Data Analytics Approach for Optimal Qualification Testing of Electronic Components

Stoyan Stoyanov, Mominul Ahsan and Chris Bailey
Computational Mechanics and Reliability Group, University of Greenwich, London SE10 9LS, UK
Email: s.stoyanov@gre.ac.uk Phone: +44 (0)208 331 8520

Abstract

In electronics manufacturing, required quality of electronic components and parts is ensured through qualification testing using standards and user-defined requirements. The challenge for the industry is that product qualification testing is time-consuming and comes at a substantial cost. The work reported with this paper focus on the development and demonstration of a novel approach that can support "smart qualification testing" by using data analytics and data-driven prognostics modelling.

Data analytics approach is developed and applied to historical qualification test datasets for an electronic module (Device under Test, DUT). The qualification spec involves a series of sequentially performed electrical and functional parameter tests on the DUTs. Data analytics is used to identify the tests that are sensitive to pending failure as well as to cross-evaluate the similarity in measurements between all tests, thus generating also knowledge on potentially redundant tests. The capability of data-driven prognostics modelling, using machine learning techniques and available historical qualification datasets, is also investigated. The results obtained from the study showed that predictive models developed from the identified so-called "sensitive to pending failure" tests feature superior performance compared with conventional, as measured, use of the test data. This work is both novel and original because at present, to the best knowledge of the authors, no similar predictive analytics methodology for qualification test time reduction (respectively cost reduction) is used in the electronics industry.

1. Introduction

The global market for electronic products is projected to reach US$2.4 trillion per year by 2020 [1]. This growth has led to intense competition between manufacturers to minimise the time-to-market and cost of their products while at the same time delivering high quality products to the customer. Assuring the robust functional performance and quality of manufactured electronics products, and respective "fit-for-purpose" characteristics, requires the adoption of qualification processes, along with reliability testing, that often are time-consuming and resource-intensive [2]. Identifying solutions for how to overcome the challenges of meeting the quality requirements specified in a cost effective manner and within the shortest possible time can provide competitive edge for many electronics manufacturers.

Qualification is an application-specific process, which means that for different products in different applications, qualification specifications and requirements are different. Generally, qualification specifications of a product are developed by the manufacturers based on the application requirements as defined by the customer. The testing is used to determine whether the product meets the specified requirements and if operational functionality/performance over intended lifetime span (i.e. reliability) can be expected and achieved [3-4].

Qualification testing of electronics products is typically conducted through measurement of various electrical parameters that are indicators of the functional state of the individual electronic component or product. A qualification test outcome is typically binary and defined as either PASS or FAIL based on the measured test values and associated specified test limits. It is a common practice in the electronics industry to archive qualification test measurements, for example to ensure traceability information is available, and as a result companies often have access to large historical sets of qualification test data for their products. Such data can potentially hold valuable information that can support the optimisation of respective qualification test procedures, for example identifying favourable sequencing of the tests and if there are any potentially redundant test that are not required. Also, test time reduction may be possible to achieve by adopting data-driven, machine learning prognostics models capable of accurately forecasting the overall qualification outcome (Pass or Fail) for a DUT.

Machine learning offers a powerful approach to problems that require analysis and decision making based on historical data analytics but the advantages it can offer are not yet fully recognised in the domain of electronics manufacturing and the associated qualification testing in particular In recent years, as result of the increasing use of Internet of Things technologies, this has started to change and applications of machine learning algorithms have become more common and increasingly important in tasks related to achieving effective and optimal outcomes by means of data analytics [5-10].

The type of qualification specs targeted with this work is the common electrical parameter testing where tens or hundreds of individual tests are executed sequentially, one after the other. The proposed data analytics-based modelling methodology can enable gaining insights into the role and significance of individual tests and their sensitivity to being reliable precursors of a DUT's pending failure in the qualification process. Use of such results is discussed and demonstrated in the context of optimising the qualification test specification and how they can feed into and support the utilisation of in-line imbedded prognostics models. The adoption of such models can make possible to forecast the "PASS" or "FAIL" outcome for an electronic device under test.

2. Opportunities with Historical Qualification Data

The type of qualification testing considered in this work requires undertaking a series of individual, test-related electrical parameter measurements on the tested electronic device. The measured test value has to be within the specification range for the test parameter in order for that test to be passed. In a sequence of individual tests, for the device to be qualified it is required that all individual tests constituting the qualification specification are passed. When a DUT fails a test in the test sequence, continuing testing of that device under the remaining tests is not performed.

