[1]A Test Station Health Monitoring System

Kevin T Fitzgibbon
Total Quality Systems, Inc.
1990 West 2550 South
Ogden, UT 84401
801-731-2150
kfitzgib@total-quality-systems.com

Larry V. Kirkland
OO-ALC/TISAC USAF
7178 4th Street
Hill AFB, Utah 84056-5205
(801)-775-4076
Larry.Kirkland@hill.af.mil

Bryan Steadman
Total Quality Systems, Inc.
1990 West 2550 South
Ogden, UT 84401
801-731-2150
bsteadma@total-quality-systems.com

Tony Pombo
Total Quality Systems, Inc.
1990 West 2550 South
Ogden, UT 84401
801-731-2150
tpombo@total-quality-systems.com

Abstract—This paper presents a process to monitor test station health using the Weibull method and statistical patterns. The methodology is currently being applied to the F-16 automated test equipment (ATE) at the Ogden, Utah Air Logistic Center (OO-ALC) maintenance depot. An automated stream of test data collected from ATEs is used to process test results and to identify improvements necessary to increase the failure forecast accuracy. The paper discusses solutions to identify causes of 're-test OK' (RTOK) due to discrepancies between software testing procedures in the line and shop repairable units. The process includes a decision support system that uses artificial intelligence methods, such as expert system and neural networks, and a knowledge database to improve the troubleshooting capability. The paper also discusses a prototype development that collects malfunction codes (MFL) originated by the aircraft bus monitoring system. The MFL information is correlated with test results to detect RTOK causes.

ACRONYMS

ATE – automatic testing equipment
CND – cannot duplicate
KB – knowledge base
LRU – line repairable unit
MC – mission capability
MCR – mission capability rate
MFL – maintenance fault list
MTBF – mean time between failures
MTTF – mean time to failure
OO-ALC – Ogden Utah Air Logistic Center
RTOK – re-test OK
SPC – statistical control process
SRU – shop repairable unit
TSHMS – test station health monitoring system

TABLE OF CONTENTS

1. INTRODUCTION

This methodology grew out of a Small Business Innovation Research (SBIR) contract sponsored by the Ogden-ALC Technology and Industrial Support Directorate. The Ogden-ALC Avionics and Electronics Repair Directorate as well as the SBIR Program Office have funded the current product enhancement phase jointly.

The original objective of the SBIR effort was to apply statistical control processes and artificial intelligence techniques to test results output from automated test equipment (ATE) in order to determine if the effects of aging and environmental factors could be detected on units under test (UUT) and the ATE itself. In order to reach the SBIR objective, several methods and techniques had to be developed or modified. The first was a means to capture, parse and archive the raw test program results in a relational database for automated retrieval and processing. Then the statistical process algorithms were modified from a controlling application to a form appropriate for monitoring. A variety of different statistical patterns, using rule based Expert System techniques, were evaluated for applicability to the present purpose. Those selected are discussed in greater detail later in this paper. Finally the added benefit of including failure and repair data to the developing automated analytical process that would extend the original objective to a broader failure-forecasting tool that will aid ATE operators and repair technicians, was realized. In order to make progress toward this expanded objective and to utilize Weibull methods for forecasting failures, a dependable source of failure and repair data is required. We turned to other ongoing efforts to reduce total ownership costs for sustainment of the F-16 Fighting Falcon, the largest fleet of combat aircraft in the world. The process developed under Acquisition Reform flexible sustainment concepts to identify, analyze and implement high return on investment (ROI) solutions for avionics systems high cost drivers was named Falcon Flex. The Falcon Flex maintenance and repair database, while not optimum, was the most dependable and defensible source of the needed data available at the time. Still some important failure data is not being effectively captured, such as assembly operating time, environmental stress levels and cause of failure (failure mode) by any existing data system. This paper discusses how some of these data shortcomings have

[1] 0-7803-6599-2/01/$10.00 © 2001 IEEE

been addressed and concludes with current and future research and development efforts required to refine the data capture process and thereby improve failure forecasting of ATE assemblies.

The following sections will present a test station health monitoring system (TSHMS) developed for the Ogden Air Logistic Center (OO-ALC) in Utah. Experimental results are presented without actual identification of part numbers, serial numbers, and station identifications for security reasons.

