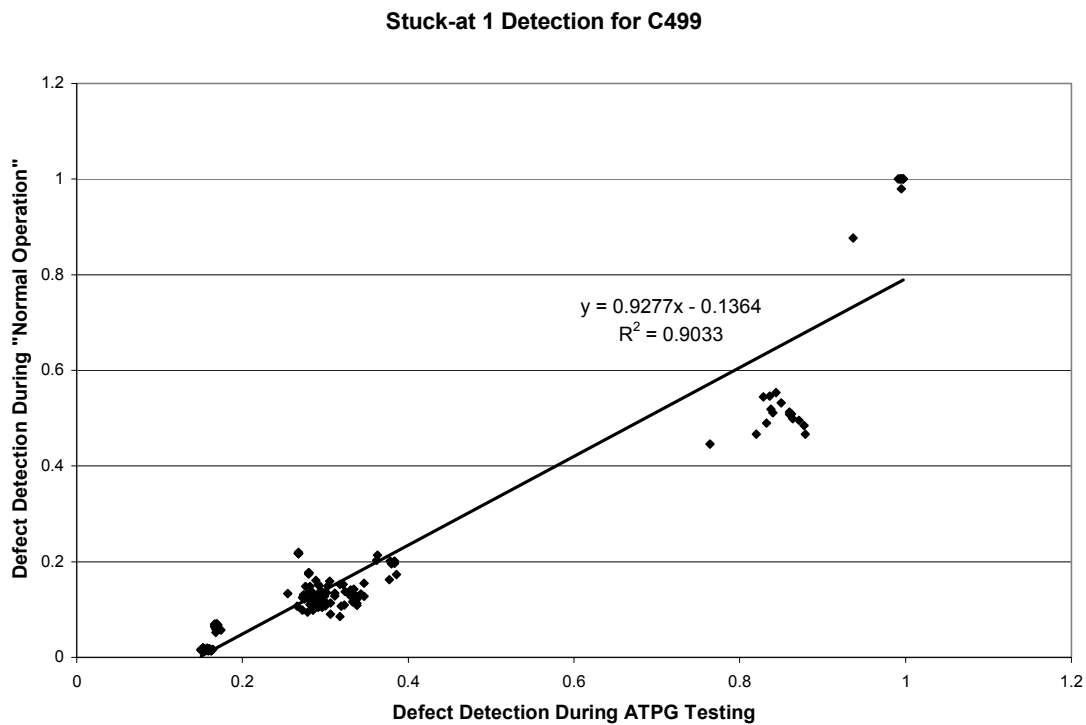
Figure 16. Stuck-at 1 Defect Detection during ATPG Testing and "Normal Operation" for C499

The final surrogate model we simulated for combinational circuits was non-feedback AND bridges in circuit C432. We chose our bridging faults to be non-feedback because feedback bridges can lead to unstable circuit operation. Figure 17 shows the results from our simulations of a subset of all possible AND bridges. While many of the AND bridges were detected more often during ATPG testing, there are a significant number of AND bridges that were detected more often during "normal operation." The variations in the probability occur at low probability of detection during ATPG testing. Also, the correlation between the two modes is weak at best. This may mean that the correlation between the two operating modes diminishes with more complex surrogate models.

