The test pyramid: a framework for consistent evaluation of RFID tags from design and manufacture to end use

David J. Puleston and Ian J. Forster
Avery Dennison RFID Division
(rfid.info@averydennison.com)

ABSTRACT

The document presents a testing framework for RFID tags in the conceptual form of a pyramid, with high precision primary response test systems using laboratory equipment at the top of the pyramid, and with other systems based on standard readers at lower levels. We explain how Avery Dennison uses a secondary test system known as a test cube to ensure consistency of measurement across functions and locations, in effect providing a pedigree for tags from design to delivery (and viably even to the point of end use). Additionally, we outline a simplified approach to the process known as “sweet spot” identification.

Introduction

Avery Dennison is the leading supplier of pressure sensitive adhesives and materials to the label conversion industry and has expanded this offering with high volume availability of RFID inlays. Inlays are manufactured in multiple lanes at very high speeds using a proprietary manufacturing process that has increased throughputs by at least an order of magnitude.

Additionally, we are an end user of RFID technology, and have implemented an integrated system within our office products division in response to retail mandates.

Our applications, research and development team has over one hundred cumulative years of experience in UHF RFID deployments from both system and tag perspectives. The team is focused upon providing customer support and improving our tag designs and manufacturing processes. We do not offer for-revenue test or integration services.

As part of our development efforts we maintain a fully equipped applications laboratory and perform extensive testing to ensure that our tag designs are responsive to market needs. In the course of this, we have performed extensive work on improving our test methodologies, both from the perspective of evaluating tag performance on products, and also to ensure that tags being delivered to end users have the intended design performance.

We have implemented a test framework that gives us end to end traceability of tag performance internally, and ensures that calibration and testing is maintained across equipments and sites in a relatively simple manner. An extension of this concept would logically allow an unbroken tag “pedigree” that extends beyond our own company and into our converter customers and into the business end users of the technology.

Several end users (as well as test and integration partners) have suggested that we disclose many of our testing processes to the EPC community and this paper has been prepared as a result of this. The information is released for the general good and we have no vested self interest in promoting one testing approach versus another, except to the extent that it drives adoption of RFID technology.

Test pyramid

Our test framework is implemented in a test pyramid concept that provides an optimal trade-off between the accuracy, cost, speed and flexibility needed throughout the life cycle, including design, manufacture and end use application.

By maintaining test consistency via calibration routines, test RFID devices and procedures, we ensure that we have repeatability and that we measure tag performance in a manner that can be traced directly back to national standards. The test pyramid is illustrated in figure 1.

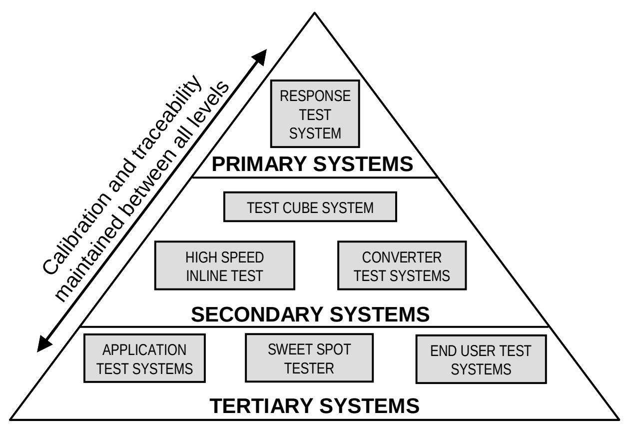

Figure 1: **The test pyramid facilitates consistent tests**

Primary Test Systems

The primary level systems are designed to give good accuracy and repeatability and to act as the reference point for all other test systems. The primary systems incorporate precision equipment with precisely calibrated functionality. However, the systems are expensive and require very highly skilled staff to operate.

Avery Dennison has primary test systems located near Chelmsford in the United Kingdom, and near Atlanta in the United States.

A block diagram of the system is shown in figure 2.