Many qualification procedures of this type require large number of electrical, logical and other functional parameter measurements, which for complex electronic parts can easily require individual tests in the order of hundreds. Hence, the overall qualification of a single electronic part can easily become time consuming given the need to perform electrical probing on such large number of parameters.

A potential way to shorten the qualification time is by taking advantage of the fact that as the testing progresses, more and more data of already completed individual tests becomes available. Through use of analytics and prognostics models on the data from such completed tests, an appealing prospect is to infer if remaining tests are likely to be passed or if one or more of them may fail. In essence, there is a clear opportunity to build machine learning models using past historical data on qualified electronic devices, and then embed the models in the qualification process for in-line use to forecast the qualification outcome.

The remaining sections of the paper detail the proposed, developed and demonstrated data-driven modelling methodology using real qualification test datasets gathered for an electronic module.

3. Computational Approach to Qualification Data Analytics

The methodology for using historical qualification test datasets aims at supporting smarter testing through optimisation of qualification test specs using knowledge that can be obtained through data analytics and from predictive modelling results. The approach is detailed with the diagram in Figure 1. Failure statistics and similarity test evaluation results can be derived offline if historical datasets are available. This information can enable the qualification optimisation by suggesting different order for the execution of tests in the sequential testing process and point what might be the potentially redundant tests in the spec. The other key opportunity with offline test data analytics is the generation of prognostic models using the data. These can enable the targeted in-line forecasting capability. The approach is to base this on test measurements from limited number of individual tests completed on a DUT, and predict the final, overall qualification status without executing the remaining tests beyond the point of the current test where the model forecast is made.

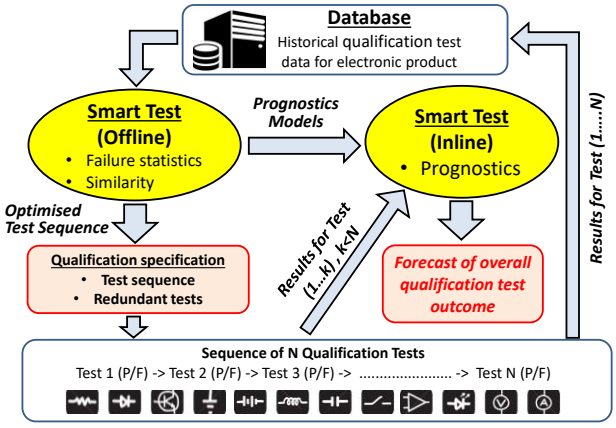

Figure 1: Approach to smart test of electronic products.

The developed numerical modelling approach is based on integrated techniques for statistical analysis of failure test data, distribution-based data modelling and data analytics for identification of (1) potentially redundant tests and (2) tests sensitive to pending failure. Machine learning based modelling for in-line qualification outcome prognostics are also used and integrated within the framework. The approach requires the following main data processing, analytics and modelling steps:

(a) The actual test data, containing records for both PASS and FAIL DUTs, is first separated into PASS and FAIL datasets, and then processed and filtered. Failure statistics for all tests using the dataset on FAIL devices is undertaken.

(b) The data is normalised over the range 0-1. This is identified as a requirement in order to enable subsequent data analytics, specifically in the context of similarity comparisons of the test data.

(c) The normalised test data of PASS and FAIL DUT datasets are used to derive, for each of the individual tests in the spec, distributions of the respective test measurements in the format of probability distributions.

(d) The Chi-square statistic and goodness-of-fit p-value are used to calculate, using the PASS data distribution for each test, how similar the data distribution is to the distributions of test data gathered from all other qualification tests. The statistical technique is used to rank all possible pairs of tests based on their "similarity". This knowledge, along with the test failure statistic information, can be used to support identifying (potential) redundant tests.

(e) The Chi-square statistic is used, similarly as in (c) above, to evaluate the similarity (respectively dissimilarity) in the measured test data in the instances of PASS and FAIL data for a given test. This analysis informs which tests are potentially sensitive (or not sensitive) to pending failure. The results can be used to inform on the existence and the potential of qualification tests to underpin the construction of predictive machine learning and fault classification models for test prognostics.

Figure 2 provides a diagram of the data analytics approach identified and developed in this study, as outlined in (a)-(e) above.

