Filtered Waviness and PSD Calibration and Correlation

By THôT Technologies, Campbell, California
and
HGST (formally IBM Corporation), San Jose, California

General Information

IBM (San Jose) personnel initiated this long-term project while Hitachi Corporation had not yet acquired the IBM Disk Drive Storage division. Throughout this document we shall refer to either of the IBM/HGST Corporation(s), or the IBM/HGST personnel, using HGST Corporation, or simply HGST.

The understanding here is that the calibration / correlation project spanned several years and what is reported here is the condensed version of "what worked", a summary of the efforts on all sides. Any blind alleys traveled down have been ignored unless a valuable lesson was learned; in which case this "wrong path" is referenced in the application note to warn the reader of potential pitfalls.

Challenges

The initial challenges for HGST were internal issues concerning;

1). What is the absolute value of any surface morphological measurement?

2). How repeatable is that measurement?

3). How repeatable is that measurement on a day to day basis?

In engineering terms we are asking about are the accuracy, precision and stability of equipment used in these measurements. In other words, What is the gauge capability of the equipment?

A second consideration (and an important one) concerns the correlation machine to machine. Issues for HGST concerned maintaining correlation between up to 9 (different) pieces of measurement equipment used in the manufacturing as well as correlation between measurement equipment of the same type used in different locations throughout the world.

What was (ideally) required was one piece of test equipment that could replace as many of the other (various and varied) pieces of measurement equipment used, thereby simplifying extant correlation problems and streamlining metrology both internally (within a single manufacturing facility) and worldwide.

Section 1. Roughness and Waviness Tests: Calibration and Correlation

Initial Planning

Initial planning for the calibration and correlation consisted of obtaining realistic estimates of the capabilities of each tester and working out a practical method of both “calibrating” and more importantly correlating both the THôT Tester and the MicroXam.

1.1). Estimated noise (measurement) floor

From past experience a realistic estimate of the noise floor for both testers was presented for both the THôT and MicroXam. Consideration included were power line interference, environmental pollution (e.g. airborne noise and 50-60Hz lighting), ground vibration and the proximity of support people.

The estimated noise floor for THôT Testers was 0.01Å-0.05Å. The reason for this (extremely) low estimate is threefold.

a). The THôT technique is to make dynamic measurements. Because the disk is spinning any static or quasi-static (low frequency) components can be ignored or filtered. What this means in practical terms is that THôT testers inherently have a very high rejection ratio for environmental pollution (60Hz power and light fluctuations) and both airborne and “groundborne” vibrations (speech, footsteps, traffic, earthquakes etc.)

b). The Doppler Laser method used measures frequency differences between the outgoing signal and the return beam. This technique is highly developed and measurement of 1 part in 10^{18} is commonplace for any standards laboratory. For the 633nm (HeNe) laser used in the THôT Tester, a very modest measurement of frequency difference (1 part in 10^{-6} is used) will give an estimated resolution of 633×10^{-6}nm (0.01Å)

c). The return signals frequency change is proportional to the instantaneous velocity of the disk surface. Integrating this velocity signal to generate displacement information. Any integrating operation inherently reduces measurement uncertainty, however it is induced (e.g. electronic noise or random vibrations).

The estimated measurement floor for the MicroXam was 2Å-3Å. The reason for this estimate is partly practical and is partly based on past experience.

a). The MicroXam is an interferometer. The measurement resolution is based on the wavelength of the light used (either white light or monochromatic red) and the number of interpolation steps that can usefully be made between the interference fringes. Assuming monochromatic red light (≈650nm) and an interpolation algorithm capable of resolving 256 steps (which is a respectable number), the best case capability would yield a measurement step of ≈26Å. Further averaging over the measurement area (typically 256 or 512 pixels) will improve this figure by √(pixel count) i.e. by a factor of ≈16-23, giving a theoretical resolution of ≈1-2Å.

b). The MicroXam is an interferometer. This implies a static measurement and any static measurement has to deal with (or compensate for) environmental pollution as well as the airborne and "groundborne" vibrations outlined above. Conventional wisdom has the MicroXam mounted on a vibration isolation table which is housed in an anechoic (or sound deadened) enclosure. With these precautions (which are always less than perfect), adding 1Å to the measurement uncertainty is very reasonable.

1.2). Measurements

The first step in the calibration/correlation was to select and/or make disks that had a range of roughness and waviness values. For roughness (Rq) and μ-waviness values >5Å the process was a simple selection from current disk production and from end of life programs (older, archived disks). However producing the very smooth disks with morphological features in the <<3Å range required co-operation from many parties (slurry manufacturers, plating and polishing personnel etc.) and was to some extent "hit and miss" as the only machine capable of reliably resolving fine detail was the THôT testers.

The differing noise floors indicated that a wise choice would be to use multiple MicroXam measurements on the same area in an effort to reduce background noise, because it was necessary to attempt measurements in the <<2Å range. The choice was made to average 25 independent MicroXam measurements in the hope that statistical averaging would yield a "averaged" measurement that was representative of the surface. Averaging in this fashion should reduce random variations by √25 and a realistic surface roughness or waviness measurement could be made as low as 0.5Å using the MicroXam.

The last item was to determine the wavelength range used in the filtering algorithm for the MicroXam. This is not as easy as it appears. Descriptions of the filtered wavelengths are not clearly stated in the MicroXam literature. However, after discussions with the manufacturer, it was determined that the approximate wavelength range, for the μ-Waviness setup used, was 100-500μm. These setting were applied to both the MicroXam and THôT testers.

Similar measurements were also attempted using the AFM. The small region scanned, 10μm by 10μm, limits the measurable long wavelength range to ≈5μm. It is possible to get some estimate of longer wavelengths (with concomitant loss of resolution) to ≈15μm. Again the same scheme (multiple averages and multiple locations) was used although AFM measurements are both tedious and time consuming.