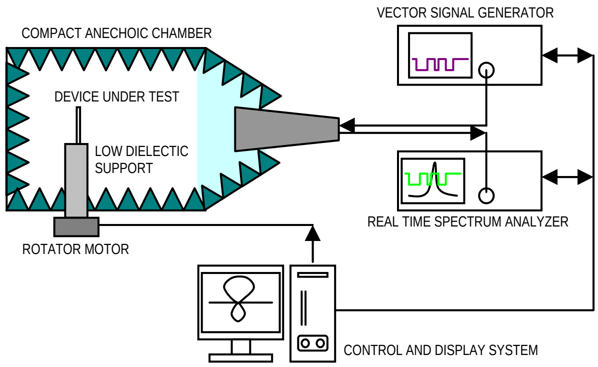

Figure 2: **Primary response test system block diagram**

The primary response test system consists of a calibrated Rohde and Schwartz SMJ100A vector signal generator capable of producing an EPC or ISO standard modulated command message over a wide range of frequencies at an accurately defined power level.

This signal drives a calibrated horn antenna mounted integrated as part of a small anechoic chamber with the device under test (for example, a tag or a tag on a product) mounted at a defined distance, so that the power level delivered to that point is, again, known to a high accuracy. Responses from the tag are received by a Tektronix RSA3408A real time spectrum analyzer. Figure 3 is an example using an EPC C1G2 waveform.

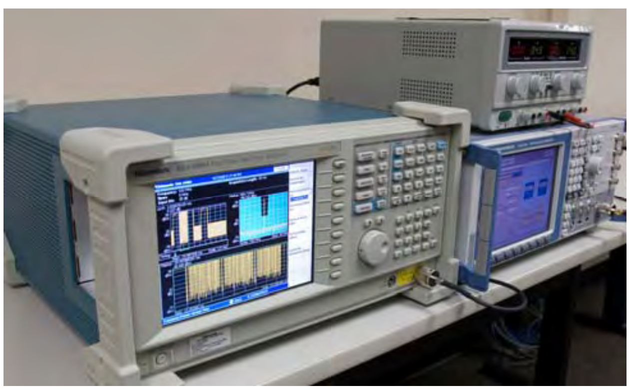

Figure 3: **Primary system detecting gen two response**

In operation, the power level of the transmitted command signal at the test frequency is increased until the tag reaches its operating threshold and responds by sending a backscatter modulated return signal. From experience, in a suppressed reflection environment and at a single frequency, this threshold often tends to be very abrupt, but where this is not the case, a statistical measure such as ratio of responses to command messages at a given power can be used.

For current RFID chip designs, the limiting factor for tag range is the delivery of power to the chip; with a well designed reader system the level of backscattered signal is unimportant. Therefore, it is only necessary for the real time spectrum analyser to detect a response. It does not need to decode the data sequence returned, although the level of backscattered signal can be recorded.

As the system allows for measurements to be taken across a wide range of frequencies, it may be used at any of the international bands, and can also allow the tag designer to study the detuning effects of different products on tags.

Secondary Test Systems

The primary test system is intended to give very accurate and repeatable results, but is expensive, slow to use and necessitates skilled operators. For more general work, we need more flexible secondary test systems that are comparatively less expensive than a primary system and can accept a slightly lower level of accuracy provided that they are held in a calibrated state that can be traced back to national standards and allow for repeatable results.

Secondary test systems can be reader based, enabling a more complete evaluation of tag characteristics, such as write power thresholds. This makes the systems more suitable for evaluating tag performance on products, and also less expensive, allowing multiple systems to be used.

We envisage two main types of secondary system: inline test systems used in manufacturing (outside the scope of this paper) and a standardized device that we call a “test cube”, which is discussed in detail below.

Test Cube

The purpose of the test cube is to ensure that all of our sites and equipments have common secondary test methodologies and reference levels. Each of the test cubes (figure 4) incorporates a standard reader, antenna, control PC, and optional external attenuation. It is configured such that precise measurements may be made of a tag at a defined distance from an antenna (normally one meter).