Non-Feedback AND Bridge Defect Detection for Circuit C432

$y = 0.4303x + 0.0715$
$R^2 = 0.4445$
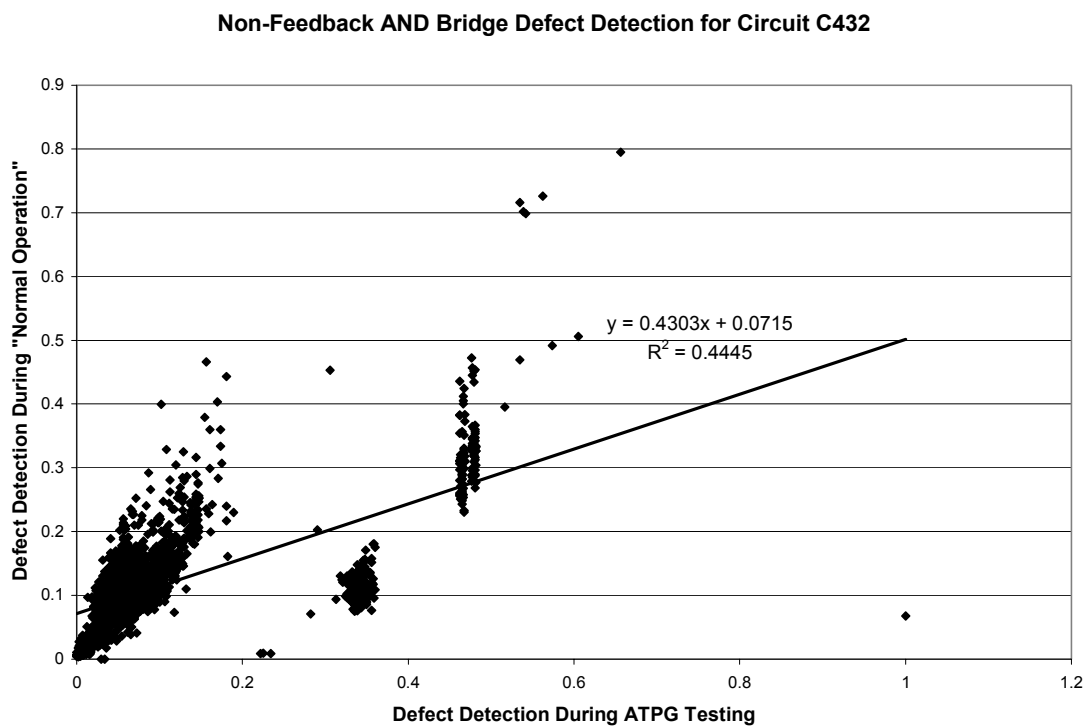
Figure 17. AND Bridge Detection during ATPG Testing and "Normal Operation" for C432

We also tested sequential circuits to see how often we detect surrogates in each operating mode. There is no correlation between detecting a defect in the sequential circuits S27 during ATPG testing and "normal operation." Also, the percentage of time that the surrogate is detected in ATPG testing does not give us any insight into how many often the surrogate is detected in "normal operation." This can be attributed to what we discovered earlier – that the observability of sites during ATPG testing and during "normal operation" have no relationship.

Stuck-at 0 Defect Detection for Circuit S27

$y = 0.2059x + 0.1481$
$R^2 = 0.0705$
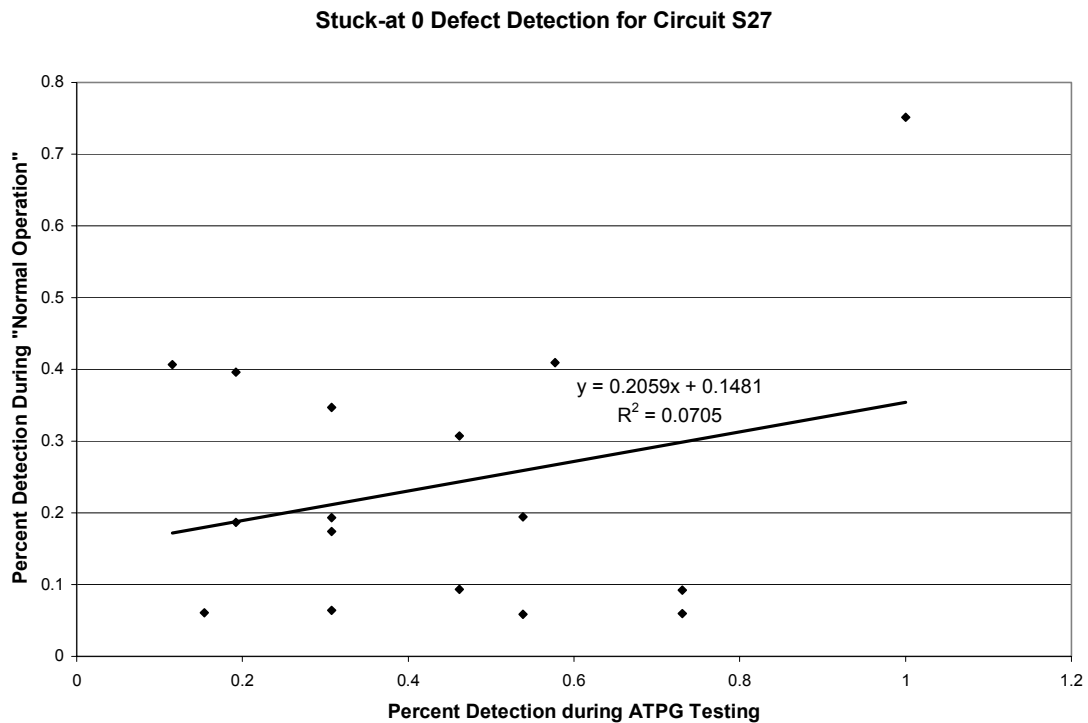
Figure 18. Stuck-at 0 Defect Detection during ATPG Testing and "Normal Operation" for S27