The next obstacle was the variation of the surface around the disk. There is a preponderance of engineers and technicians in the disk (and drive) manufacturing fraternity that pointedly ignore these circumferential variations and assume that two or three AFM and/or MicroXam type measurements averaged together will be a representative measure of the surface morphology. Why should this situation have occurred? Three reasons immediately come to mind:

1). Ignorance of mechanical polishing techniques,

2). Sensitive measurement techniques have not been available; the “ If you cannot measure it you cannot fix it” mentality

3). Multiple measurements take too much time to perform in a high throughput environment.

Logically this situation is surprising. Consider this, the disk is used “dynamically” i.e. it spins, the head travels in a circumferential direction around the disk. Therefore one question springs to mind. How can the engineer know how the head and disk will perform, as a system, from a single, spot, static measurement? This is the one “wrong path” traveled that we (THôT and HGST) must warn against if successful evaluations are to be made of surface morphology in the <3Å regime.

This is the major “wrong path” traveled. If static measurements are used to measure surface morphology in the <3Å regime, the user MUST use multiple averaging to reduce random variations AND make multiple measurements around the circumference.

After much engineering work a realistic (workable) solution was found. The bottom line is that good correlation was achieved if (for the MicroXam):

1). Each measurement region was measured 25 times and the spot readings averaged, and

2). The minimum number of circumferential samples is set to 24 (i.e. circumferential sample interval of 15°)

Examples showing circumferential variation for both THôT and MicroXam are shown below. Note that these are examples only and are not the same disk surfaces. The data set for the THôT example has been averaged (both radially and circumferentially) to mimic the measurement region of the MicroXam.

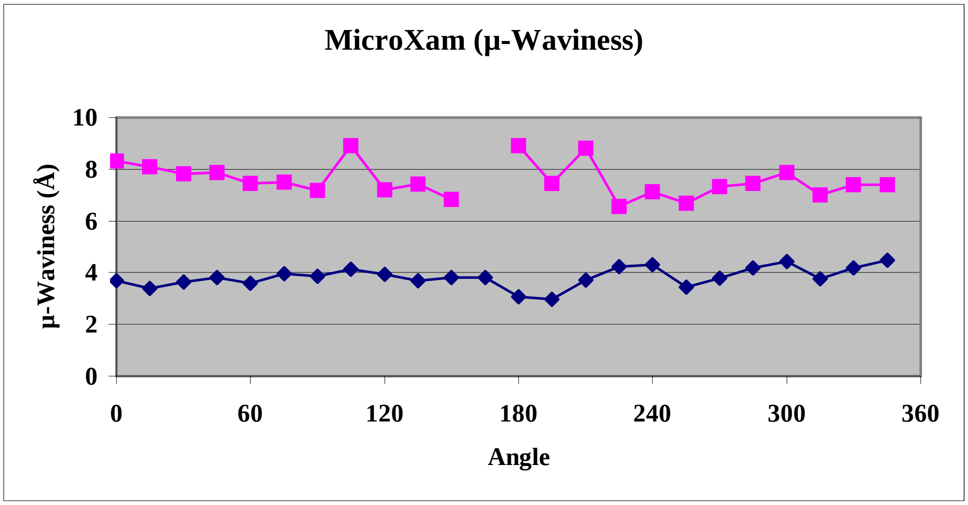

Figure 1. A typical MicroXam data showing circumferential variation

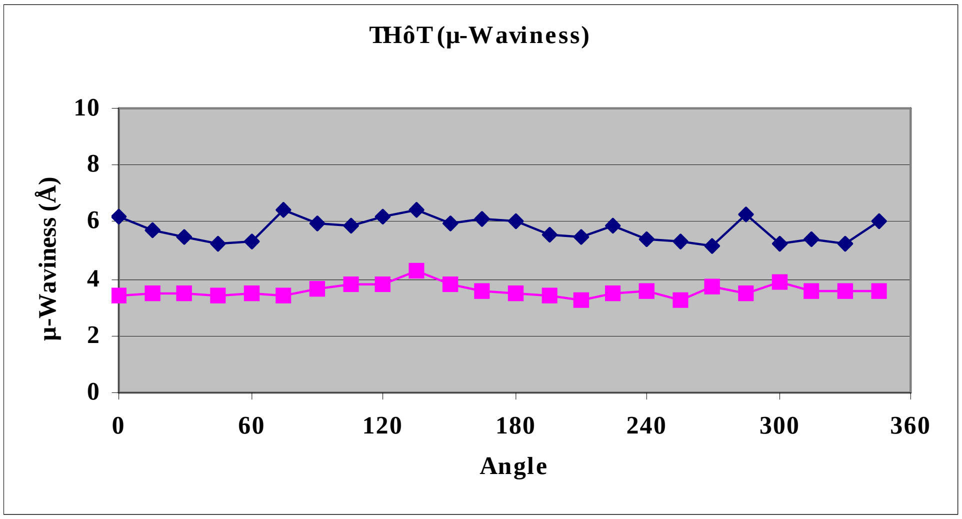

Figure 2. THôT data example showing circumferential variation. (Data set is reduced to "mimic" the MicroXam sampling interval)

1.3). Correlation

Once the initial engineering groundwork had been finished then it was ready for “show time”. Using a selection of disks with a range of “known” roughness and waviness values, the correlation exercise began.

For both the roughness and the µ-Waviness ranges it is necessary to firstly determine the wavelength range that is inherent in the measurement tool. In some cases this is not easy to determine. In other cases (such as the THôT testers) the wavelength range is programmable and obvious. For the AFM, the best estimate is given using the scan length and the number of micro-steps (usually 128 or 512). For a 10µm x 10µm scan this would give a longest wavelength of ≈10/√2 (≈7µm) with partial information up to ≈2x10√2 (≈15µm). The shortest wavelength possible is ≈2x(10/512)µm (≈0.05 0.2µm).

This is another “wrong path” often traveled. If different measurement devices are used to measure wavelength sensitive surface morphologies such as roughness and µ-Waviness, the user MUST have knowledge of the valid wavelength range *of that piece of test equipment,* before attempting any correlation.

For the intermediate and high values of roughness and µ-Waviness, the correlation exercise proved successful, a tribute to the extensive groundwork from all parties involved. Data is shown below.