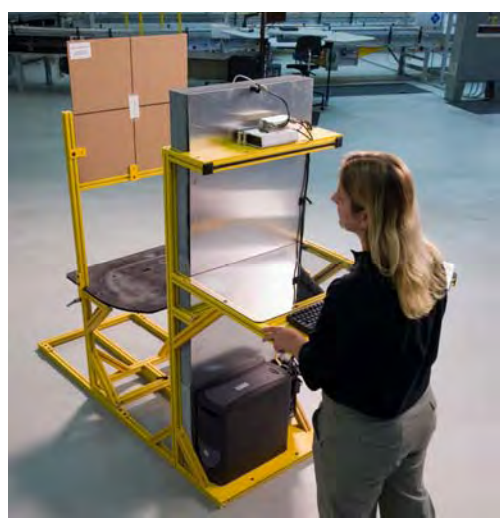

Figure 4: **Standardized test cube in use**

All our sites have at least one test cube system, and maintain a consistent calibration via calibration tags generated using the primary test systems. Using these systems and the inline manufacturing test systems, we can provide tags with a through life pedigree that ensures that tags leaving our manufacturing facility are meet the performance standards that our design and applications teams envisaged. It is a logical extension of this scheme to extend the pedigree such that we are able to offer the same level of tag certainty to our converter customers, to equipment manufacturers and even to end users.

We use the ThingMagic Mercury 4e reader within the test cube for a number of reasons that include flexibility, stability and multi-protocol operation. It is also readily available, which is essential given the large quantities we require, particularly in our manufacturing operations. However, the solution outlined here is not restricted to that reader and could have been constructed using other readers.

It is generally accepted that read range is one indicator of tag performance, but can be difficult to measure exactly because of multi-path effects such as reflections. Also, range measurements by their nature involve moving part of the system and this is therefore a slow process in comparison to a purely electronic approach. Therefore many users have opted to measure the link margin of the system, namely the maximum attenuation that can be added to the communication channel and the tag still respond. Using such a technique it is possible to reconstruct the theoretical range as follows:

$$R_{theoretical} = R_{measurement} \times 10^{-margin/20}$$

For example, 12dB link margin at 1m would give a theoretical range of 4m.

We use a similar technique but have added calibration mechanisms that allow us to reference readings back to an absolute scale sensitivity measured in dBm, effectively making the reading independent of range. Figure 5 shows an example of such a sensitivity sweep as measured on the test cube using our windows based control software.

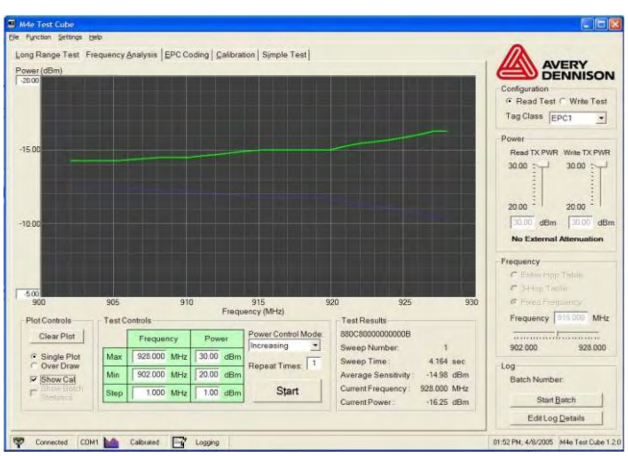

Figure 5: **Test cube link margin plot**

In reality, real world systems will hop in frequency, and in the US the tag will be read at a random frequency between approximately 902 and 928MHz. We therefore measure the link margin over the full spread of frequencies for a number of reasons:

- Tags (particularly smaller tags or tags placed near metal) can become narrow band, and it is appropriate to check that the tag can be read at all frequencies the reader might use;
- Taking measurements at a number of frequencies gives us a greater statistical population and improves the overall accuracy of measurement;
- The slope of the response allows us to make inferences regarding tag detuning by carton contents and possibly refer the test back to the primary test system for more detailed analysis.