Stuck-at 1 Detection for Circuit s27
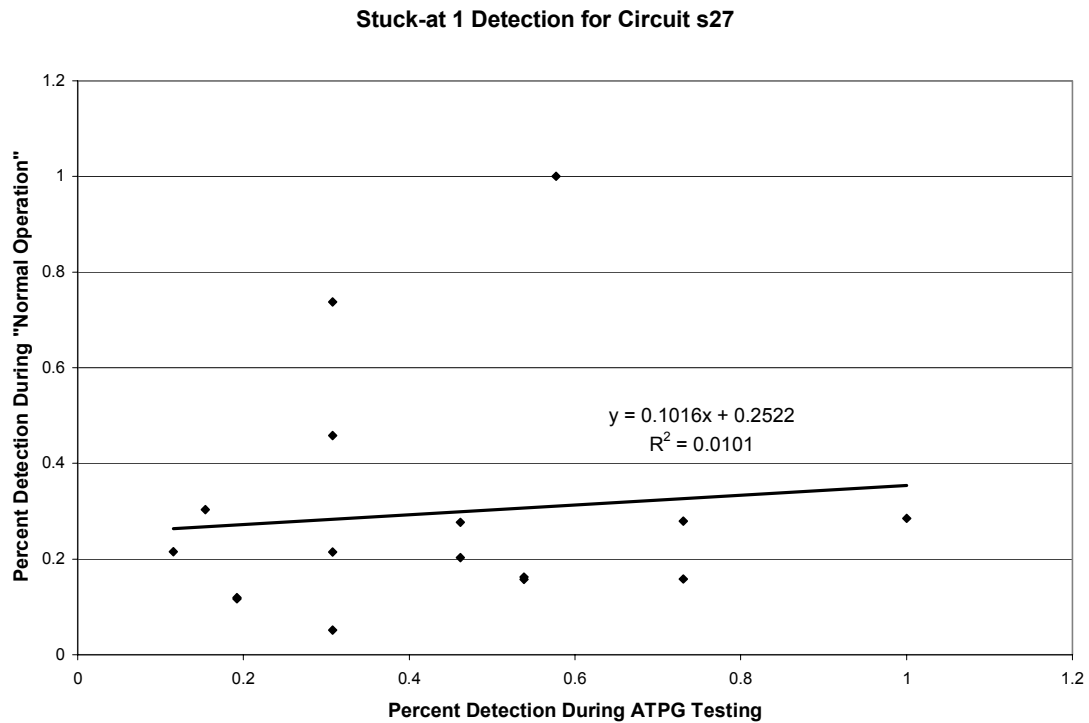
Figure 19. Stuck-at 1 Defect Detection during ATPG Testing and "Normal Operation" for S27

Figures 18 and 19 are evidence that there is no correlation for defect detection between the two operating modes. Figures 20 and 21 show the same lack of correlation for circuit S344. The probability of detecting defects during normal operation bears little resemblance to the probability of detecting defects during ATPG testing.