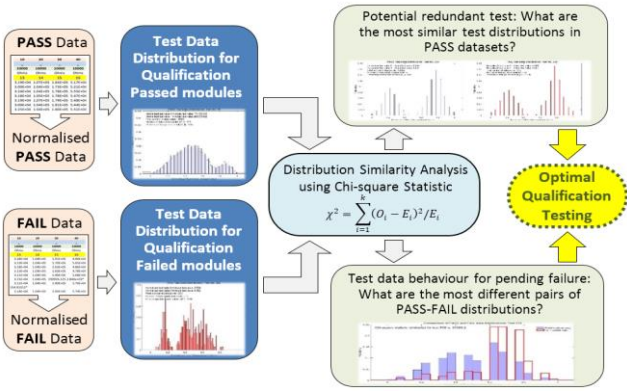

Figure 2: Data analytics approach for optimisation of qualification testing of electronic devices.

(f) Where possible, a revision of the qualification spec should be realised so that tests sensitive to pending failure are undertaken first in the sequential testing process. Machine learning techniques can be used to develop prognostic models that need only limited number of completed tests, those undertaken first and sensitive to pending failure, and offer predictive accuracy that is superior compared with other strategies for test data use.

Figure 3 illustrates the concept behind adoption of machine learning prognostic models within the qualification process execution.

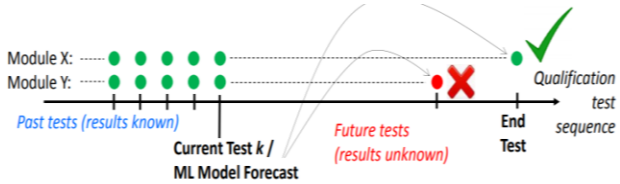

Figure 3: In-line prognostics for qualification outcomes for DUTs using Machine Learning (ML) models.

4. Qualification Test Data Modelling

4.1. Qualification Datasets used in the Study

Historical qualification test datasets for 50,000+ electronic modules, referred here as Device under Test (DUTs), are investigated as part of the reported research. The proposed numerical approach for data analytics is applied to assess a qualification procedure that requires 150+ individual qualification tests. As an example, large number of tests involve measurements that have the test outcome as real value numbers, for example voltage, current, time durations, power ratios, signal power strength and frequency. There are also tests for which the test parameter is Hex or integer value. There are also logical tests that output *True* or *False* values. Some of the tests have double sided limits for the PASS test condition and others are single sided.

Performing a numerical-based analysis for such a range of diverse tests that follows a generic (non-specific to the test) computational approach is challenging. Following preliminary investigations of the datasets, it is decided to develop the data analytics approach on the basis of distribution modelling of the qualification test data and mining the data behaviour/relationships using suitable techniques. Such approach can offer robustness and generalisation of the proposed computations which are judged to be the two most important attributes of the proposed Smart-Test framework.

The data analytics studies are undertaken only on a subset of tests for which the test result varies and can be modelled as a distribution. With the qualification spec of the investigated electronic module data, there are 111 qualification tests out of the total 150+ tests which meet this selection criterion. All of the following studies detailed in the paper use test data that is gathered only from these 111 tests. Remaining tests are excluded from further observations.

Figure 4 shows a simplified, illustrative sample of near-raw measured tests data, in the format of normalised values, from the qualification of the investigated electronics module. The measured parametric values for each sequential test are arranged column wise, and test results for each electronic module appear in a row of the presented table. It should be noted that no further test is carried out once a module is failed. Hence, no test data will be available for a module onwards from the test of failure as indicated by a value with an asterisk (*).

