Using Statistical Process Control To Improve Yield and Traceability for Automated Production Test

S. Hughes, K. Lamond, J. Mackay, Agilent Technologies UK Ltd.

Abstract

Automated testing in Agilent has historically consisted of racks of measuring and test equipment (ATE systems) operated remotely by humans via software programs. In such an environment, certain tasks still require manual intervention such as initiating test programs, making connections to the device under test and zeroing/calibrating power meters.

Responding to wider demands for lower manufacturing overhead costs and increased production capacity, a fully automated production environment had already been designed and successfully implemented by a team in the Sonoma County division of Agilent, allowing 24/7 testing without human intervention. In principle, the migration of different production test processes to this automated environment at Agilent's Queensferry site was straightforward, but a number of unforeseen problems conspired to reduce local operational performance to an unacceptable level.

This paper describes how statistical process control (SPC) was employed to identify and overcome these problems, allowing almost continuous station operation with close to 100% process yield and traceability to national standards.

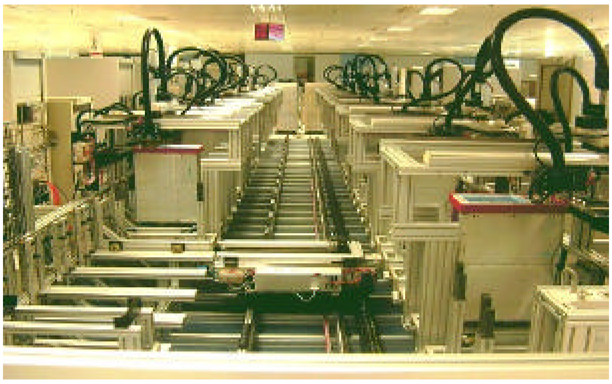
Introduction

Two years ago the equipment manufactured on production lines within Agilent's RF Communications division at Queensferry, Scotland were tested against specifications by processes comprising a number of function-specific systems operated by engineers. Typically, they were responsible for 'walking' a product through each step of the production test process, ensuring that the appropriate test schedule was followed for the product and managing various on-line demands from the test systems. This has been the model of production test procedure for many years and although increased product complexity has driven greater test capability, the implementation has remained largely the same.

Traditional production test environment

With increasing demands being placed on the production department, this model was found inadequate to meet targets for production yield and, therefore, capacity. It was apparent that a number of factors were limiting the production performance, specifically: unrepeatable measurements, erroneous results (no faults found), manual processes and test system calibration frequency. In addition, manufacturing costs associated with capital investment in test hardware and the overhead cost of production personnel were becoming hardware and the overhead cost of production personnel were becoming increasingly unsustainable.

In consequence, a decision was made to change the mode of operation to one of near complete automation where the task of testing a product would be handled by a robotic system. This system, named *Yellowstone*, had been proven elsewhere on other Agilent product lines and was readily available at reasonable cost.



Yellowstone robotic test environment

The main advantages of operating a system such as *Yellowstone* are:

- Maximum utilisation of assets through continuous production increases return on invested capital
- Reduction in manufacturing overhead through fewer production test operators
- Redeployment of skilled test engineering staff to more value added work (instrument rework & repair)
- A manufacturing test process that is more consistent and repeatable
- Easing of production congestion and bottlenecks through defined automation rules
- Greater control of processes leading to a more predictable output.