Stuck-at 0 Defect Detection for Circuit S344

$$y = 0.5732x + 0.0597$$
$$R^2 = 0.0932$$

Figure 20. Stuck-at 0 Defect Detection during ATPG Testing and "Normal Operation" for S344

Stuck-at 1 Defect Detection for Circuit S344

$$y = 0.0869x + 0.2752$$
$$R^2 = 0.0018$$

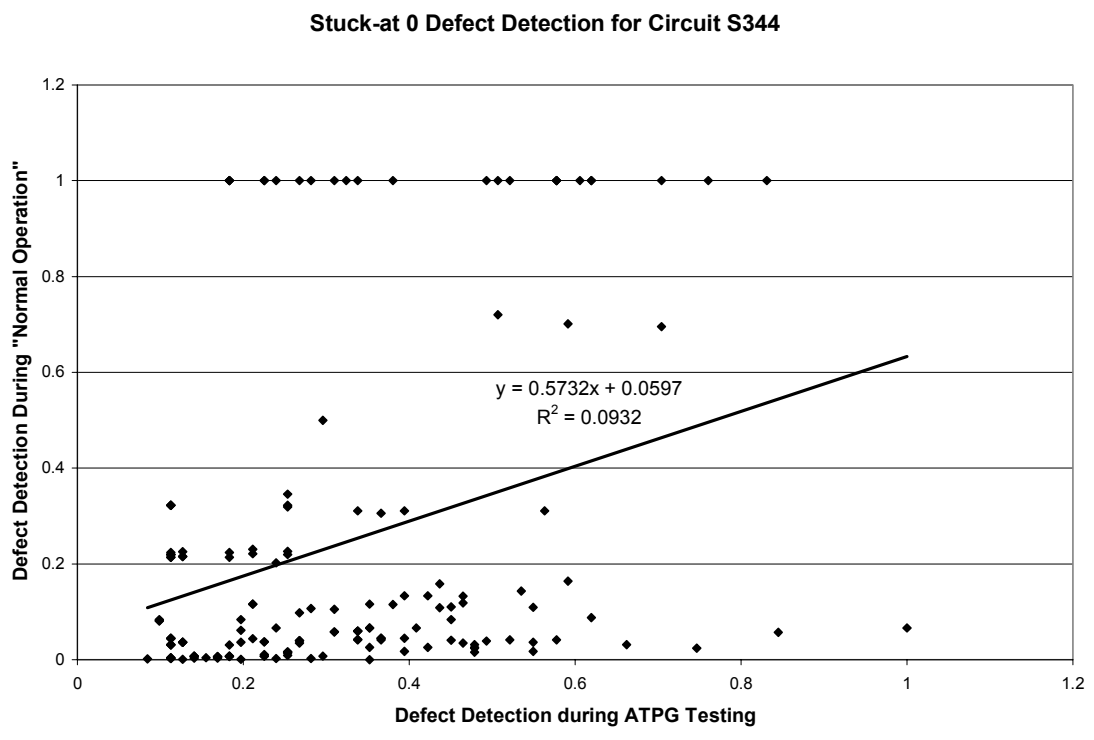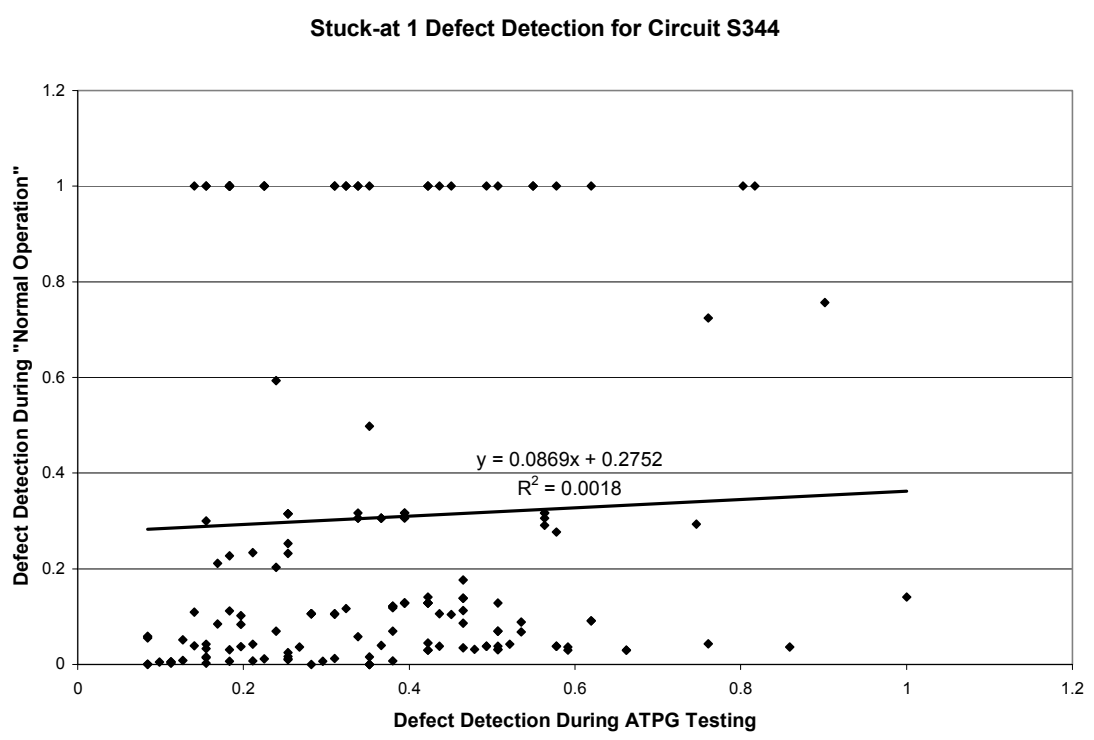**Figure 21. Stuck-at 1 Defect Detection during ATPG Testing and "Normal Operation" for S344**