Note: The only failure is that initial target was to extend the upper measurement region to 20Å. Unfortunately, the only archived samples suitable for measuring, that were readily available, had an upper limit of ≈12Å.

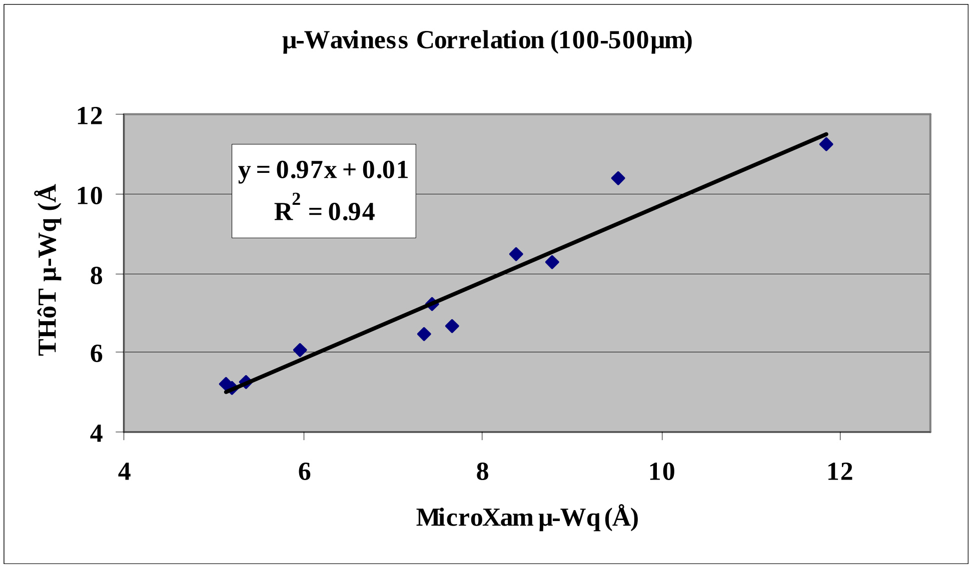

Figure 3. µ-Waviness Correlation, (THôT with MicroXam)

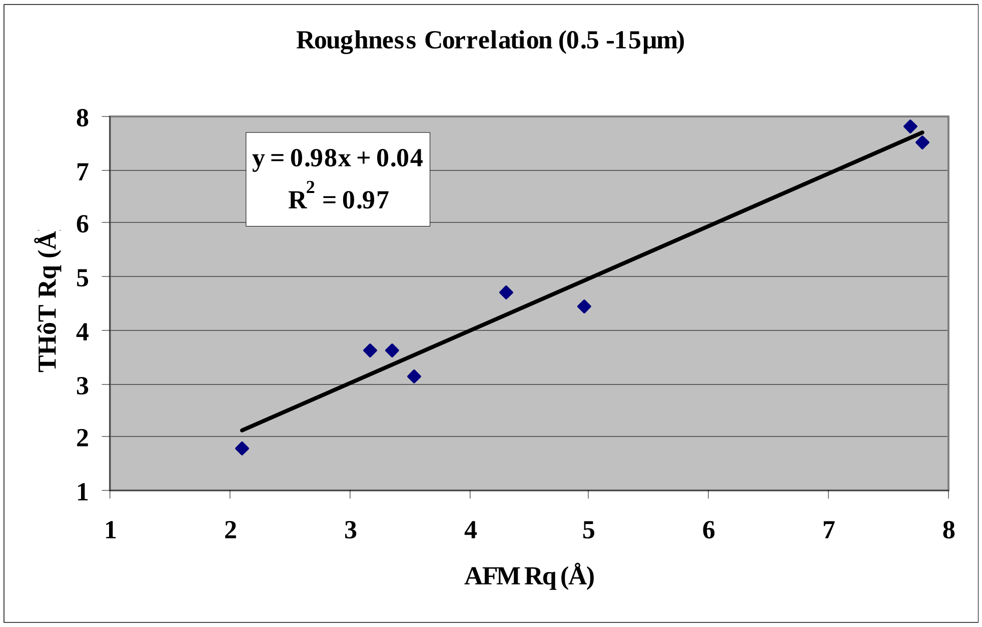

Figure 4. Roughness Correlation, (THôT with MicroXam)

1.4). Correlation over all wavelengths and repeatability

The final step in this exercise was a correlation of surface morphology measurements over all wavelengths of interest. For this exercise, because of the equipment available, it was convenient to choose:

Roughness (< 5µm),

n-Waviness (60 to 160µm),

µ-Waviness (100 to 500µm),

Waviness (500µm to 2mm).

This exercise required that different pieces of measurement equipment AFM, SMS, MicroXam, and Optiflat, all be correlated to the THôT tester. Remember that part of the initial requirement was that, wherever possible, a single measurement device should replace multiple pieces of test equipment. (Refer back to the “second consideration” described on the title page). Correlation data is shown below in Figure 5.