	Test Number									
	1	2	3	4	...	10	11	...	N	Known outcome
Upper limit	1	1	1	1	...	1	1	...	1	
Lower limit	0	0	0	0	...	0	0	...	0	
Module 1	0.11897	0.88349	0.48914	0.53865	...	0.18535	0.48090	...	0.02012	PASS
Module 2	0.08448	0.99432	0.55154	0.56794	...	0.17126	0.57001	...	0.02273	PASS
Module 3	0.08046	0.91191	0.51449	0.53167	...	0.14965	0.65764	...	0.01976	PASS
Module 4	0.03448	0.90622	0.41700	0.43496	...	0.16493	0.54107	...	0.02335	PASS
Module 5	0.13594	0.88349	0.18204	0.22847	...	-1.11831				*FAIL*
Module 6	0.08851	0.99432	0.55154	0.57819	...	0.18305	0.57365	...	0.02172	PASS
Module 7	0.08621	0.91191	0.49792	0.50350	...	0.18197	0.65706	...	0.01984	PASS
Module 8	0.08793	0.99147	0.52619	-0.92373	...					*FAIL*
Module 9	0.09195	0.99147	0.45112	0.51727	...	0.12586	0.57273	...	0.02080	PASS
Module 10	0.12586	0.88349	0.48524	0.54832	...	0.20171	0.47633	...	0.02104	PASS
Module 11	0.03736	0.90338	0.43552	0.45752	...	0.12853	0.53894	...	0.02517	PASS
Module 12	0.09195	0.99147	0.50179	0.54334	...	0.16112	0.57304	...	0.02041	PASS
Module 13	0.08793	0.91191	0.52814	0.54480	...	0.16745	0.65678	...	0.02014	PASS
Module 14	0.12701	0.88349	0.51449	0.56677	...	0.22218	0.47361	...	0.02074	PASS
Module 15	0.03966	0.90338	0.42187	0.42354	...	0.15133	0.54013	...	0.02167	PASS
Module 16	0.09195	0.99147	0.49283	0.53279	...	0.11682	0.57115	...	0.02096	PASS
Module 17	0.08851	0.91191	0.45405	0.46513	...	0.16317	0.65714	...	0.02278	PASS
Module 18	0.12644	0.88349	0.44430	0.52459	...	0.19251	0.47399	...	0.02379	PASS
Module 19	0.07759	0.88633	0.51254	-0.99374	...					*FAIL*
Module 20	0.09310	0.99147	0.49889	0.52723	...	0.17413	0.57326	...	0.02128	PASS
Module 21	0.08908	0.91191	0.47939	0.49403	...	0.15793	0.65560	...	0.02003	PASS
Module 22	0.12471	0.88349	0.37313	0.48915	...	0.20590	0.47338	...	0.02207	PASS
...	...	...	...	...	...	...	...	...	...	...

Figure 4: Illustrative example of qualification test data in normalised format.

4.2. Distribution Modelling of Test Data

4.2.1 Normalization of Test Data

Measurements from different qualification tests are different. Some measurements are numerical decimal continuous values, others give the result as an integer number or as a Hex value. The magnitude/order of the measured value (where numerical) can also be very different. Measurement units from test to test change. Some tests are double side limited, some have a limit only on one side. The best strategy in numerical analysis to handle such differences is to subject the data to normalisation. The normalisation scheme used in this

study makes data transformation using normalised limits of 0 and 1.

Accounting for the overall approach and the need for a robust data handling, the following 2-step normalisation strategy is implemented:

1. The raw test data is first processed and filtered into two, PASS and FAIL, datasets. The data is then screened for extreme outliers. Such extreme outliers are filtered out from the data.

2. The data for DUTs that PASS the qualification is associated with the "normal", expected test behaviour in the context of measured values. For each test, a strategy to decide on the actual lower and upper limit values for the data normalisation over 0 to 1 is required. The values for normalisation are selected as a percentile of the entire PASS data set for a given test. After some testing, the low limit of the data was selected as the 0.1 percentile and high limit as the 99.9 percentile of the data.

4.2.2 Distribution Modelling

Modelling the distributions of the normalised data for each test enables to observe the behaviour of the data (how the data is spread, the nature and magnitude of variation, etc.) and to compare qualification tests. Examples of modelled test data distributions for the DUT using the histogram (binned data) approach are illustrated in Figure 5. Note that normalised test values less than 0 are binned in a single bin and similarly a single bin holds all values above 1. Detailed distribution is generated within the 0-1 interval. The vertical axis of the histogram charts denotes 'probability'.

The test from the sequential qualification spec detailed in Figure 5 is given as an example to show how the historical measurements results from that test, on DUTs that passed and failed the overall qualification, are modelled and prepared for subsequent use and evaluation. Note that the normalisation of the latter (FAIL) dataset is based, as explained previously, on the same actual lower-upper limits for data normalisation derived from the PASS-dataset for the test.