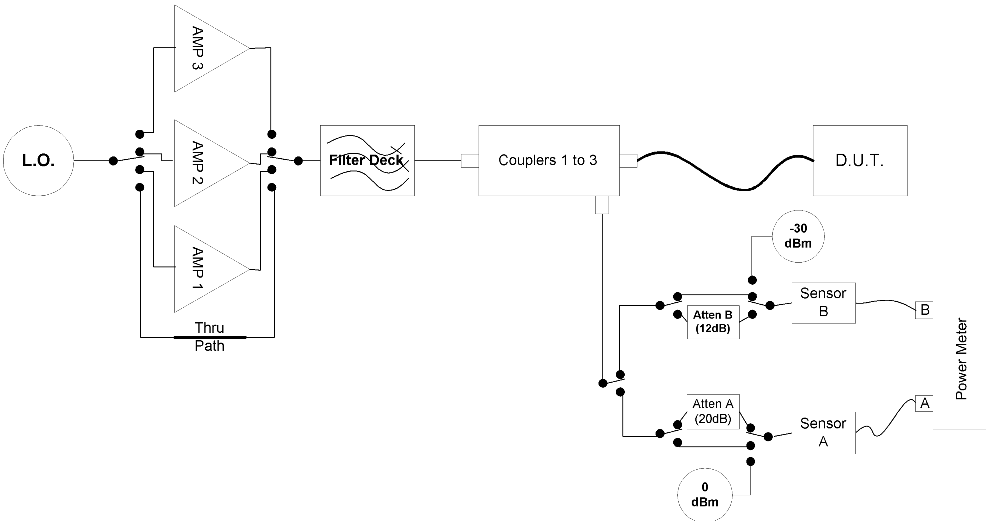
The evolution of one particular production test system, from a legacy-type ATE system to the fully-automated robotic version is now described.

Testing for Continuous Wave (CW) Power Specification

The ET42803 (*Figure 1*) Power Test Station is used to verify the accuracy of CW power detector circuitry within RF receivers that are key products for customers manufacturing wireless communications equipment.

Figure 1

ET42803 Simplified Block Diagram



The system is capable of characterising the power detector against a traceable standard from 100kHz to 4GHz at –30dBm to +40dBm, with typical uncertainties of 0.1dB (95.5% confidence). This performance has been achieved through careful attention to system design, mismatch contributions, harmonic content and the accuracy and drift of power sensors and reference. Algorithms for the calibration of the RF path losses within the test system and monitoring of certain performance parameters are an integral part of its operation.

To Automate or Not?

The proposed automation project coincided with demands for improved RF detector specifications in a flagship wireless communications test product and an investigation was undertaken to assess the feasibility of the proposal. A measurement uncertainty (MU) analysis based on the ISO GUM[1], using the general s-parameter model for a 3-port device[2], was carried-out.

By comparing the weighted contributions of the various terms of this expression, potential problem areas were identified with the new system configuration. These anticipated problems in changing the environment for the test system might be summarised as follows:

i. The increase in length (1m to 2m) of the primary RF cable connecting the device-under-test (DUT) to test system would increase the uncertainty due to mismatch (VSWR), and repeatability (flexure) without careful selection.
ii. Push-fit N-type connections between the test system and DUT, power sensors and power references might show poorer repeatability than torqued connectors.
iii. Stand-alone reference sources for 1mW and 1μW would be required since the manual calibration and zeroing of power sensors was no longer possible.

Solutions to these problems were identified in the form of:

i. Armoured cabling of high specification.
ii. Specially adapted metrology-grade connectors mounted on a mobile ‘vehicle’ with force control.
iii. The availability of new reference sources intended for standalone use.

Based on these solutions, the decision to automate was agreed.

The Outcome

Whilst none of the anticipated problems caused any significant effects in the test system configured for the automated environment, it was apparent that the performance of this system was significantly less than those systems still operating in the manual environment. The cause(s) of the degraded performance was not obvious. The intention of a more robust and repeatable measurement had in this particular instance, not been realised and the station had to be withdrawn from the Yellowstone environment.

In order to determine the causes of the poor performance of this system without extensive experimentation involving a team of engineers, some form of reference measurement was essential to compare the station’s performance with known good values.

Introduction of SPC

The use of statistical process control (SPC) had been employed on production lines at Queensferry for some years but in nearly every case the purpose had been to indicate trends or relative changes in the test system performance over a period of time. A product, representative of those being

---

[1] *Guide to the Expression of Uncertainty in Measurement* 1993 (E), International Organisation for Standardisation.

[2] *Understanding Microwave Power Splitters*, R.A. Johnson of Hewlett-Packard Company, Microwave Journal 1975.

manufactured, was generally used as the nominal standard for this testing. However, this approach was insufficiently accurate or repeatable to identify the cause of the Power Test Station problems. To facilitate the investigation, a more rigorous testing scheme employing a more fundamental standard was required.