2. SYSTEM SPECIFICATION

The first requirement to implement a test station health monitoring system is the capability to collect test results from the station. Old legacy stations do not have this capability built in. Test results can be collected using 'dumb terminal' scripts that write the results to a local file, or by the testing software that writes directly to a data repository subsystem (database). The main problem in either case is the data format. There are currently no standards that apply to test output format. As a result, the test data might not be uniform and will require additional parsing or data processing. We will assume in this paper, that raw test data is formatted and available in appropriate format. The importance of test station health monitoring systems in diagnostics is critical to the capability of isolating actual causes of failure in cases where the cause can be either the station instrumentation, or the unit under test, or stack tolerance problems. Diagnostics can be supported using models and/or actual test and repair information. Repair data can also be used to identify causes of failure, or failure modes. Operating time, failure history, or failure modes are necessary to forecast failures. Actual failure and repair data can be used to feedback for adaptive models. Models are useful to provide information for cases where no previous failure or repair information is available.

3. DATA ANALYSIS METHODS

There are three types of data used in the TSHMS: 1) test data, 2) repair data, and 3) failure data. Test data consists of test measurements and test results. Repair data consists of component replacement information. Failure data consists of all data related to the component's operation and failure. Data analysis algorithms are applied to raw test, repair, and failure data. These algorithms consist of statistical methods such as statistical process control (SPC) rules and failure analysis methods. SPC algorithms were modified for monitoring (instead of controlling) purposes.

3.1 Test Data

Test data consists of test measurements output by the testing software. Test data includes: item identification (part and serial number), test date and time, test code and name, measurement upper and lower limits, units of measurements, and the measurement value. Table 1 presents a typical set of test results.

Table 1 – Typical Test Data

Part Number	Test Date	Test Time	Test Code	Test Name	Upper Limit	Lower Limit	Units	Measurement
12609	08-Dec-99	11:12AM	T.4	28VDC SELFTS	28.5	26.5	V	27.883
12609	08-Dec-99	11:12AM	T.6	+15VDC S/T	15.5	14.7	V	15.158
12609	08-Dec-99	11:12AM	T.7	-15VDC S/T	-14.7	-15.5	V	-14.979
12609	08-Dec-99	11:12AM	T.31	K1-A CLOSE	5	9.99E+37	OHM	1.1247
12609	08-Dec-99	11:12AM	T.37	K2-A CLOSE	11.7	6	OHM	7.971
12609	11-Dec-99	3:46PM	ST1-40	Q1 7.5AMP	7.75	7.4	A	4.81394
12609	11-Dec-99	3:46PM	ST1-41	Q2 7.5AMP	7.75	7.4	A	7.52503
12609	11-Dec-99	3:46PM	ST1-1	+5V PS9	5.15	4.85	V	5.0264
12609	11-Dec-99	3:46PM	ST1-2	+15V PS3	15.3	14.7	V	14.937

Statistical analysis (presented in section 3.4) is performed using measurement values versus time. In our F-16 application, test measurements are grouped by part and serial numbers, test code, and test name. Test results presented in Figure 1 belong to an actual F-16 avionic system.