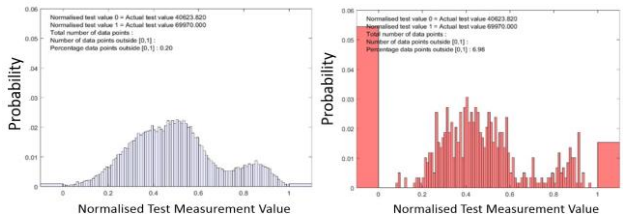

Figure 5: Example of data distributions for a given test gathered from PASS-status electronic modules (left) and FAIL-status electronic modules (right).

Note that a test data distribution associated with failed modules, as with the example of the right diagram in Figure 5, is based on the available data for that test. These are limited amount of measurements from failed devices which successfully passed the test under consideration (and indeed all previous tests) but have failed at a subsequent test in the procedure. The interest in these distributions is to see if there is some distinctive behaviour in the test data preceding the one at which failure has occurred.

To strengthen our confidence in accuracy of these FAIL test data distributions, larger dataset of failed modules is ideally required. The limited amount of fail data in this study should be noted in the context of all subsequent observations and results reported in the paper.

4.3. Data Similarity Assessment

The availability of data distributions enables understanding the behaviour of the qualification test measured datasets. The use of histograms is convenient as it enables also comparing different qualification tests, and also PASS and FAIL data for a given test. By comparing, in a quantitative manner, how similar or different are two data distributions, important observations and conclusions regarding a qualification test procedure can be made. The quantitative approach to similarity assessments of data distributions is based on the use of Chi-square statistic.

In statistics, the Chi-square goodness-of-fit test is often used to test if a sample of data comes from a population with a known distribution. An attractive feature of the Chi-square test is that it can be applied to any univariate distribution for which the cumulative distribution function can be calculated. The Chi-square goodness-of-fit test is always applied to binned data. As in our case the distributions are already in the format of histograms, the application of Chi-square statistic technique is very convenient and straightforward.

With standard use of Chi-square test goodness-of-fit, the hypothesis that a set of data (i.e. observed values) comes from a population with a given (specified) distribution (i.e. expected values) is tested. This assessment uses the Chi-square test statistic χ^2 which is defined as

$$\chi^2 = \sum_{i=1}^{k} (O_i - E_i)^2 / E_i$$

where Oi is the observed frequency of data for bin i and Ei is the expected frequency for bin i. In the case of using Chi-square test for goodness-of-fit evaluation, expected frequency refers to the given distribution and is calculated as

$$E_i = N(F(Y_U) - F(Y_L))$$

where F is the cumulative distribution function for the distribution being tested, Y_U is the upper limit for bin with index i, Y_L is the lower limit for data bin i, and N is the sample size of the observed data.

The goodness-of-fit part of the computation is based on the fact that χ^2 follows approximately Chi-square distribution with (k-c) degrees of freedom where k is the number of bins and c is the number of estimated distribution parameters (c=1 in this study). With a specified significance level α, a Chi-square critical value ($\chi^2_{1-\alpha,k-c}$) is obtained using the Chi-square distribution with (k-c) degrees of freedom. The test statistic and the critical value are used to check the condition

$$\chi^2 > \chi^2_{1-\alpha,k-c}$$

If this relation is true, then the hypothesis that the data are from a population with the specified distribution is rejected.

In this work we adapt the use of Chi-square test statistic calculation to meet the requirement to have a similarity measure, or similarity index, that shows how similar is the test data distribution given with one histogram to the data modelled with another histogram. The use of a simple metric termed Similarity Index (*SI*) is proposed. The *SI* is the Chi-square statistic value χ^2 normalised to the Chi-square critical value $\chi^2_{1-\alpha,k-c}$ assuming significance level α=0.01 and (k-1) degrees-of-freedom where k is the number of non-empty bins in the histogram pair:

$$SI = \frac{\chi^2}{\chi^2_{1-\alpha,k-c}}$$

Larger values of S*I* indicate less similar datasets given with a pair of histograms while small values of *SI* are indicators of greater similarity in the respective datasets. In this study, the proposed similarity index is used to rank the qualification tests given similarity between their respective PASS and FAIL data distributions.

4.4. Sensitivity of Tests to Pending Failure

In order to apply some form of computational intelligence to the data to enable in-line prognostics prediction for the overall qualification test outcome, distinctive data relationships and data behaviour of the electronic modules that are successfully qualified and those that fail a test must be present. From data analytics point of view, this would mean the distribution of PASS data for a given test will need to be different to an extent from the data distribution for the same test that is associated with FAIL devices.