It was then realised that SPC, which would help engineers to identify and then monitor/control test system performance issues, had wider potential benefits. In addition, a means of comparing a number of similar test systems against a single reference was highly desirable. A reference calibrated at a standards lab might further provide direct traceability from production measurements to national standards.

There are a number of possible gains built on the foundation of this form of SPC:

- Calibration of test system RF paths driven by SPC rather than maintenance schedule.
- Test system downtime reduced due to increased calibration interval.
- Maximum test system yield (target > 99%).
- Clear identification and segregation of product and test process problems.
- Confidence in measurements is maintained (ISO17025 requirement).

SPC Building Blocks

The elements necessary to realise the foundation of SPC include:

Measurement System Understanding

The key aspect that SPC is required to verify is the test system measurement uncertainty. This requires the engineer to have a comprehensive understanding of system operation down to the smallest contributor of uncertainty such as switch and connector repeatability, power sensor drift and so on. The investment in time by the engineer is not insignificant.

The Reference or 'Gold Standard' Instrument

The proper selection of the Gold Standard instrument is fundamental to SPC implementation. The repeatability of the instrument must be significantly less than the test system uncertainty for the SPC process to add value. For this reason, the traditional approach of selecting a representative product is inappropriate in most cases. The criteria that must be considered when selecting the Gold Standard include:

- It must operate over the parameters and range of the test system.
- Must have a 2-sigma repeatability that is less than the system MU.
- Calibration uncertainty that is appropriately less than the system MU.
- To maintain confidence, its assigned calibration interval should be shorter than recommended by the manufacturer.
- Cost (two units enable process continuity, for example if one is being calibrated).

Operational Requirements of the Gold Standard

- The unit should be continually powered, even whilst not in use.
- Frequent maintenance should be performed e.g. connectors gauged and cleaned, fan filters cleaned, etc.
- It must not be opened, adjusted or used for diagnostics.
- Handled, stored and transported in a manner that will not affect the calibration or physical condition of the instrument.

The SPC Limits

For warranted DUT specifications, the SPC limits should be no greater than the MU value.

Type-B MU analysis will yield 95.5% confidence limits that are generally more conservative, but apply to a number of test systems of the same type throughout the recommended calibration interval. Type-A analysis on the other hand may only be valid for a single system with defined system trace equipment, Gold Standard and environment. Variation from system to system needs also to be accounted-for.

Multiple Gold Standards will introduce further variation and it may be necessary to add a term to the SPC limit values to account for this.

SPC Test Times and Frequency

Adding SPC to an existing test process will impact production time. The cost of running and maintaining an SPC methodology must be balanced against the risk to quality through incorrectly calibrated end-product. It is therefore important to select the smallest number of test points that will fully exercise the system through its entire operating range.

A similar argument applies to the frequency of Gold Standard runs and consideration should be given to the following:

- Production volume fluctuations.
- Running daily for guaranteed DUT specifications.
- Bi-daily or weekly intervals for non-critical specifications.

SPC Failures

An SPC failure can have many possible causes; operator error, extraneous signals, contaminated connectors, cable wear, etc. The procedure for flagging a failure must take into account all these mechanisms and drive a corrective action process. In our case, an SPC failure has been defined as two consecutive runs that do not pass all test points. In the automated environment, the Yellowstone controller immediately puts the test system off-line, preventing production throughput.

Reporting SPC Results

Trends in performance and SPC failures are clearly seen by presenting results graphically. This aids the engineer in detecting anomalies in system performance, the prediction of system drift or when a test system requires calibration. Reviewing large data sets on a frequent basis is, however, time consuming and prone to error and so test support personnel are notified automatically by the system when an SPC failure occurs. Such a reporting system balances the need for prompt remedial action with the collection of data for analysis.