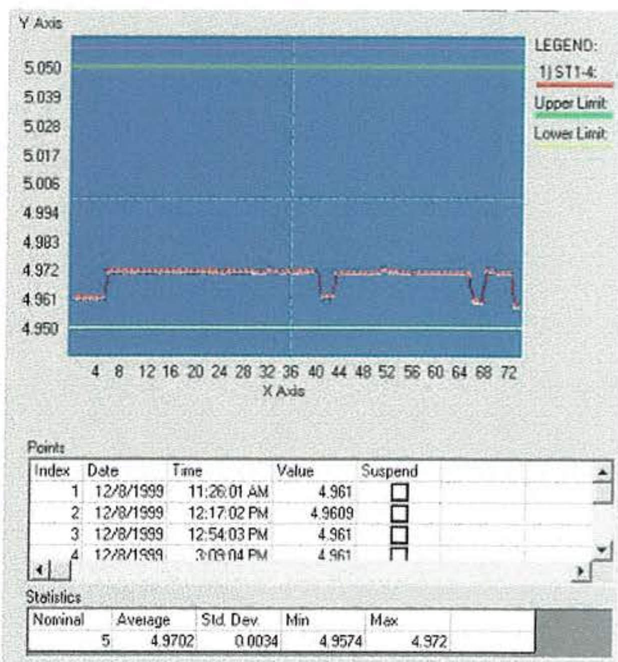

Figure 1 – Typical Test Measurement Results Along Time

In order to avoid possible disclosure of classified information the actual item's noun, part and serial numbers, and failure mode information were omitted. The graph presents 74 data points of 5 VDC power supply measurements collected over a period of 16 months. Step function events are clearly identified between periods of relative low variability. A consistent 'bias' towards the lower limit is also clear. The sensitivity (standard deviation) of the actual signal can be compared to the upper and lower limits included in the plot.

3.2 Repair Data

Repair data consists of all information related to repair actions after a failure is detected during the test phase. It usually includes: part number, serial number, repair date, description and item repaired. Table 2 presents typical repair data records. Repair information populates a Knowledge Base (KB) that supports diagnostic procedures. The KB accumulates repair history for items identified by part and serial numbers. Test failures are linked in real-time to the KB and provide support to ATE operators and repair technicians. Past failure and repair information can be retrieved. For 'never-before' cases the failure model provides the support. Failure models are adjusted using the KB data. In our specific F-16 application, the **Repair Date** field in Table 1, which corresponds to the 'date of repair', is often used as the 'date of failure'. In this case, there is an unknown (undocumented) time gap between the actual date of failure (in the aircraft) and the date of repair. The **Repair Date** is the best representation for failure datum available. This implies that our failure estimate method includes inaccuracies due to the uncertainty of the failure datum. The **Item Repaired** field in Table 2, which represents the items that were repaired, adjusted or replaced, is often associated with failure modes. This information misrepresents the actual cause of failure, or the failure mode. The accuracy of the failure forecast is also affected by the uncertainty of the failure modes.

Table 2 – Typical Repair Data

Part Number	Serial Number	Repair Date	Repair Description	Item Repaired
12609	C102	8/27/1998	Logic Assembly #2	U6
12609	C102	6/11/1998	Logic Assembly #2	Resolder P1
12609	C102	9/19/1998	Logic Assembly #2	U6
12609	C102	3/2/1998	Logic Assembly #2	U6
12609	C102	2/2/1998	Logic Assembly #1	Repair pin
12609	C102	1/26/1998	Logic Assembly #1	U6
A0762-5	E0924	12/9/1997	Power Supply #4	Q2, 2N2222A
A0762-5	E0924	5/19/1998	Power Supply #4	Q7, 2N2905A
A06738-7	005823	11/25/1997	Waveform Generator #1	U14
A06836-4	C0034	4/22/1998	Input/Output #2	U27

The failure analysis method described in section 3.3 uses the **Repair Date** as failure datum ('ages'). In order to convert the dates to ages, the initial date of operation (or service) must be known. The age is counted from the initial date of operation until the date of failure. In our F-16 application, the date of initial operation was not available and affects the prediction of future failures.

3.3 Failure Data

Typical failure data includes time of failure, operating conditions, failure modes, and induced failures. The Weibull analysis method is a well-known and mature method to estimate the mean-time-to-failure (MTTF) given failure data (failure times). The MTTF is frequently referred as the mean-time-between-failures (MTBF). The Weibull model is discussed in several references such as Refs. 1-6. The Weibull model includes the most common distributions usually associated with a life cycle including the exponential and the normal distributions. Weibull parameters are used to forecast failures using direct forecast or probabilistic estimating methods such as Monte-Carlo simulations. The Weibull parameters are bounded depending on the confidence level selected by the user. Several algorithms and methods can be used to define the confidence bounds for the Weibull parameters. The binomial method is used in this work. The binomial method, described in Ref. 1, calculates the upper and lower bounds for given failure data points. Intermediate bound data points are calculated using linear interpolation. The confidence bounds can be used to determine the accuracy of a failure forecast. The uncertainty in the MTTF is calculated using the two confidence bounds for a given confidence level.