The rationale to use these computational evaluations in order to imply certain test significance in the context of this test indicating pending failure is based on the following notion: if there is similarity between PASS and FAIL data distributions of a qualification test then the test does not produce data with behaviour which can support prognostics computations. Similarly, if the data distributions of PASS and FAIL data differ then the test can be seen as generating measurements that can be potentially usable in the context of predicting the qualification outcome. By calculating the Similarity Index for each pair of PASS and FAIL data using the respective data histograms, across all analysed qualification tests, it is possible to rank the tests with regards their data distribution similarity and thus assess their sensitivity to detect pending failure.

Figure 6 shows two representative tests with dissimilar pairs of PASS and FAIL data distributions. The difference between the PASS data and the FAIL data distributions is illustrated by overlapping the two histograms and the value of the *SI* is included with each graph. Larger SI value means greater difference in the test PASS and FAIL data distributions.

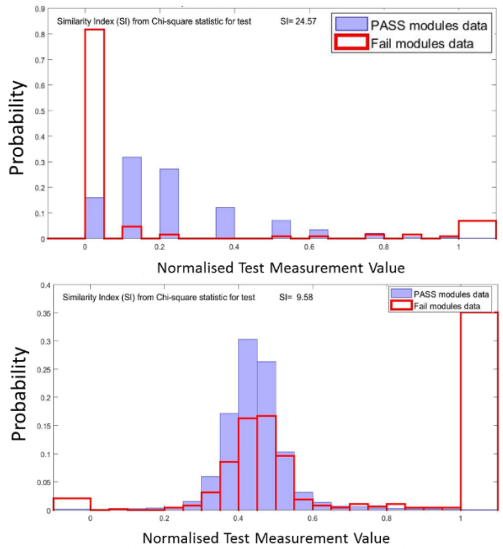

Figure 6: Example of two qualification tests with dissimilar distributions of PASS and FAIL module data indicating tests are sensitive to pending failure.

Tests with the smallest *SI* are those tests for which distribution of PASS and FAIL data are similar. We can consider such tests as being less sensitive to detecting pending failure on the basis of the test result. The test outcomes from testing good modules and modules that fail the qualification do not differ notably and hence test measurement data do not contain useful information to support prognostics modelling. Figure 7 details two examples of similar pairs of PASS-FAIL data distributions taken from the full set of 111 analysed qualification tests.

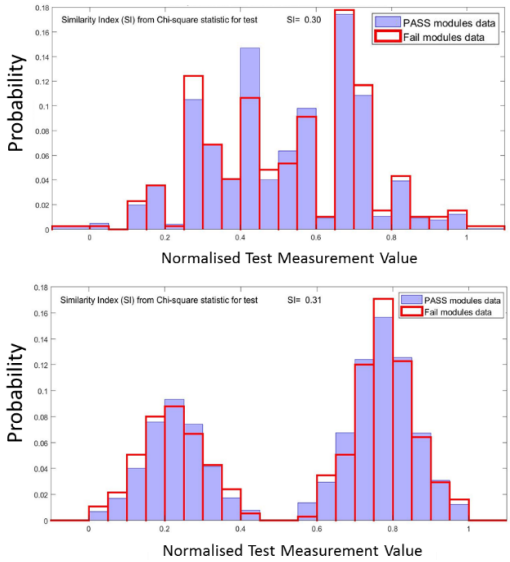

Figure 7: Example of two qualification tests with similar distributions of PASS and FAIL module data indicating tests are not sensitive to pending failure.

5. In-line Prognostics for Qualification Testing

5.1. Machine Learning Approach

The main benefit of the data analytics detailed in the previous section is the knowledge generated on the tests defined as being sensitive to pending failure. This information can feed into and support what is expected to be potentially more robust and efficient predictions from developed prognostics models.

In this work and for the application discussed, a prognostics model is a data-driven model that takes as input a number of measurements from tests already completed, and hence available, for a DUT. The model then predicts the expected final outcome of the qualification, i.e. makes a prediction if the DUT will pass all remaining tests in the test sequence (tests yet to be undertaken for the DUT) or will it fail at any of these coming tests and hence will fail the qualification.

The prognostics models considered here are only data-driven models that can be built from the available tests data and executed at a given test in the sequential procedure, as chosen by the user. In particular, machine learning type of algorithms are of prime interest. In this paper, the demonstrations rely on the use Support Vector Machine (SVM) model for binary classification [11].