SPC Implementation

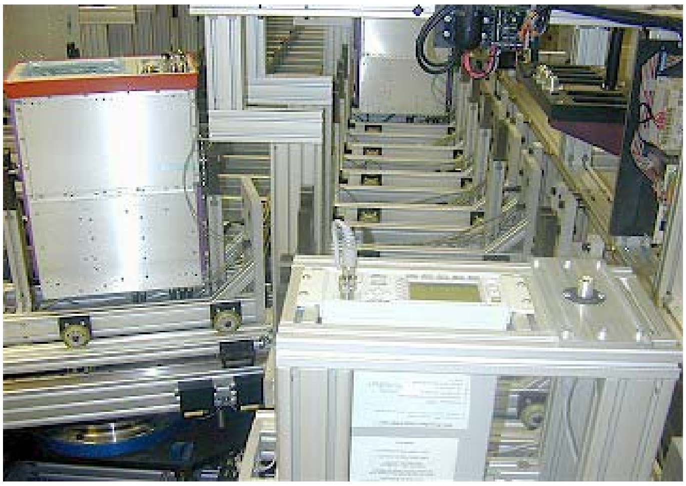
The implementation of a robust SPC process for the purposes of guaranteeing station performance, requires the measurement of key test points with a repeatable working standard. This, in effect, is an accurate calibration of the test system itself. This provides an error value at the key operating conditions that, in turn, gives an indication of the system ‘health’. The instrument selected as the Gold Standard was an Agilent E4419B power meter and 8482A power sensor. Because of the requirements for automated connections, the sex of the power sensor was changed from N-type male, to female creating unusual calibration requirements.

Table 1 – SPC Gold Standard Calibration Requirements for Yellowstone Station

Equipment Model No.	Equipment Description	Test Points	Required Uncertainty	Test Proc. or Cal Lab.
ET54001 E4419A/B N-type(m)	Power Meter	As manual except for 1mW Ref. Output. Use 478A N-type (f) for Stds Lab calibration.	< 0.4%	ET54001-90001
ET30896 8482A N-type(f)	Power Sensor 100kHz – 4.2GHz	Cal. Factor & VSWR 0.1, 0.3, 0.5, 1, 3, 5, 10, 30, 50, 100, 200, 300, 500, 650, 800, 1000, 1500, 1800, 2000, 2500, 2600, 3000, 3500, 3700, 4000, 4200	Cal Factor < 0.7% Input VSWR as manual	Nat stds lab ET30896-90001
ET30897 5065- 4616	1mW Reference Oscillator	Output power level	< 0.4%	ET30897-90001

Both the sensor and power meter are mounted on a permanently powered cart (UPS cart), which never leaves the Yellowstone system. In order to facilitate the automatic calibration of the 8482A power sensor by the Yellowstone robot, use of the power meter's 1mW reference was substituted by a standalone reference located at one of the Power test stations. Limits of ±0.015dB were applied to the automated sensor calibration in order that any problems with the UPS cart may be detected.

Reference 1mW source on a mobile device



Gold Standard E4419B and 8482A in UPS cart

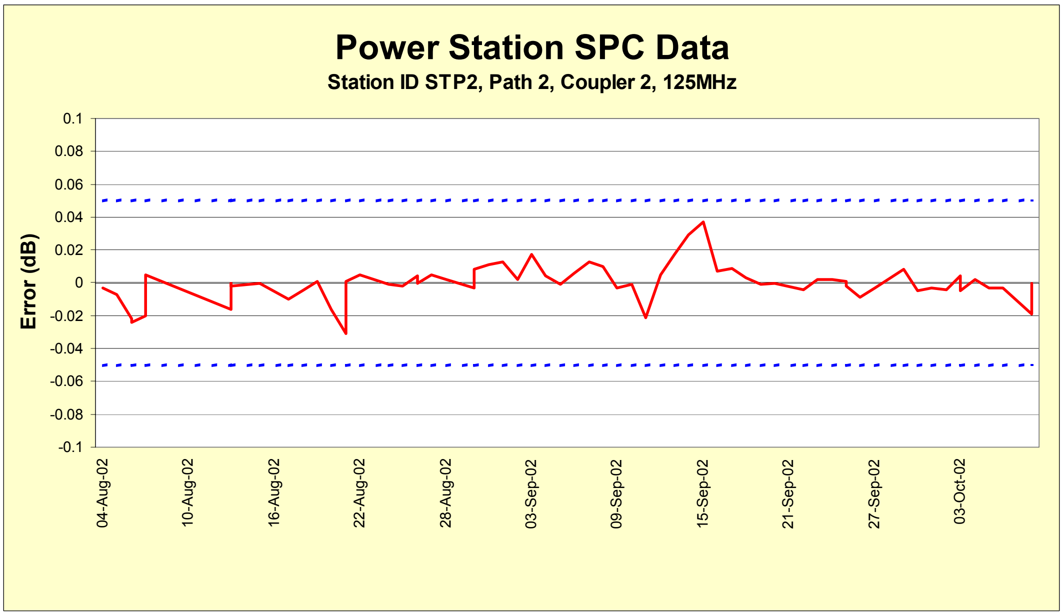
The test points covered by the SPC program were selected to ensure that the majority of the operating conditions and critical RF paths were exercised. The test plan covers the ranges of paths by amplitude, couplers and filters.