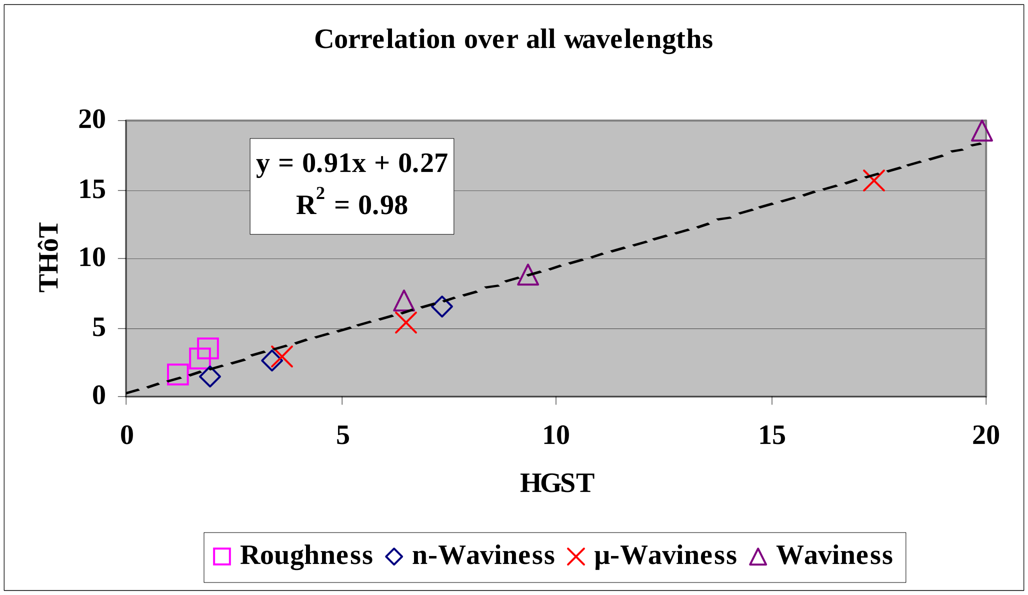

Figure 5. Correlation over several wavebands.

The last pertinent data necessary is the question of day to day repeatability. What is the stability of my test equipment and what variation can I expect in my measurements? The graph below (Figure 6) demonstrates a series of 10 measurements made over a period of 2 weeks using the same set of 10 disks. The data taken are Waviness data (400µm to 2mm). Waviness wavelengths represent the “worst case” in terms of repeatability as environmental factors can begin to affect static measurements and mechanical considerations such as disk flutter will begin to creep into these measurements if the disk is tested dynamically.

So using this waveband sets up “equal” disadvantages for the THôT tester, dynamic measurements with disk flutter influencing the repeatability, and the MicroXam, static measurements with airborne noise and floor vibration influencing the repeatability. In the graph shown, the error bars are the 1σ limits for both the THôT Tester (vertical error bars) and the MicroXam (horizontal error bars)

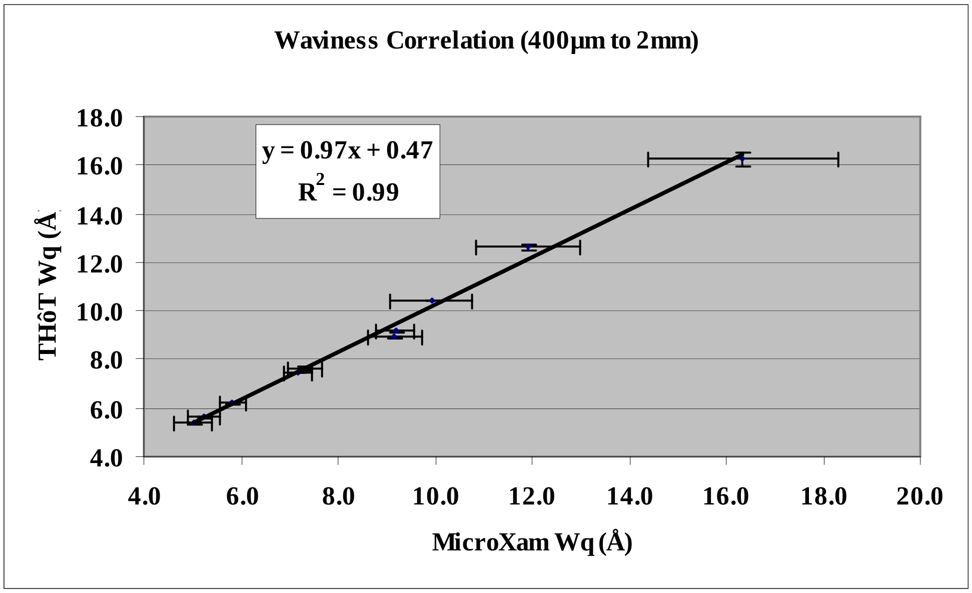

Figure 6. Test equipment repeatability and long term stability.

1.5). Linearity at <1Å and electronic calibration

The next (and final) hurdle is to evaluate the linearity of the MicroXam and AFM to truly determine their capability to measure low values of µ-Waviness and roughness respectively.

1.5.1). Linearity at <1Å

Figures 7, 8, 9 and 10 demonstrate the linearity of the AFM and MicroXam compared with the THôT tester. Both testers showed good linearity and correlation when measuring within the capabilities of the testers. However, in both instances the MicroXam and the AFM plateau at a noise floor of ≈1Å and ≈3Å as expected. Only the THôT tester has the capability to differentiate surface morphology in the <1Å range.

It should be noted that an attempt was made to assess the true “zero” for each tester. So the first datum is not a measurement made on a disk but a “null” measurement to determine baseline noise.

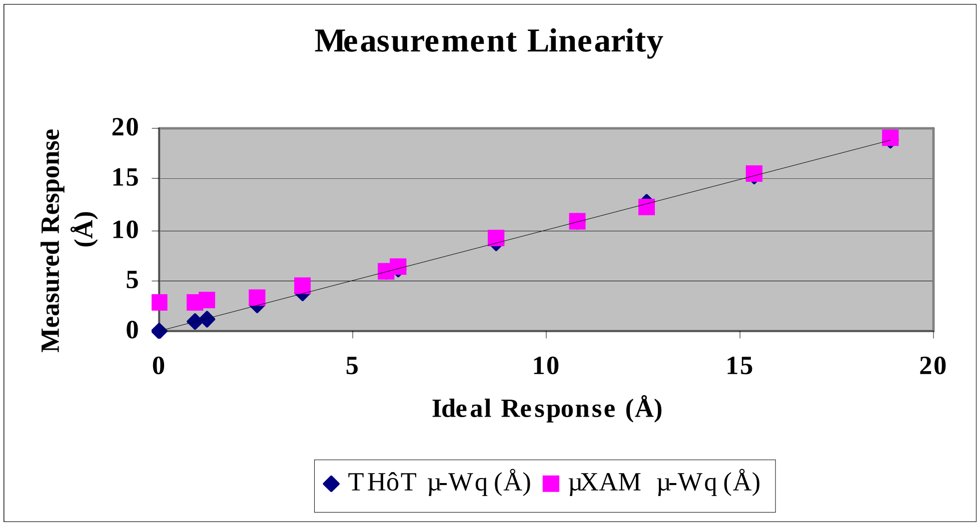

Figure 7. MicroXam linearity.