As with all machine learning methods, a model is developed from so-called training dataset so that the model structure parameters that are unknown are calculated through solving an optimisation problem that provides the smallest error between the model predictions and the actual target outcomes. The predictive accuracy and performance are validated on a separate, so-called validation, datasets.

5.2. Study Case: Prognostics Performance with Tests Sensitive to Pending Failure

The study presented in this section evaluates the benefit of the approach of using tests sensitive to pending failure and verifies the advantages of the proposed data analytics. The machine learning models developed are based on sizes of training datasets of 1,070 DUTs and the validation data sets include 150 DUTs. For both training and validation datasets, the number of the DUTs that PASS the qualification is same as the number of DUTs that FAIL the qualification (50:50 split in the data). This is an important requirement to ensure balanced information is provided when the model is developed and constructed as well as to ease the comparisons of the models' performance.

Two SVM models are developed, each using information from 20 completed tests out of the total 111 tests in the sequential test procedure. A practical way to approach this model developments and allow for model comparison in a like-to-like manner, is to assume the scenario when the DUT's have passed the first 40 tests in the sequence of tests, and test No. 40 is the current test. We use the Similarity Index (SI) results from the data analytics investigation and split the first 40 tests into two equal groups of tests with size 20 tests each. In the first group we select the 20 most sensitive tests, according to their SI, among the originally sequenced tests (test 1 to test 40) in the spec. Similarly, the second group represents the 20 tests identified as being the least sensitive to pending failure among the first 40 tests.

With each of the two groups of tests, in an identical manner (using exactly the same training datasets, as size and as DUT data records, and same cross-validation of generated models) we obtained the best preforming SVM model for each of the two options. The inputs to each model are the 20 test results (measurements) from the respective tests, which, assuming current test completed is No. 40, are all available results. With the training data we use the known final qualification outcome, FAIL (0) or PASS (1), to construct the models. The SVM models make binary predictions, 0 or 1, for a given input of measurements on the respective 20 tests for the SVM model.

The model predictive accuracy for the SVM that uses 20 sensitive to pending failure tests is detailed in Figure 8 (left) in the form of confusion matrix plot. This is an average result obtained from validating the model on large number (in this instance 100) different validation datasets.

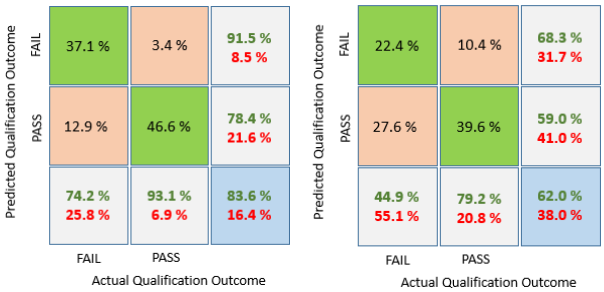

Figure 8: Prognostics performance of two SVM models each using as input results from 20 qualification tests (in range of tests 1 to 40): (1) most sensitive to pending failure tests (left) and (2) least sensitive to pending failure (right).

On the confusion matrix plot, the rows correspond to the predicted qualification status with the SVM (Predicted Qualification Outcome), and the columns show the actual (verified with testing) qualification status (Actual Qualification Outcome). The diagonal cells show the percentage of DUTs for which qualification is predicted correctly (actual and predicted qualification status match). The off-diagonal cells detail the % of DUTs which are not classified correctly. The far right column shows the accuracy for each predicted status and the bottom row shows the accuracy for each actual outcome. The cell in the bottom right of the plot shows the overall accuracy, in this instance 83.6%.

In a similar way, the results from the SVM model with identical complexity and same size of input information, but now based on 20 least sensitive to pending failure tests in the range of tests from 1 to 40, are summarised with Figure 8 (right). Clear difference in the predictive capability between the two models is observed. With this prognostics model, built with data from tests which are

not sensitive to pending failure, the prediction accuracy for the DUTs qualification status has decreased dramatically to 62% only.

This study confirms that the approach of formulation and identification of the so-called tests sensitive to pending failure, on the basis of similarity index attributes, is a key, integral part of the in-line prognostics strategy to smart-test execution, and offers clear improvement in the accuracy of the constructed machine learning models.