Table 2 – SPC Test Plan

ET42803 and ET42932 Station Path Matrix								
STATION PATHS			PATHS USED TESTING PRODUCT					SPC
Coupler	**Amplifier**	**Filter No.**	**Filter Freq (MHz)**	**Product A Pre-Test**	**Product A Final Test**	**Product B Pre-Test**	**Product B Final Test**	**SPC Test Freq (MHz)**
1	1	1	0.4					
1	1	2	1					
1	1	3	10			used	used	10
1	1	4	15					
1	1	5	20					
1	1	6	30			used	used	**Note 1**
1	1	7	50			used	used	40
1	1	8	80			used	used	60
2	1	9	100			used	used	100
2	1	10	150	used	used			125
2	1	11	200	used	used			175
2	1	12	300	used	used			250
2	1	13	500	used	used			400
2	1	14	800	used	used			650
2	1	15	1000	used	used			950
2	2	16	1500	used	used			1250
2	2	17	2000	used	used			1850
3	3	18	3000	used	used			2720
3	3	19	4000					
Note 1 : No calibration point exists to allow SPC on this path.								

The initial thought for the SPC limits was to use the MU values of ±0.1dB. However, because the process effectively repeats the measurement of the RF path loss carried-out during a test system calibration, analysis of the original measurement equation indicated that a number of the uncertainty contributors cancel out. The revised analysis indicated that limits of ±0.05dB were appropriate.

Figure 2 – Example control chart



Results of Implementing SPC

As previously discussed, a variety of implementation problems over and above those anticipated had seriously impacted the performance of the fully automated Power test system. Immediately after the introduction of the described methodology, an SPC failure was found to occur almost daily which confirmed, as previously suspected, that the Power test system configured for the Yellowstone environment was out of control. Investigation revealed several causes:

i. The primary system RF cable is in this instance, attached to the DUT or Gold Standard, using a metal plate housing a metrology-grade push-fit N-type connector (mobile device). The robot uses an electromagnet to pick-up and position the plate and then maintain the connection during measurement. It was found that heat transfer was taking place from the plate, through a Delron connector mount, the N-type connector itself and finally, to the 8482A's internal thermocouple sensor. When calibrating the system, DUT or Gold Standard this caused a temperature related measurement drift, inducing an offset in the results. To overcome this, 12V fans were fitted to the electromagnets.

ii. The connection of two additional power sensors (sensors A and B) to N-type bulkhead connectors mounted on the test system had been replaced with mobile devices, held in position during test by passive magnets. The 8lb force of connection with a spring-loaded N-type (f) was insufficient to ensure good repeatability. This arrangement was replaced with the two power sensors permanently located in the test system. With switched paths for measurement, sensor zeroing (50Ω termination) or sensor calibration (reference source), there is no disconnection of the power sensor and non-repeatability error is minimised to that of only the matrix RF switch.

iii. Test operators initially performed a manual calibration of the Gold Standard 8482A power sensor to the 1mW reference. This was subsequently replaced with a robot-assisted calibration, found to be considerably more repeatable and consistent.

iv. High power RF amplifiers used in the Power test system were found to be generating DC offsets when not in circuit. This became more apparent when an existing intermittent problem of sensor damage was not eliminated within the robot controlled Yellowstone environment. Additional terminated switch paths were added to prevent destructive discharges.

v. With improving consistency and performance of the test system configured for automation, the effects of RF switch non-repeatability, connector quality and maintenance/cleanliness issues became more apparent. Component changes, previously deemed insignificant, were carried-out and the maintenance schedule modified.