3.4 Statistical Analysis of Test Data

The statistical method referred to in this paper was derived from the statistical process control (SPC) method and adapted to 'monitoring' instead of 'controlling' process. The method uses the concept of 'patterns', which are groups of data points that satisfy certain conditions or have specific properties. The main difference between the SPC and the patterns used in this method are the threshold definitions that are based both on statistical parameters and user-defined limits. The following patterns are used in our method:

Out-of-Limits – one data point above the upper limit or below the lower limit

Close-to-the-Limit – a group of M consecutive data points (out of N total data points close by less than P% from the limit (upper or lower); P is a percentage of the range

Low Variability – a group of M consecutive data points (out of N total data points) has a standard deviation less than the overall one-sigma (this pattern is used in the identification of step function patterns)

Step Function – the average of a group of K consecutive data points with statistical 'low variability' is separated by a minimum 'distance' from the average of a second group of J consecutive data points with statistical 'low variability'

Trend Line – the linear regression model of a group of M consecutive data points

A number of other patterns not listed in this work were investigated and later ignored since they did not contribute to this performance monitoring and failure-forecasting problem. In this work the linear regression was used to calculate the parameters of the 'trend line' pattern. Models other than the linear model can be used to characterize the trend line pattern. Patterns can indicate performance status and detect degradation. The existence of certain test result patterns can indicate that the performance of a unit is degrading and thus failure probability is increased. For example: 'close-to-the-limit' patterns can anticipate an imminent failure: 'Trend line' patterns can be extrapolated to intercept limit lines and estimate failure events.

3.5 Failure Forecast

The method used in our TSHMS combines the results from the Weibull and the statistical analysis to estimate failure dates and the associated uncertainty. The Weibull method provides the MTTF, and the upper and lower bounds associated with a confidence level. Given the date of last failure, the estimated date of next failure is given by adding the MTTF. The uncertainty is given by the upper and lower bound times added to the date of next failure. The statistical method uses the trend line pattern to extrapolate the time to the next failure. The variance of the fitting line provides the uncertainty of the estimated failure time. In this case, calculation of the trend line parameters is made using only the data points after a previous failure or other event that indicates a 'reset' in the life cycle, i.e., the cycle for that failure must be re-initiated. Typical patterns that indicate a reset are 'step function' and 'out-of-limit' patterns.

3.5 Test Pass/Fail Criteria

Test results include a pass/fail result, which is computed by comparing each measurement value with its associated lower/upper limits. Software engineers establish the limits used by the test software, after careful analysis of the circuits that will be tested. Engineers rarely have any feedback on their limit definitions. Experimental results indicated that many of those limits are ill-defined in terms of biases and sensitivity. Statistical analysis tools can support engineers in their limit definitions. In practice it was identified that, very often, the test measurements are within the limits and statistically stable (low variance). However, as shown in the typical example in Figure 1, they are biased towards one of the limits, and present sensitivity much higher or lower than the limit ranges. The TSHMS prototype developed for the F-16 avionic systems provides tools that allow the software engineers to inspect and evaluate the measurements limits.

3.6 Causes of Re-Test OK

Re-test OK (RTOK) and 'cannot-duplicate' (CND) cases occur when systems, sub-systems and units are tested using different ATE or test procedures. A failure is detected at one level but cannot be detected at another level of testing. In the F-16 case there are two testing levels: 1) the flightline testing where sub-systems (referred as line-repairable units, or LRU) are tested as a whole, and 2) the depot testing where sub-systems can be tested at a card level (referred as shop repairable units, or SRU). There are many causes of CND/RTOK. In this paper we address only the case where experimental results have indicated that stack tolerance is the cause of CND/RTOK. In this case the CND/RTOK can be identified and eliminated by redefining the limits and the range parameters of the 'pass/fail' criteria. These parameters are identified using the results provided by the statistical analysis tools.

3.7 Health Status Criteria

Station or instrument statuses are classified based on test failure rates. Statistical coefficients such as averages and sigma (or standard deviation) are used to define thresholds that separate good or bad station items (instruments). In our specific F-16 application, top-level status indicators for stations and instruments use averages and one-sigma parameters to classify 'green', 'yellow' and 'red' situations and are presented in real-time using a web-based user interface. Detailed levels of information and data are presented using hierarchically structured web pages.