However, both Figure 20 and Figure 21 show sites that were detected all the time during "normal operation" but not during ATPG testing. Obviously, these sites do not fit the pattern of detecting a defect more often during ATPG testing than "normal operation." After investigating these points in more detail, the reason for there existence is due to their proximity to the outputs of D flip-flops. The circuit is initialized by a reset signal that sets the output of the D flip-flops to 0. This causes some of the sites close to the D flip flop not to be affected by the primary inputs of the circuits for the first clock. However, in ATGP testing, the outputs of the D flip-flops are primary inputs and thus can be varied much easier. For example, if there is a stuck-at 1 fault at the output of a D flip-flop, that defect will always be excited when the circuit leaves reset. But in ATPG testing, since that site is now a primary input, there is a greater chance that the input will also be a 1, which would not excite the defect.

From this data, we can conclude that a simple mapping between defective part level using MPG-D and ELF-MD is not adequate. The MPG-D model emphasizes observing every site as many times as possible. However, we have shown that for sequential circuits the probability of observing the defect sites during normal operation appears to no relation to the probability of detecting defects during ATPG testing. This indicates that an accurate estimator of ELF-MD will need to include information on the probability of observation of different circuit sites during normal operation in addition to the likelihood that each of those sites are likely to still contain defects. Fortunately, data collected while calculating the defective part level using the MPG-D model introduced earlier may prove useful in predicting ELF-MD.

EXTENDING MPG-D TO PREDICT ELF-MD

The requirements for detecting a defect during testing are identical to the requirements that must be met for a defect to cause incorrect behavior during normal operation. Specifically, the defect must be both excited and observed for this to occur.

Recall that the probability of exciting an undetected defect given that a site is observed is modeled as a decaying exponential in MPG-D. Therefore, after test patterns have been applied, many of the circuits containing "easy-to-detect" defects have already been identified and removed from consideration. The remaining defective parts are most likely to contain those defects that are harder to excite and consequently detect. It is how quickly these defects cause observed errors during normal operation that will determine our value of ELF-MD.

Ideally, when calculating ELF-MD, we would assume that only one defect occurs on any given defective chip and would consider the weighted average of the test set sizes for each of the remaining potential defects (as depicted in Figure 1) based upon probability of occurrence of that defect while calculating the probability of exciting a defect while observing the site. However, detailed information of the remaining defect types and their corresponding test spaces is unknown. Accordingly, we propose to use the probability of exciting an undetected defect given that the site is observed calculated according to the MPG-D formula:

$$P_{excite|obs_i} = e^{-\frac{\#obs_i}{\tau}} \tag{7}$$

where $\#obs_i$ is equal to the number of times that site was observed during ATPG, as an upper bound on the probability of exciting an undetected defect at a site given that that

site is observed during normal operation. This is possible because most defects were detected more often during ATPG testing. Since ATPG testing has a better chance of detecting most defects, the probability of exciting a defect given to us by the MPG-D formula can serve as an upper bound. We will use this value in our ELF-MD calculations.