As in the case of link margin, it is possible to reconstruct a theoretical range from the average sensitivity resulting from the frequency sensitivity sweep. For example, at 915MHz when using a 36dBm EIRP:

$$R_{theoretical} = 10^{(4.33-sensitivity)/20}$$

In the past we have attempted to make probabilistic measurements of read rate, but found inconsistencies between different types of readers and also between readers of the same type.

For example, it is known that many class one tags are not read consistently by what is probably the most commonly used class one reader. An example is shown in figure 6, which plots the read probability of such a tag (with red representing 100% success, blue 0% and other colors some probability between these values) versus power and frequency. The results are not what would be expected.

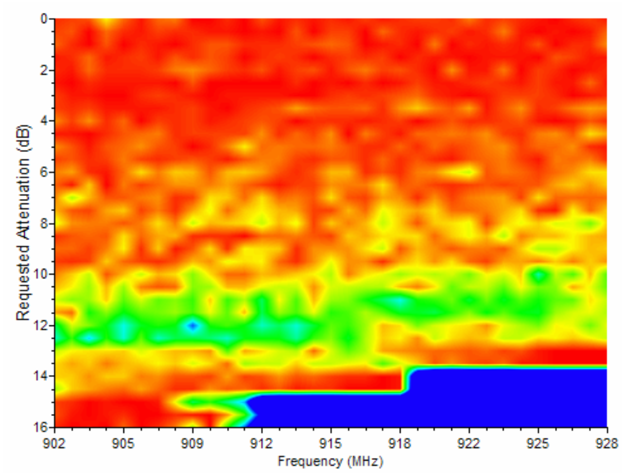

Figure 6: **Inconsistent read probabilities of a reader at various powers and frequencies.**

We have seen similar issues with other readers and classes. We often see major changes using exactly the same tag and reader, with only the firmware changing – leading us to question whether we were in reality measuring the tag or the reader. We therefore do not favor using probability measurements in our own testing given the current rate of change in current reader systems.

As reader based systems do not have the same precise calibrations as laboratory grade equipment, the key challenge is to keep measurements consistent across different test systems. In the past, we distributed specially selected reference tags so that users could determine the offsets of the various systems attributable to variations in reader power output, cabling or antenna gain.

More recently, in implementing the test cube, we used a different technique designed to facilitate auto-calibration and to remove the need to use specially performing tags. Calibration tags are enclosed in hardened plastic cases for protection and are written not with an EPC code but with the response that would be expected from that tag in a perfect system, namely a 30dBm power output into a 6dBi antenna with zero cable and insertion losses.

In this manner, the control software need only read the reference tag to determine the ideal values, compare these with the actual values measured of the reference tag. The difference is the offset needed to maintain calibration across systems. This correction factor is automatically applied by the control software. In our implementation, this data is held at eleven different frequencies, using the encoding scheme shown in table 1.

Bits	Data	Comment
95 (msb) - 88	0x3A	Header identifying calibration tag
87 - 76	Date	
75 - 61	f_{res}	Resonant frequency of tag
60 - 55	Base	0-31dB
54 – 50	903.0	Offset values for eleven different frequencies containing response as fractional dB. Actual = Base + Offset
49 – 45	905.4	
44 – 40	907.8	
39 – 35	910.2	
34 – 30	912.6	
29 – 25	915.0	
24 – 20	917.4	
19 – 15	919.8	
14 – 10	922.2	
9 – 5	924.6	
4 – 0 (lsb)	927.0	

Table 1: **Calibration tag data format**

Our test processes require that test cubes be recalibrated at least weekly. The control software will automatically prompt the user to recalibrate by inserting the calibration tag and pressing the calibrate button (figure 7).