3.8 On-Demand Preventive Calibration

Calibration routines are usually performed at fixed intervals of time. However, experience has shown that degradation of performance depends not only on time but also on the duty cycle, i.e., the stress level to which equipment is subjected. An alternative option is to perform calibration procedures based on the status of the instrument. Monitoring performance and anticipating degradation can detect calibrations.

3.9 Aircraft Malfunction Codes (MFL)

Aircrafts have a failure detection system that monitors the avionic equipment's performance during flight operation. Monitoring the communication bus and identifying malfunctions in the equipment communication process perform this function. Each malfunction is identified by a code (the MFL) that is stored and can be retrieved in a post-flight operation. The aircraft maintenance staff at the flightline level identifies possible failures using the MFL codes. In practice, the MFL provides a source of information that is not often trusted by the maintenance staff due to the weak correlation between the MFL code and actual failures. Very often MFL are ignored and are used only to refer the suspected failing system or sub-system to the next level of testing. We are currently developing a prototype to track and link the MFL to all levels of testing and repairing processes. This link will allow the aircraft maintenance staff to associate MFL codes with actual failure and repair actions at the flightline maintenance level, thus avoiding another source of CND/RTOK.

3.10 Test Station Mission Capability (MC)

The tools implemented in the TSHMS can also be used to provide the testing capability of an ATE. Statistical results from the station's instrument self-tests are used to rate the functional capability and quality of the testing equipment. Failure rates and performance status of the instruments are then used to calculate a coefficient, which is used to define the mission capability rate (MCR) of the station.

3.11 User Interface

All TSHMS information can be immediately and securely accessed via an Internet browser. End users can securely access all results and information stored in the database with no additional software applications except the browser. Outputs include textual and graphical (printable) reports.

4. EXPERIMENTAL RESULTS

Actual F-16 avionic systems repair and test data are currently being applied to validate the method. Results shown here were obtained during the prototype development phase. Repair data has been collected since September 1997. Test data has been collected since December 1999. Repair data and test data are not available for all items at all times. There are gaps of data collection. However, there is data enough to provide conclusions to this work. Typical MTBF for this kind of item ranges from 2,000 to 3,000 hours. At a maximum usage rate of 200 hours per month, items are expected to fail and show up to repair every 10 to 15 months. Therefore, only a few items failed more than three times during this period, and had the same failure modes (same components replaced or repaired). The results presented in this paper correspond to items with enough repair and test data to investigate the accuracy and validity of the method. Other results are also presented in Ref. 8.

5. CONCLUSIONS

A test station health monitoring system was successfully implemented using test, repair and failure data, and statistical methods. The methods are capable of monitoring the performance, forecasting failures, and calibration inspection needs of instruments in the test station. The method is capable of identifying stack tolerance problems and detecting some causes of CND/RTOK due to stack tolerances. The statistical and Weibull methods are appropriate but still require better sources of operating time and failure mode information. The system can be implemented in any equipment capable of outputting test results and capable of accessing repair information.

6. REFERENCES

[1] R. B. Abernathy, *The New Weibull Handbook*, Dr. Robert B. Abernathy, author and publisher, 536 Oyster Road, North Palm Beach, Florida 33408-4328, USA, October 1994.

[2] R. B. Abernathy, J.E. Breneman, C.H. Medlin, G.L. Reinman, *Weibull Analysis Handbook*, Pratt and Whitney Aircraft, USA, November 1983.

[3] D. Kececioglu, *Reliability Engineering Handbook*, PTR Prentice Hall, Englewood Cliffs, NJ, USA, 1991.

[4] D. Kececioglu, *Reliability & Life Testing Handbook*, PTR Prentice Hall, Englewood Cliffs, NJ, USA, 1993.

[5] R. A. Dovich, *Reliability Statistics*, ASQ Quality Press, Milwaukee, WI, USA, 1990.

[6] D. K. Lloyd and M. Lipow, *Reliability: Management, Methods, and Mathematics*, American Society for Quality Control, USA, 1984.

[7] "Non-electronic Parts Reliability Data", RAC Reliability Analysis Center, Rome, NY, USA.