The other requirement that must be satisfied is observation of the site where the defect occurs. As suggested by the data in the previous section, the observation probability used should be the observation probability of each site under the conditions of normal operation. In the absence of application data, probability of observation data for random vectors may be used. The probability that the value of a circuit site in a single clock cycle (when the defect was excited) would affect the outputs in either that clock cycle or in a subsequent clock cycle would need to be determined. If it generally takes many clock cycles for that value to affect the output, then the average number of clock cycles needed should also be collected so that this can be factored into ELF-MD calculations. In either case, some sort of simulation will likely need to be done to collect this data.

Once the probability of observation under normal operating conditions and the probability of excitation of an undetected defect given that the site is observed and given that it was observed for a certain number of times in testing have been obtained for every site, we multiply the two values together. This will give us the probability of detection of an undetected defect at each circuit site given that that site is where the defect occurs. This gives us:

$$P_i = \left(P_{obs_i|\text{normal operation}}\right)\left(P_{exc|obs_i}\right) \tag{8}$$

Therefore, the probability of detecting a defect that occurs at site i for the first time with the first pattern is P_i. Similarly, if independence between patterns is assumed, the probability of detecting a defect that occurs at site i for the first time with the second pattern is:

$$P_i(1-P_i) \tag{9}$$

If we extend this, then the probability of detecting a defect that occurs at site i for the first time with the nth pattern is:

$$P_i(1-P_i)^{n-1} \tag{10}$$

We can use this to find the average number of patterns that will be applied (number of clock cycles) before the defect at site i causes an error in normal circuit operation

$$\sum_{n=1}^{\infty} nP_i(1-P_i)^{n-1} = \frac{1}{P_i} \tag{11}$$

Thus, as expected, the average number of clock cycles that we can expect to pass before the defect at site i is detected is inversely proportional to the probability of that defect being detected. If many additional cycles are expected to be needed before an error will appear at an output, these can be added to our expected value at this point.

However, each site is not equally likely to be the site where the defect occurs. Sites that were observed many times during testing are much less likely to contain the undetected defect than sites that were observed few times, if at all. Thus, when we find the average patterns to detection for the entire circuit, we will need to take a weighted average where the weights are based upon each site's likelihood of containing the defect. For this, we can use the DL contribution of every site calculated by MPG-D.

$$DL_i(n) = DL_i(n-1) - \Delta site_i - \Delta share_i \tag{12}$$

where n is equal to the number of test patterns applied during ATPG testing.

We can calculate the expected number of clock cycles before failure for a circuit given that the circuit is defective and has been tested with a manufacturing test pattern set of given characteristics as:

$$\sum_{i=1}^{number\ of\ sites} \frac{DL_i}{DL} \frac{1}{P_i} \tag{13}$$

Obviously, we can convert this to time and thus ELF-MD using the clock speed. We may also find that we want to take sites with incredibly low DL contributions and remove them from consideration if we are fairly confident that no defects could reasonably occur there. This could reduce the simulation time required for determining the probability of observation values during normal operation.

FUTURE WORK

To better understand the correlation between the defective part level and ELF-MD, more simulations should be conducted to collect data on larger circuits. OR-bridge surrogates should be modeled for simulation as well. Even more complicated surrogates in sequential circuits, such as delays and coupling effects, should be considered for future investigation. These simulations will be useful because the AND-bridge surrogates had a weak correlation between the circuit in "normal operation" and being tested with ATPG patterns. Further investigation will reveal if this was only for that circuit model or if the correlation does diminish with more complex surrogate models. Also, these simulations can be done much faster than they were done throughout most of this work. Recently, techniques have been developed to test many defect sites in parallel. This drastically reduces the amount of time needed to simulate the entire circuit.

In addition to the extra simulations, extensive testing of the model we developed needs to be done. This work does not include an analysis of the effectiveness of the developed model. Also, data should be collected to quantify the uncertainty in the ELF-MD predictions. Intuitively, as the probability of detecting a defect decreases, precisely when that defect can be expected to be detected becomes less certain. Future work should contain a more detailed analysis of the precision with which ELF-MD can be predicted.