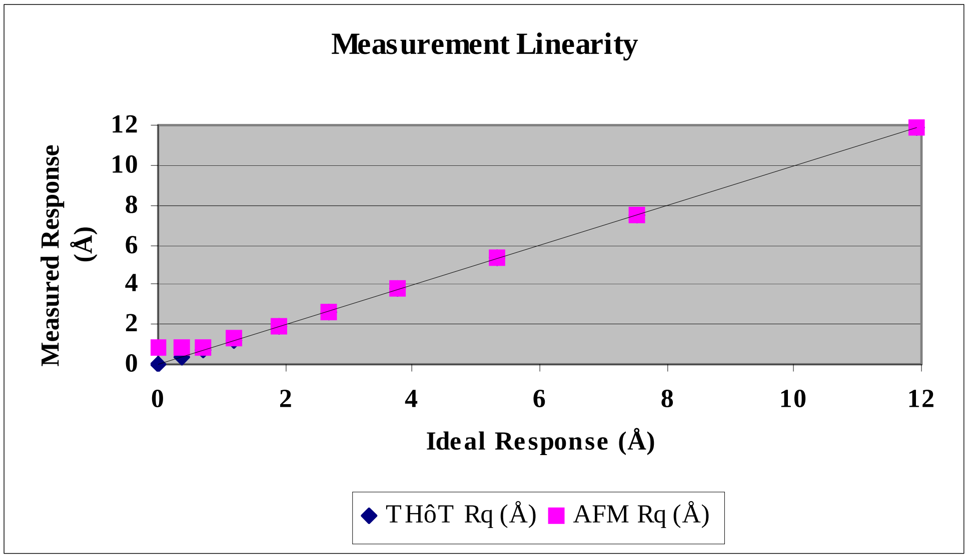

Figure 8. AFM linearity.

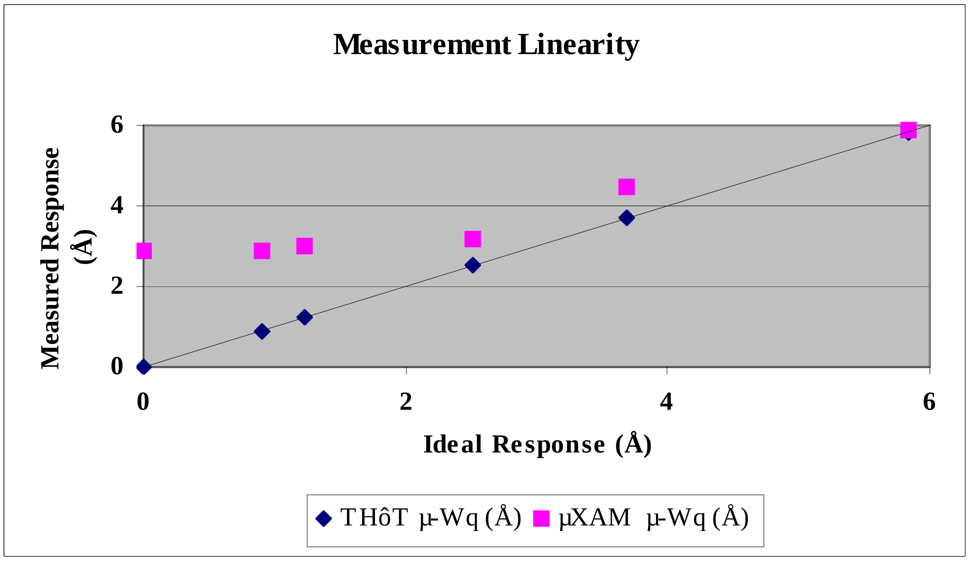

Figure 7. MicroXam linearity below 6Å.

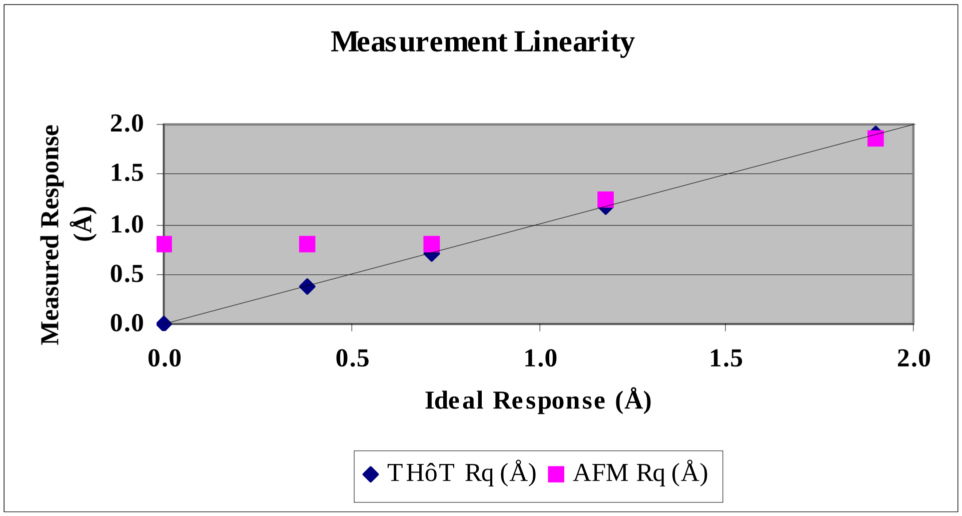

Figure 8. AFM linearity below 2Å.

1.5.2). Electronic Calibration

There are no standards available for sub-Ångstrom calibration. The best available transfer standards for AFM calibration are ≈8nm (≈80Å). Conventional wisdom dictates the use of a standard that measures in the "ballpark" of the attempted calibration range; in fact calibration should fall between 30 to 90% of the available range. Clearly using a standard that measures toward the top of the range and attempting to "calibrate" in the lower 1-2% of the range (the non-linear region) is risible.