As problems were identified, the fixes were rolled-out across the large number of Power Test systems, both within and outside of Yellowstone. The performance of the systems, monitored through daily SPC runs, increased to a hitherto unrealisable level. The specific improvements being:

i. Yield increase from around 50% to 90% for the product tested at the Power Test System.

ii. A reduction in the number of test systems from 11 to 7; an effective saving of 30% in capital expenditure.

iii. The MU has been reduced by around 20% to ±0.072dB (±0.082dB below –27dBm).

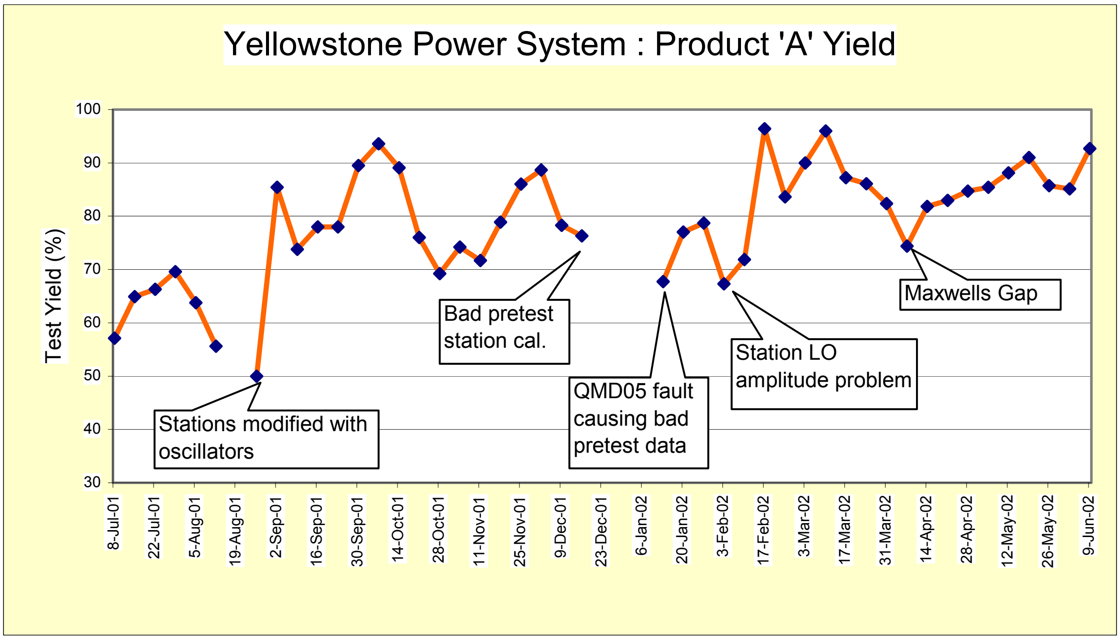
iv. Manufacturing tolerance interval reduced by around 20% (i.e. tightened test limits), in line with demands for similar reductions in the customer specification on a key product.

v. The systems' calibration intervals are now driven by SPC as drift in accuracy is monitored through the daily SPC runs. The interval is currently 1 month with no SPC failures. Limited experimentation indicates that an interval of 8 weeks and possibly more may be achievable.

vi. By being able to effectively separate the performance of the test system and product (DUT), the number of 'no fault found' conditions has decreased dramatically. Passing the exacting demands of the daily SPC run means that any DUT test failure now confidently be attributed to the product rather than the test system.

vii. Manufacturing and production resources are more effectively utilised.

Figure 3 – It's important to correlate problems and their resolution with system performance



Stations modified with oscillators

Bad pretest station cal.

QMD05 fault causing bad pretest data

Station LO amplitude problem

Maxwells Gap

The SPC philosophy described here has realised the intention of helping to identify the problems associated with automating the Power Test System and subsequently, has provided a means of keeping the test process in control with a high degree of confidence.