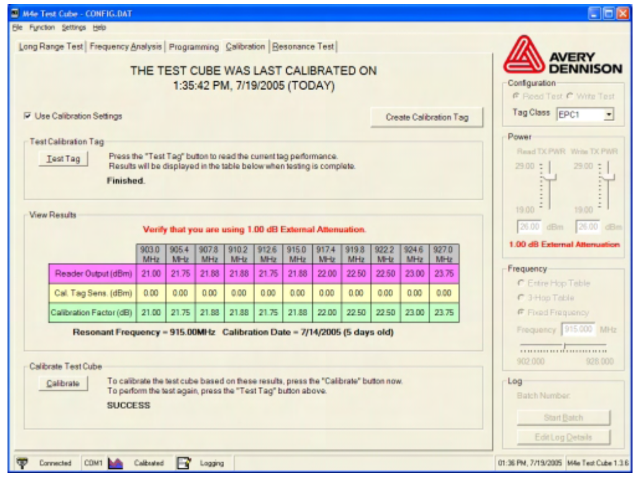

Figure 7: **Test cube calibration procedure**

We can also store the resonant frequency of the tag, namely the point of maximum power absorption or 0º phase transition when close coupled to an antenna and network analyzer system. To do this, we have ported the test cube software to run on an Agilent N5230A network analyzer as shown in figure 8.

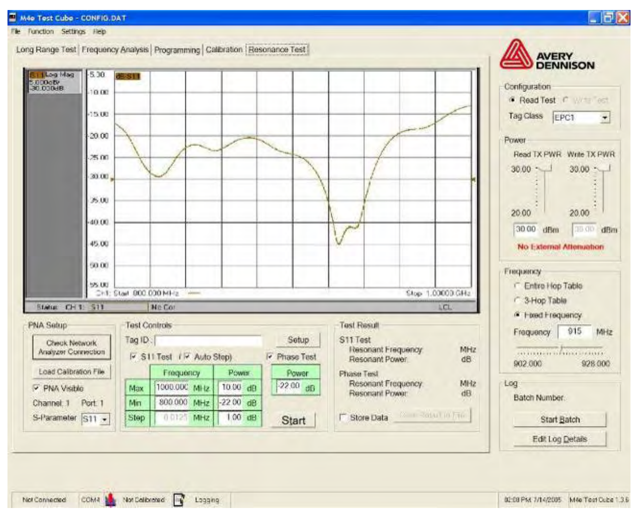

Figure 8: **Measuring resonance frequency by running test cube control software on network analyzer**

Resonant frequency measurement is a specialist test primarily used in the tag design process and should not be confused with the frequency response of the tag in an application. Both the resonant test and similar tests using multi-tone approaches are measuring only one element of the tag performance, such as the response of the tag rectifier structure, and ignore the significant impact on tag performance of circuitry like charge pumps and also operational shifts in tag characteristics as power changes.

It is not required that test cubes be operated in an anechoic chamber, and it is very possible to get consistent results by spacing the cubes at a distance to each other in an environment which does not have a low height ceiling. In our applications lab, which has plenty of space available around each system, multiple test cubes track each other within ½dB when reading the same tag.

If however, sufficient space is not available, then it is recommended to provide screening with anechoic foam behind and to the sides of the cube. At one of our other sites with less space available, we have observed that we are able to influence test results measurably on an unscreened system by opening and closing metal blinds some tens of feet from the test area. We would also caution users to avoid low ceilings and fluorescent lighting where possible when trying to make precise measurements.

A further implementation note is that the limiting factor in tag range performance (assuming a well designed reader) is the link to the tag i.e. if you can power a tag you should expect to be able to read it. The implication of this is that up to a certain point, the same result will occur whether or not the receive path is attenuated. Thus, it is possible to postulate a minimal system that relies solely on varying the reader power output and does not require additional external attenuation.

A final word of caution here is that the performance of a tag really needs to be tested at multiple angles against a variety of dielectric and metallic substrates with varying three dimensional structures and air gaps. Over emphasis on face-on free space link margins can be interesting from a competitive perspective but potentially does a disservice by placing design emphasis in the wrong places. If one relies solely on free space tests there is a real possibility that “better is the enemy of good enough”. It is far more important that the tag work well in the application i.e. when applied to products that will be read in motion at uncontrolled angles.