Using a pure sine wave it is possible to set up an electronic "measurement". Given the rotation rate (RPM), radius, sine wave frequency and the sine amplitude, a numerical calculation can be made to determine an “absolute” value (for a single wavelength) and this can be compared with the measured value. The limitation here is only the measurement accuracy of the sine amplitude. Figures 9, 10 and 11 show the results of these tests (for the THôT tester only!) and it is clear that the linearity is outstanding for measurements above 0.1Å and only slightly degraded when measurements <0.05Å are considered. In statistical terms the gauge capability far exceeds any alternate test methods.

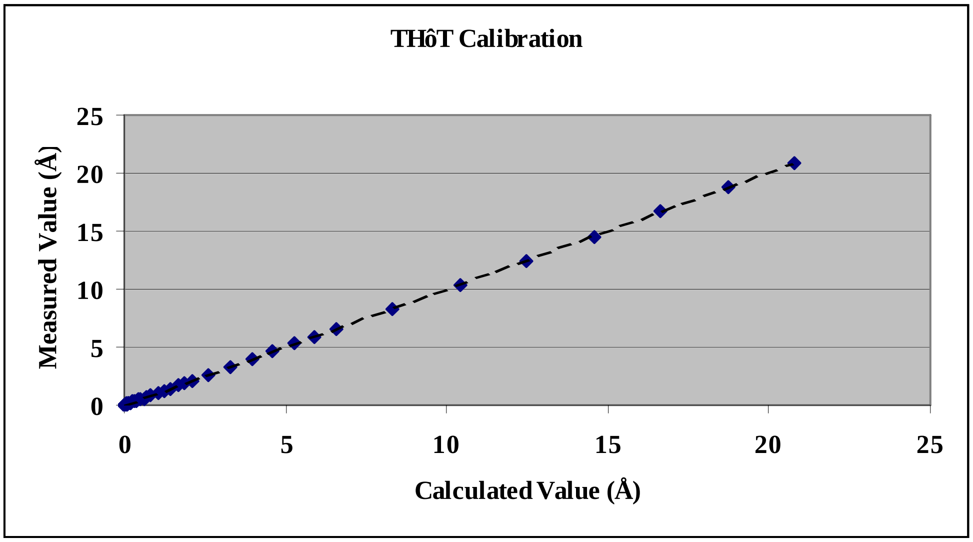

Figure 9. Electronic Calibration (0 to 20Å).

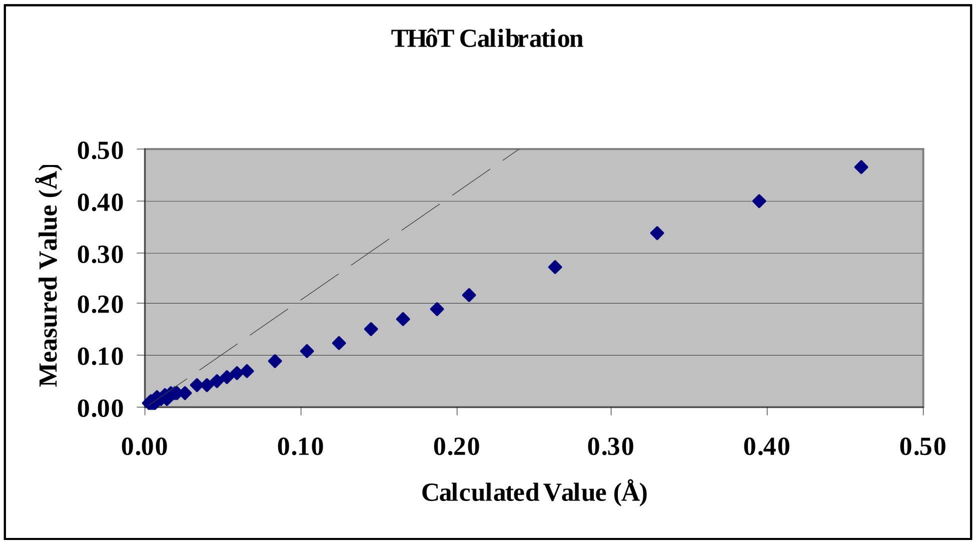

Figure 10. Electronic Calibration (below 0.5Å).

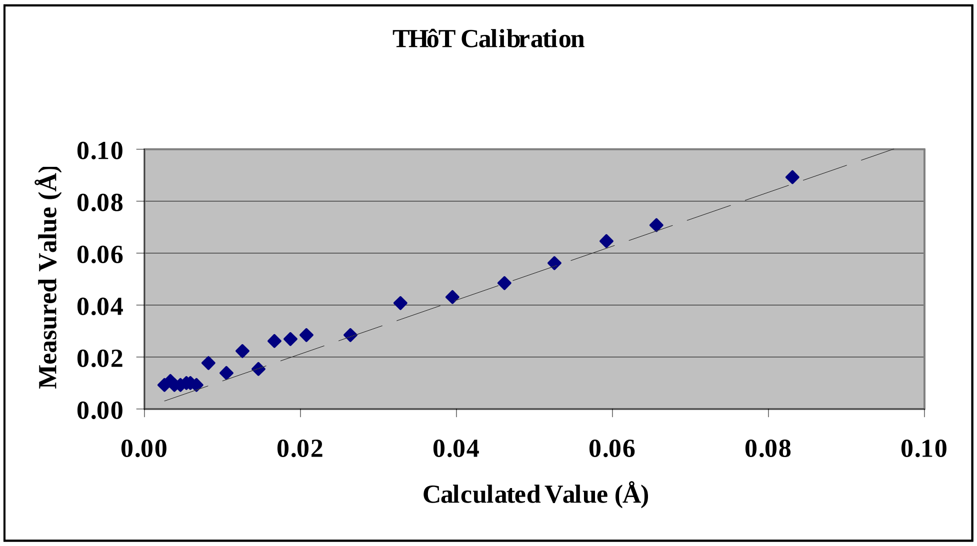

Figure 11. Electronic Calibration (below 0.1Å).