6. Conclusions

This study aimed at the formulation and the development of a novel computational approach that can be used to optimise qualification testing of electronic products by reducing test times and cost through off-line data analytics and imbedded in-line model-based prognostics. The proposed approach and the associated models were developed and tested with rigour using comprehensive datasets of real historical qualification data on an electronic module.

Major capability that is seen as new significant development within the proposed unified approach is the formulation and evaluation of qualification tests' sensitivity to pending failure. The results showed that there are tests for which test data behaviour differs when devices pass the qualification compared with the case when they fail. This was an important finding for the behaviour of the analysed available test datasets and as data attribute is expected to be present in most qualification test data. In the context of imbedded in-line prognostics capability in electronics product qualification, the existence of sensitive to pending failures tests can support efficient use of test data with machine learning algorithms in forecasting the likely outcomes for tested devices in the overall qualification tests procedure. This approach has been demonstrated successfully with the analysed test data, and the time and cost benefits of the proposed test optimisation strategy were clearly presented and validated.

The proposed approach to optimise electrical and functional qualification test specifications of electronic devices through use of data analytics and machine learning techniques, and by adopting imbedded model-based prognosis for qualification test outcomes has the potential to transform the current practices in the industry of undertaking comprehensive and time consuming testing. The major impact of this research is associated with the clear benefits of adopting the discussed technologies in terms of cutting cost and time of qualification testing. The Smart-Test solutions demonstrated in this work are feasible to realise in any application that can tolerate the accuracy of the machine learning model forecasts associated with the test data, and where the flexibility to design/optimise existing and future product qualification test specifications is present.

Acknowledgments

This work is undertaken under Smart-Test research project funded by University of Greenwich (UK) in close collaboration with Microsemi (www.microsemi.com). The authors thank Microsemi Corporation for providing access to real qualification test datasets and for the valuable discussions on the subject of this work.

References

1. Pecht, M., Shibutani, T., Kang, M., Hodkiewicz, M. and Cripps, E., "A Fusion Prognostics-based Qualification Test Methodology for Microelectronic Products", *Microelectronics Reliability*, Vol. 63 (2016), pp. 320-324.
2. Ruidong, D. and Chun, G., "Researches and Investigation on Manufacturers Reliability Test Data of Electronic Parts", *Energy Procedia*, Vol. 127 (2017), pp. 242-246.
3, Wang, W., Azarian, M.H. and Pecht, M., "Qualification for Product Development", *International Conference on Electronic Packaging Technology & High Density Packaging*, Shanghai, 2008, pp. 1-12.
4. Vichare, N., Zhao, P., Das, D. and Pecht, M., "Electronic Hardware Reliability", In: *Avionics: Development and Implementation*, Edited by Cary R. Spitzer, CRC Press, 2006, pp. 414-422.
5. Stoyanov, S., Bailey, C. and Tourloukis, G., "Similarity Approach for Reducing Qualification Tests of Electronic Components", *Microelectronics Reliability*, Vol. 67 (2016), pp. 111–119.
6. Park, S.H., Park, C.S., Kim, J.S., Kim, S.S., Baek, J.G. and An, D., "Data Mining Approaches for Packaging Yield Prediction in the Post-fabrication Process", *IEEE International Congress on Big Data*, Santa Clara, CA, 2013, pp. 363-368
7. Kupp, N. and Makris, Y., "Integrated Optimization of Semiconductor Manufacturing: A Machine Learning Approach", *IEEE International Test Conference, Anaheim*, CA, 2012, pp. 1-10.
8. Kim, H.G., Han, Y.S. and Lee, J.H., "Package Yield Enhancement using Machine Learning in Semiconductor Manufacturing", *Advanced Information Technology, Electronic and Automation Control Conference* (*IAEAC*), Chongqing, 2015, pp. 316-320.
9. Chou, P.B., Rao, A.R., Sturzenbecker, M.C., Wu, F.Y. and Brecheret, H.B., "Automatic Defect Classification for Semiconductor Manufacturing", *Machine Vision and Applications*, Vol. 9 (1997), pp. 201-214.
10. Boubezoul A., Annanou, B., Ouladsine, M. and Paris, S., "Defective Wafer Detection Using Multiple Classifiers", *Proc. European Control Conference*, Kos, Greece, July 2-5, 2007, pp. 2929-2934.
11. Christianini, N. and Shawe-Taylor, J. "An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods", Cambridge, UK: Cambridge University Press, 2000.