Tertiary test systems

Tertiary test systems are additional test systems that are used to facilitate application testing, either in terms of typical end use applications such as conveyers and portals, or in terms of equipments like printers and applicators.

These systems are typically also calibrated to the primary and secondary test systems, but are often configured for particular functions and locations with less of a need to replicate them at all locations. An exception to this is a sweet spot tester, which is discussed separately.

Application test systems

An example of a tertiary system used in applications testing is shown in figure 9. This is a test cube that has been reconfigured with the addition of a rotating platform such that a tagged product may be rotated in front of the antenna and its angular response determined.

Figure 9: **Angular response tester**

The tester is calibrated using an identical method to that used to calibrate the test cube, and uses essentially the same control software, except that extra functions are provided to control the rotary table and to generate polar plots of tag performance as shown in figure 10 below.

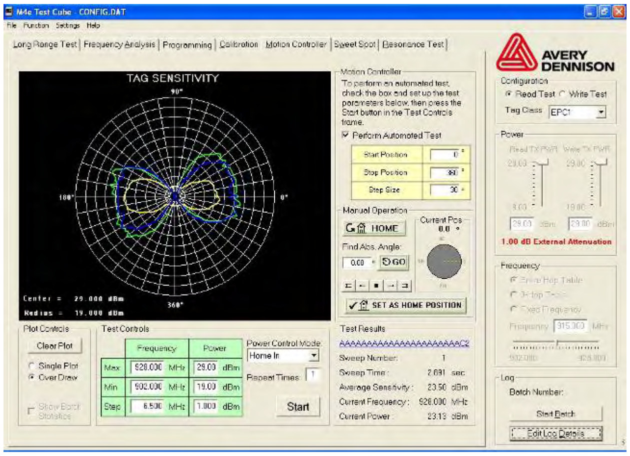

Figure 10: **Angular response controls and plot**

Our objective here is to derive meaningful correlations between tag performance and real world applications such as portals or conveyers. To do this, we are in the process of collecting large amounts of data from test systems that emulate typical end use applications, and cross correlating this with data gathered from more controlled test rigs.

This is relatively straight forward to do at portals, but conveyer performance is currently highly environment specific and results are greatly influenced by such factors as metal rollers, guard rails, etc. An example of the challenge is shown in figure 11, which maps the read count at a conveyer against height (y axis) and time/position (x axis). The plot shows both the effect of metal rollers when close to the tags and also a secondary hot zone probably due to local reflections.

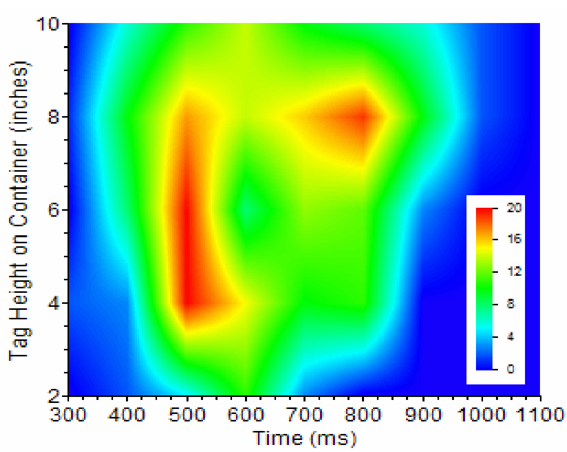

Figure 11: **Conveyer performance at different heights**

Sweet spot tester

With a very large number of different products packaged in different, and changing, ways, putting all of these through full application testing to determine optimum RFID tag selection and "sweet spot" (optimal mounting position) using current test systems presents a serious challenge to the industry.

A number of companies have done an outstanding job in pioneering specialist test systems that allow the so called "sweet spot" to be identified.

Our European team came at the problem from an alternate direction in response to a request from a European retail customer, who has successfully been using a prototype system for a number of months. This system provides a simple and quick method for end users to select a tag, and to determine the optimum mounting position and to gain an estimate of the associated performance.