Section 2). PSD Tests: Calibration and Correlation

2.1). IDEMA Wavebands

Establishing the basic capabilities with respect to accuracy, repeatability and stability using any wavelength range is 90% of the battle, The remaining 10% is essentially "detail" work but nevertheless important. The first task is to define certain wavelength ranges as common ground such that everyone can speak about surface morphology using a common language. This is not to say there can be no deviation from theses standard definitions, in fact deviations are common and necessary because heads, disks, and drives are designed for varied capacities and applications. IDEMA, the disk and drive advisory body has established 8 wavelength bands that represent the various morphological categories. These are:

Name	**Symbol**	**Wavelength range**
Macro-Waviness	(MW)	2mm to 5mm.
Waviness	(W)	400µm to 2mm.
Micro-Waviness	(µW)	100µm to 400µm.
Nano-Waviness	(nW)	50µm to 100µm.
Macro-Roughness	(MR)	25µm to 50µm.
Roughness	(R)	15µm to 25µm.
Micro-Roughness	(µR)	5µm to 15µm.
Nano-Roughness	(nR)	< 5µm.

If one establishes these categories, the logical question becomes: What is the effect on measurement if I have one dominant wavelength compared with a collection of smaller, less obvious wavelengths?

Approaching this from a mathematical perspective, a longer wavelength of a fixed height has more energy than a shorter wavelength with the same fixed height. The reason is because the average power of the wave depends not only on the square of the amplitude (considered fixed here) but also the square of the frequency as well. Frequency is inversely proportional to wavelength so the shorter the wavelength, the higher the frequency. So, within a specific waveband, a longer wavelength is weighted more heavily than a shorter one.

What is needed is a method to isolate and measure a narrow waveband to be able to assess its individual contribution to the overall morphological structure. This is akin to a spectrum or a Fourier analysis where a complex waveform is broken into frequency (or wavelength) components each with its own amplitude (and phase descriptor in the case of Fourier series).

2.2). PSD Testing

For PSD testing using the THôT tester, 128 narrow bands are measured. Each band “width” is adjusted in a logarithmic width progression to provide equally spaced “point” measurements. A wavelength “spectrum” can be built up using these 128 measurements and a curve generated that allows the user to see and assess the various components within any of the standard IDEMA wavebands. A typical PSD graph is shown below in Figure 12.

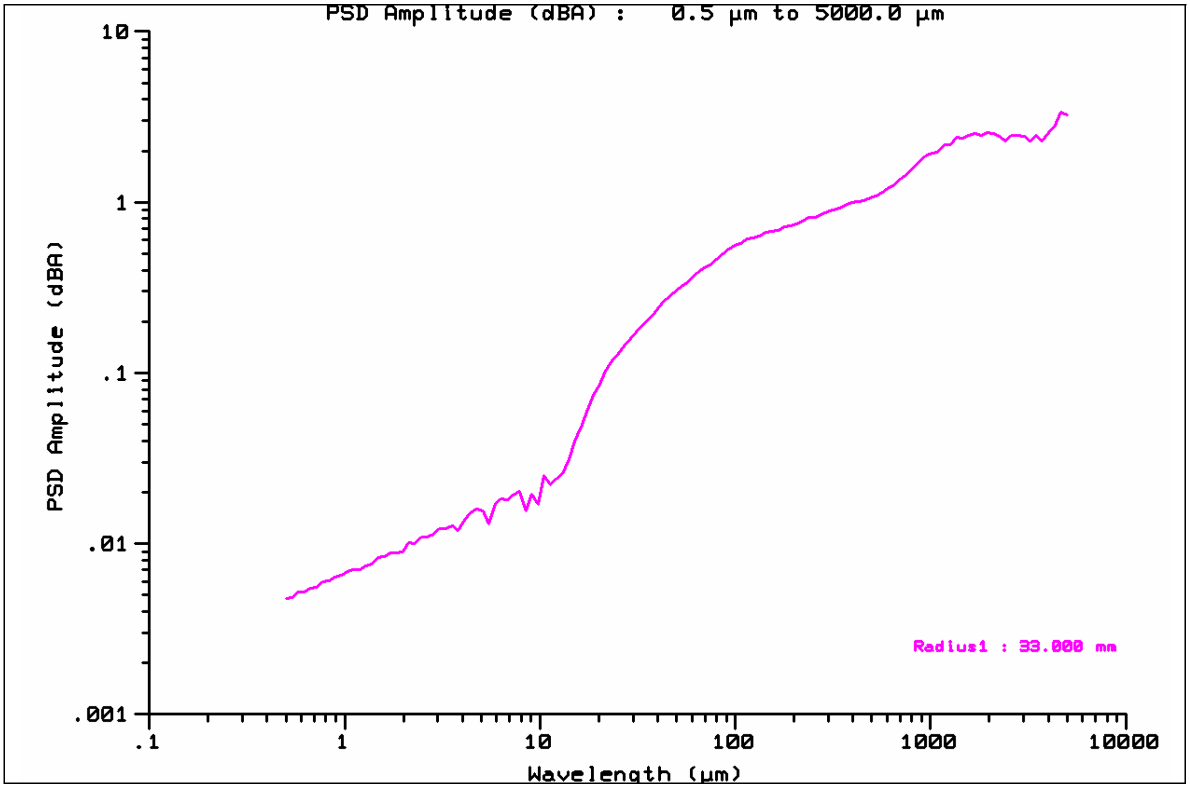

Figure 12. PSD output from THôT tester.

For the THôT PSD test, the range of wavelengths has been fixed to cover all IDEMA wavebands from 0.5µm to 5000µm. The horizontal (wavelength) axis is scaled in µm while the vertical (magnitude) axis is scaled as dB(Å). The 0dB reference has been fixed at 1Å thereby making conversion (integration) to any waveband a simple exercise. ASCII text files are automatically generated and the user has access to each PSD datum (magnitude and wavelength).