This process can be applied to any test system to monitor the performance of over a period of time, relative to a known good operating condition. It is the relative 'drift' in the measurements and the growth in uncertainty from that point that is under examination and, if the measurement exceeds the SPC limits, then there is either an immediate and (hopefully) identifiable problem or possibly recalibration of the system is required.

Looking forward, if the reported results of the SPC test itself, rather than the calibration of the Gold Standard power meter and sensor, could be traced directly to national standards, then it is unlikely that confidence in the production measurements can be further improved upon. In effect, the SPC run is calibrating the test system against a higher standard at frequent intervals, through the Gold Standard (transfer device). This method could also be considered to be an inter-laboratory comparison (ILC).

Inter-Laboratory Comparison

The potential implications of performing a frequent ILC where one body is, for example a national standards laboratory such as NPL or NIST, and the other is Agilent production are:

- Extended calibration intervals of test system equipment due to increased confidence in measurement through SPC and ILC.
- Direct traceability of measurement through the Gold Standard, rather than multiple items of test equipment in the system, implies that test systems may not need to leave the production line for scheduled equipment calibration.

ILC Specifics

An ILC process is a comparison of calibrations undertaken by two or more participating calibration entities. In this case the two calibration bodies might be Agilent Queensferry production (Power Test System), a national standards laboratory and utilising the artefact itself, the Gold Standard unit. Previously it has been stated that one of the criteria for selecting a Gold Standard was repeatability. This is critical to achieving the most accurate comparison.

Specified parameters, test conditions, instrument settings and associated uncertainties first need to be agreed; essential as the standards laboratory will almost certainly use different methods and equipment in order to perform the measurements. The ability of the laboratory to make measurements at all of the required test points may not in some instances be possible. An ILC is needed whenever the Gold Standard is calibrated as this provides an absolute reference with least uncertainty at that point in time.

Comparison of Results

With the ILC measurements completed by both parties, a comparison of the data generated by the Power Test System and the standards laboratory may be made. Calculating the difference between these reported values will not in itself produce useful information, unless accompanied by some form of acceptance limits.

$$\text{ILC Result} = \text{Results}_{\text{Power Test System}} - \text{Results}_{\text{Standards Lab}}$$

These limits will be calculated by combining, in some manner, the expanded uncertainties from both parties. In doing this, the performance of both measurement systems is accounted for in a single term. A possible method for defining the acceptance limit involves the quadratic sum of the two uncertainties for the measured parameter, together with the Gold Standard's repeatability:

$$\text{Acceptance Limits} = Factor \times \sqrt{MU_1^{\,2} + MU_2^{\,2} + \text{Repeatability}_{\text{Gold Standard}}^{\,2}}$$

where MU_1 and MU_2 are the expanded uncertainties for each party and *Factor* depends on the confidence required.

The value of *Factor* must be small enough (< 1.0) to reflect the fact that the acceptance limits apply to the difference of the two measurements. Additionally, the limits must assure the required confidence in measurement traceability of the Power Test System.