This sweet spot tester has therefore been incorporated within the test pyramid as a tertiary system and is usually used as a precursor for application test in scenarios where a user is prepared to accept an exact tag placement (in some cases this is not attractive because of either applicator limitations or the cost of manual labor).

As may be seen in figure 12, the tester has two main parts; a test head, consisting of a block of low dielectric constant material where the tag being tested can be mounted on the front compliant surface. This is driven by a near field coupler mounted in proximity, and a reader / software system, attached to the head via a coaxial cable which measures the performance of the tag.

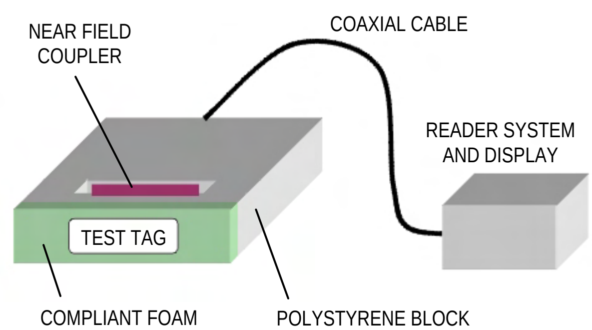

Figure 12: **Sweet spot tester configuration**

The main advantage of this technique is that it allows conventional tags to be stuck on the front of the tester (see figure 13) and does not require any additional equipment other than a conventional reader. In use the tester is moved across the surface of a product and the tag performance is measured. We have found that in most cases the results correlate very well with the application test results carried out with a test cube system.

Figure 13: **Tag mounted on prototype sweet spot tester**

One note of caution: although the sweet spot is generally considered to be the location at which the tag works best on the product, we have seen examples where the ideal sweet spot identified statically is not the same as that in real end use applications – again due to the local environment.

However, in most circumstances the tester will give good results and we have therefore also enabled sweet spot functionality within a derivative of the test cube control software. This allows for products to be probed either in a manual or automated manner and will provide an instantaneous reading of tag performance plus a trailing history indicator, and optional area maps. The skyline seen in figure 14 shows the performance of a tag as it is moved across a tray of bottles containing liquid. As would be expected, good performance is seen when the antenna tips are in air, and lesser performance is seen when over the liquids.

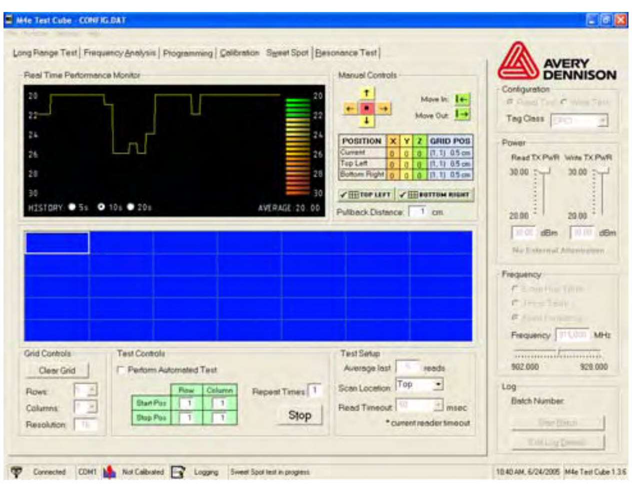

Figure 14: **Display of sweep of sweet spot tester across a tray of bottled liquids**

Concluding remarks

The test pyramid concept provides a framework for Avery Dennison's continuing efforts to develop improved test methods and technologies that ensure the highest possible quality and consistency for our products. By sharing our experience, we hope to assist the RFID community move ahead in this key area. The use of common and well controlled test methodologies will allow users to make valid comparisons between RFID products, and improved quality and performance will enhance user experience and speed deployment and adoption.

The authors would like to thank EPC Global for the opportunity to share this information as well as the many friends and colleagues across the industry, who have supported and encouraged us in this work.