2.3). PSD Standards Generation

To date we have used the IDEMA wavebands as a transfer reference between the THôT tester, the AFM, and the MicroXam. We have also established the THôT tester as (presently) the only tool to easily resolve sub-Ångstrom differences in surface features. This tool and these same techniques can also be used to generate standards that can be used as transfer standards. The PSD data is in fact the ideal transfer data set since it can be made to apply to any waveband for any other measurement tool – static or dynamic.

Generation of the PSD standard is now straightforward. Using a calibrated THôT tester (as outlined in Section 1), one has to generate a reference PSD curve and establish measurement limits, rather like running a control chart. HGST engineers and THôT personnel selected five “identical” Gold Standards that are used to generate Silver standards. These Silver standards are used to generate the working standards that are shipped with each tester.

The THôT technique is to run the PSD test 36 times and from these data calculate a mean value and the standard deviation for each of the 128 data. The mean PSD curve for the candidate standard has to closely match the archived data from the Gold standard. Note that the Gold standard is only measured under exceptional circumstances and is not used as the reference. If the candidate falls within acceptable limits (within ± 0.5σ deviation from the Gold standard) then this is considered a working standard. At this point the mean PSD line and the measurement deviations are recorded. The data file is burned to a CD and both the disk and CD are identified with a unique, matching, serial number.

2.4). PSD Correction Factors applied to the PSD graph

Because of the outstanding repeatability of the THôT tester, the correction factors are typically very small. Typical correction factors are between zero and 0.02Å. When applied to a PSD curve, these correction are simply added to (or subtracted from) each individual measured value. These factors are automatically applied when a PSD calibration is performed on the THôT tester. Calibration is straightforward and verification of the calibration is given at the end of the calibration routine. A typical Calibration/Verification graph is shown in Figure 13.

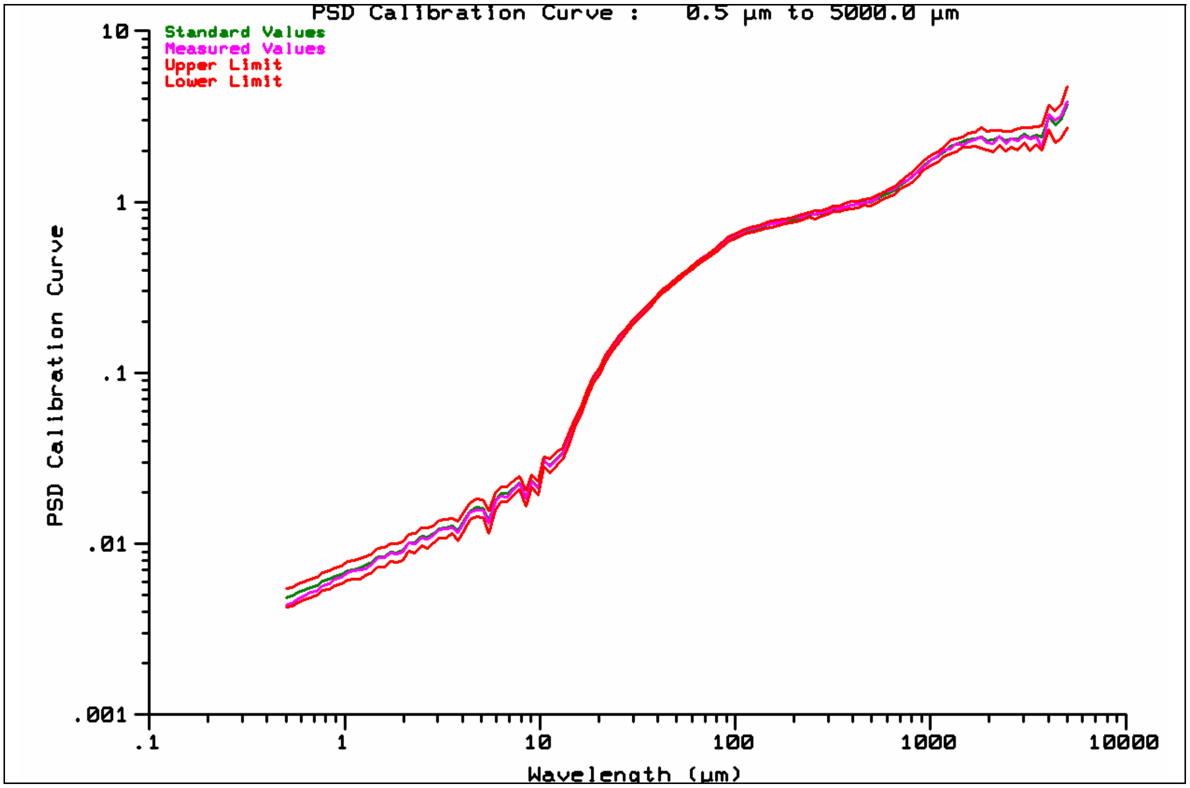

Figure 13. PSD Calibration Verification from a THôT tester.

2.5). PSD applied to IDEMA wavebands

How to apply correction factor from PSD tests to the 8 band IDEMA wavebands? Firstly it should be noted that PSD corrections can be applied to ANY waveband, not just the IDEMA wavebands specified in Section 2.1.

Technically each waveband should be de-constructed into a Fourier series with each element of the series analyzed for magnitude and phase. The wavelength limits should be the same as the PSD intervals and the magnitude and phase information preserved. Each correction value is then applied to the argument of each individual element of the Fourier series. Then the waveband has to be re-constructed (again maintaining any phase information) to give the corrected measurement integrated over the specified waveband.

Complicated, time consuming and requires more data storage (RAM) than a computer can currently handle!

Fortunately there is a simpler method that gives a very close approximation to the theoretical method. This method could be called “RMS value” method. Since we are measuring an RMS value for roughness (Rq) or waviness (Wq), a close approximation is achieved if each of the individual correction factors, i.e. the ones that fall inside the specified waveband, are squared, summed and the square root taken. This value is then directly applied (added to or subtracted from) the measured value. The trick here is to preserve the sign of the individual correction factors since sign information is lost on squaring. This technique works only if the correction values are small compared with the measured value. Fortunately the standards generation method used (cf. Section 2.3 above) ensures that this condition is met.