In addition to testing the benchmark circuits, simulations should also be done on "real world" circuits. This should be done because random patterns may not accurately simulate "normal operation" on a "real world" circuit. A hardware design

follows certain specifications and applying random patterns may violate the specifications. This leads to unpredictable results and may uncover defects that would otherwise be masked forever. Understanding what the hardware does may lead to a better simulation of "normal operation" and this in turn may lead to more accurate results when trying to determine ELF-MD.

Unfortunately, while software simulation data is very useful, it is far from perfect. Its main limitation is that it is time consuming to collect simulation data. Even with the techniques mentioned earlier to reduce simulation time, collecting the data is still time consuming. This limits the number of clock cycles that can be simulated and the complexity of the simulated circuit. Highly complex circuits are extremely time consuming to simulate and gathering data on hard-to-observe locations is difficult. Simulations for certain defect sites in circuit C2670 took as long as five days. Software simulation also restricts the defect types studied to those that are modeled as surrogates in the simulation. An even better understanding of ELF-MD can be accomplished if experiments are done using actual manufactured integrated circuits in hardware. Several hours worth of software simulation can be accomplished in only a few microseconds of hardware testing. Also, an actual manufactured IC would not be limited by our surrogate models. A manufactured IC could contain any type of defect. A hardware experiment would contain more meaningful data because it uses "real world" hardware and testing procedures.

A hardware experiment would consist of testing parts that were identified as defective during manufacturer testing. These defective parts are tested in parallel with a series of "gold standard" chips that have been thoroughly tested and are assumed to

be non-defective. By comparing the outputs of the defective and "gold standard" chips, we can obtain actual field detection of manufacturing defects results. These results can be compared with what our developed model predicts. This information can help us further refine and expand our model while possibly triggering more interesting questions to be investigated.

CONCLUSIONS

This research investigated what information is needed to predict the expected latency to failure due to manufacturing defects. ELF-MD could evaluate the quality of ICs and the test pattern sets that were used to test them. If test pattern set A produces a higher ELF-MD than test pattern set B, then test pattern set A could be considered better than test pattern set B. A model for predicting ELF-MD based on MPG-D was also introduced.

It was shown that the probability of observing a site and consequently detecting a defect during ATPG testing does not give enough information to predict ELF-MD for sequential circuits and AND-bridge defects. We can conclude from this that any ELF-MD predictor must include information about the circuit in "normal operation."

For that reason, this research presented a preliminary model which attempts to relate the following to ELF-MD.

1. The observability of circuit sites during normal circuit operation

2. The probability of exciting an undetected defect at one of those circuit sites given that it is observed and given that a certain number of observations of that site occurred during manufacture testing, and

3. The probability that a defect remains at that site

However, this model needs to be tested and further evaluated using both surrogate simulation and hardware experiments using actual defective circuits.

Developing a model to predict ELF-MD is not a simple task. The problem is quite complicated and there are still many aspects that have yet to be explored. However, the preliminary results presented collected by this research has given valuable

information and direction in formulating an initial model. Additional experimentation should lead to even more valuable insights.

Despite the complexities involved with predicting ELF-MD, the ability to predict the ELF-MD would be valuable to both manufacturers and customers. Manufacturers could adjust their maximum allowable defective part level for each product and use their testing resources better. If a company knew the estimated ELF-MD for their product, they could adjust their warranty period or to relax their defective part level requirement. This would turn products that were unnecessarily eliminated into profitable products.

REFERENCES

[1] T. W. Williams and K. P. Parker, "Design for testability—A survey" *Proceedings of the IEEE,* vol. 71, no. 1, pp 98-112 Jan. 1983.

[2] R.D. Eldred, "Test routines based on symbolic logic statements." *Journal of the ACM,* vol. 6, no. 1, pp. 33-36, 1959.

[3] T. W. Williams and N. C. Brown, "Defect level as a function of fault coverage," *IEEE Trans. on Computers*, vol. C-30, no. 12, pp. 987-988, 1981.

[4] S.D. Millman and E.J. McClusky, "Detecting bridging faults with stuck-at test sets," in *Proc. Int. Test Conf.*, 1988, pp. 773-783.

[5] P.C. Maxwell, R. C. Aiken, V. Johansen, and I. Chiang, "The effectiveness of I_{DDQ}, functional and scan tests: how many fault coverages do we need?" in *Proc. Int. Test Conf.*, 1992, pp. 168-177.