[8] K. T. Fitzgibbon *et allii* - *"A Performance Tracking Methodology and Decision Support Model"*, AutoTestCon 2000, Anaheim, CA, USA, September 2000.

7. BIOGRAPHY

Kevin Fitzgibbon *earned his MS in Aeronautical Sciences at Stanford in 1985. He also earned an MS in Electronic Engineering in Control Systems at the Aeronautical Institute of Technology, Brazil, in 1983, and a BS in Electronic Engineering at the Aeronautical Institute of Technology, Brazil, in 1978. As a Senior Systems Engineer for Total Quality Systems, Inc. since 1999, Mr., Fitzgibbon managed several SBIR programs. Mr. Fitzgibbon has also specialized in the integration of aeronautical systems for over 20 years. He participated as a navigation specialist at the International Civil Aviation Organization's (ICAO) Tenth International Air Navigation Conference in Montreal, Canada, 1991. From 1978 to 1989 he worked as an Assistant Researcher for the Aeronautical Technological Center (ATC) in Brazil. As a Research Professor at the ATC from 1989 to 1994, he taught under-graduate and graduate classes in Digital Control, Optimal Control Theory, Avionics, and Navigation Systems. He has been a professional member of The Institute of Navigation (ION) since 1986 and received the Samuel Burka Award from the ION in 1987.*

Larry Kirkland *is a Senior Electronic Engineer at the USAF Ogden Air Logistics Center (OO-ALC), Hill AFB Utah. Mr. Kirkland has over 35 years experience with the United States Air Force in Test/Diagnosis and ATE. He worked as an Electronic Repair Technician on the F-4 & F-16 aircraft. He has worked as an Electronic Engineer in Test/Diagnosis on the F-4, F-16, F-111, & B-1 aircraft. Mr. Kirkland has numerous papers published at Test, Software Technology and AI conferences. He has been published on numerous occasions in IEEE Engineering Magazines. Mr. Kirkland is a past Chairman of the IEEE "AI Human Interface Group" and co-chairman of the IEEE "Interoperability in Knowledge-based systems for sensor-based applications" group.Mr. Kirkland has received the following awards: AFMC Four Star Commanders Coin (2001) – remarkable performance; AFMC Senior Engineer of the year award (2000); IEEE ATE Future Concepts Award (1994); IEEE Walter E. Peterson Award for new technology (1992) - (Machine-based Intelligence); IEEE Using Neural Networks on ATE Award (1990); AFLC Engineer of the year award (1989); OO-ALC Programmer of the year award (1988); OO-ALC Suggestor of the year award (1978). Mr. Kirkland holds an U.S. Patent for* *new development in test/diagnosis techniques. Mr. Kirkland holds a B.S. degree from Weber State University. His graduate work is from Utah State University and the University of Phoenix.*

Bryan Steadman *acquired twenty-five years of management experience while serving as an officer in the US Air Force space program. He was the Program Manager for the Department of Defense NAVSTAR Global Positioning System Joint Program User Equipment Office. In this position, he was directly responsible for the software development and hardware design and acquisition of the GPS receivers procured for virtually every mobile DOD platform (ships, aircraft, ground vehicles, soldiers, etc.). Mr. Steadman earned an MS in Systems Management from the University of Southern California, an MBA from Auburn University, and a BS in Electrical Engineering from the University of Utah. He has been the Sr. Project Manager for developing the Falcon Flex process, and performed the majority of the analyses that has led to a $500 million sustainment cost avoidance over the projected life of the F-16.*

Tony Pombo *earned a BS degree in Physics at the Syracuse University in 1958. He has served as the President of* ***TQS Inc*** *since 1996. As the Design Engineering Program (DEP) II Program Manager from February 1996 to the present date, Mr. Pombo facilitated Acquisition Reform in the Company and Phases I, II and III of the Falcon Flex program. He also has overall responsibility for this SBIR (AF98-262) project. From 1986 to 1996 he adapted several reliability techniques for USAF Weapon System Master Planning. Mr. Pombo has been attracted to the diagnostic sciences, throughout his career, as a means to reduce operation and support cost by identifying process quality and system reliability improvements.*