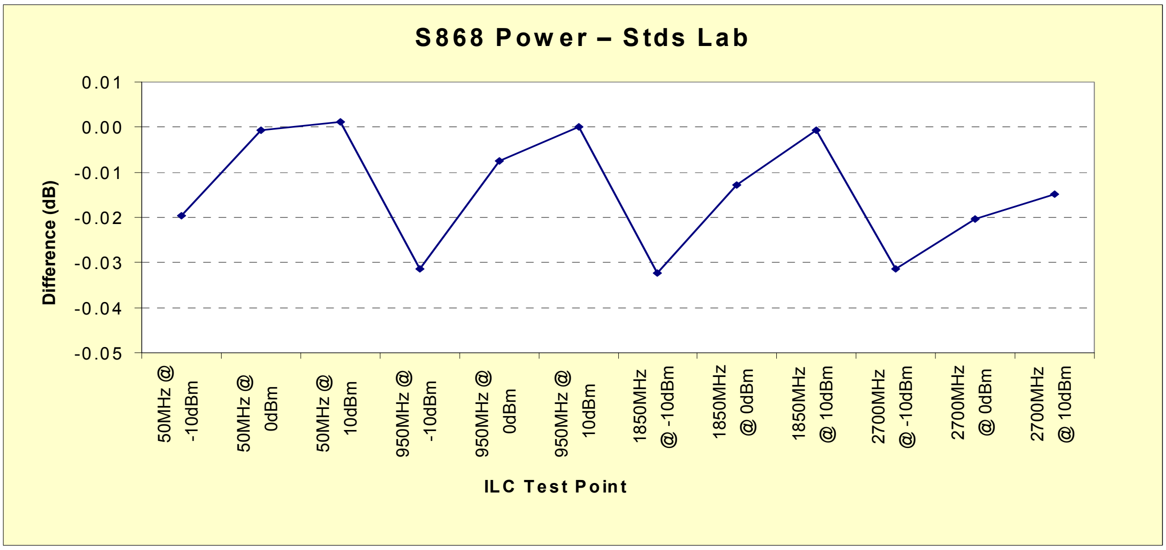
If the ILC results generated by the multiple Power Test Systems fall within the acceptance limits, we can be confident that each is making absolute measurements correctly. However, if any system produces results outside of the acceptance limits, then there is a problem with the measurement and the test system will be put out of commission until corrective action has been carried-out. The results in *Table 3* show the ILC data for the first Power Test System.

Table 3 – Initial ILC Result for Power Test System

Freq (GHz)	Indicated Power (mW)	Standards Lab		SQF (Manual) Power System S868		Difference S868 Power – Stds Lab (%)	Possible Acceptance Limits (%)
		Power Diff (Inc–Ind) (%)	MU (%)	Power Diff (Inc–Ind) (%)	MU (%)		
0.05	0.1	0.3	0.5	-0.15	1.7	-0.45	1.2
	1.0	0.2	0.4	0.18	1.7	-0.02	1.2
	10.0	0.0	0.4	0.03	1.7	0.03	1.2
0.95	0.1	1.4	0.6	0.67	1.7	-0.72	1.2
	1.0	1.2	0.5	1.03	1.7	-0.17	1.2
	10.0	1.0	0.5	1.00	1.7	0.00	1.2
1.85	0.1	2.4	0.7	1.64	1.7	-0.74	1.2
	1.0	2.3	0.6	2.00	1.7	-0.29	1.2
	10.0	1.9	0.6	1.88	1.7	-0.01	1.2
2.70	0.1	5.3	0.7	4.54	1.7	-0.72	1.2
	1.0	5.2	0.6	4.71	1.7	-0.47	1.2
	10.0	5.0	0.6	4.64	1.7	-0.34	1.2

The ILC results indicated by the Difference term may be seen more clearly in the chart of *Figure 4*.

Figure 4



The chart shows that the measurements made by the S868 Power Test System are within 0.74% (0.032dB) of those carried out by the standards lab. The chart also shows that it may be possible to make further improvements to system performance by investigating what may be a systematic effect, present at levels less than 10mW.

The ILC has enabled acceptance limits to be placed around the absolute performance of the Power Test System. When monitored and controlled with the regime of SPC testing, the performance of the system can be traced directly to the standards lab on an on-going basis with a high degree of confidence.

Conclusion

The concept of applying a more metrologically based SPC process has provided a means of identifying and then controlling system performance issues within well-defined and exacting limits. This method met the original aim of solving the problems associated with the migration of a legacy test system, critical to RF Communications production, to a new automated environment. Benefits to production have been realised in the form of reduced capital equipment requirement, improved system performance (quality and yield) and finally, the redeployment of production resource. This work has had a positive impact on the factory-cost of Agilent products, and provides a clear justification for extending this philosophy to other production process within the factory.

By introducing the inter-lab comparison, measurement traceability is more direct and has been enhanced over the traditional reliance upon an unbroken chain of calibrations. The possible implications of this will be fully realised in the future.

Acknowledgement

The authors would like the acknowledge the time and effort given by the engineers and metrologists of manufacturing engineering, PL13 production and the Queensferry standards lab that have allowed this work to be performed. In addition, the authors would also like to recognise the investigation and detailed MU analysis carried out by Dave Leestma of Spokane Test Engineering.