[6] S. C. Ma, P. Franco, and E. J. McCluskey, "An experimental chip to evaluate test techniques: experiment results," in *Proc. Int. Test Conf.,* 1995, pp. 663-672.

[7] P. Franco, W. D. Farwell, R. L. Stokes, and E. J. McCluskey, "An experimental chip to evaluate test techniques: chip and experimental design," in *Proc. Int. Test Conf.*, 1995, pp. 653-662.

[8] K. M. Butler and M. R. Mercer, "Quantifying non-target defect detection by target fault test sets," in *Proc. Eur. Test Conf.*, Munich Germany, April 10-12, 1991, pp. 91-100.

[9] Li-C. Wang, M. R. Mercer, and T. W. Williams, "On efficiently and reliably achieving low defective part levels," in *Proc. Int. Test Conf.*, 1995, pp. 616-625.

[10] Li-C. Wang, M. R. Mercer, T. Williams, S.W. Kao, "On the decline of testing efficiency as fault coverage approaches 100%," in *Proc. VLSI Test Symp.* Princeton, NJ, 1995, 74-83.

[11] Li-C. Wang, M. R. Mercer, T. Williams, "A better ATPG algorithm and its design principles," in *Proc. Int. Conf. on Computer Design*, 1996, pp. 248-252.

[12] R. Kapur, J. Park, M. R. Mercer, "All tests for a fault are not equally valuable for defect detection," in *Proc. Int. Test Conf.*, 1992, pp. 762-769.

[13] P. C. Maxwell, R. C. Aitken, V. Johansen, and I. Chiang, "The effect of different test sets on quality level prediction: when is 80% better than 90%?" in *Proc. Int. Test Conf.*, 1991, pp. 358-364.

[14] J. Park, M. Naivar, R. Kapur, M. R. Mercer, and T. W. Williams, "Limitations in predicting defect level based on stuck-at fault coverage," in *Proc. VLSI Test Symp.,* 1994, pp. 186-191.

[15] M. R. Grimaila, S. Lee, J. Dworak, K. M. Butler, F. Stewart, H. Balachandran, B. Houchins, V. Mathur, J. Park, Li-C. Wang, and M. R. Mercer, "REDO − Random Excitation and Deterministic Observation − first commercial experiment," in *Proc. VLSI Test Symp.*, 1999, pp. 268-274.

[16] J. Dworak, J. Wicker, S. Lee, M. R. Grimaila, K. M. Butler, B. Stewart, L-C. Wang, and M. R. Mercer, "Defect-oriented testing and defective-part-level prediction," *IEEE Design and Test of Computers*, January-February, 2001, Vol. 18, No. 1, pp. 31 - 41.

[17] J. Dworak, M. R. Grimaila, S. Lee, L-C. Wang, and M. R. Mercer, "Enhanced DO-RE-ME based defect level prediction using defect site aggregation – MPG-D," in *Proc. of the 2000 Int. Test Conf.,* Atlantic City, NJ, October 3 - 5, 2000, pp. 930-939.

[18] F. Brglez and Fujiwara, "A neutral netlist of 10 combination benchmark circuits and a target translator in FORTRAN," In *Proc. Int. Symp. Circuits Syst., 1985,* pp. *663-698.*

[19] F. Brglez, D. Bryan, and K. Kozminski, "Combinational profiles of sequential benchmark circuits," in *Proc. Int. Symp. Circuits Syst., 1989,* pp. 1929-1934.

VITA

David Michael Dorsey was born August 14th, 1980 in Mt. Holly, New Jersey. He graduated from Brazoswood High School in Clute, Texas in 1998. He started his college career at Texas A&M University in the fall of 1998. Starting at the beginning of his junior year in 2000, he started working as a year-round intern for Compaq Computer Corporation, later Hewlett Packard, where he worked until August 2003. During his senior year, David participated in the University Undergraduate Research Fellows Program. David worked with Dr. M. Ray Mercer during this program. In May of 2002, David graduated summa cum laude with a Bachelor of Science in computer engineering from Texas A&M University. He started his work on his Masters of Science degree in computer engineering in the fall of 2002 under the direction of Dr. Mercer. He can be reached at the following address: 102 Silverlace, Lake Jackson, TX 77566.