AFM Length Analysis of Data Marks:
Measuring Jitter, Asymmetry, Process Noise and Process Position

Donald A. Chernoff and David L. Burkhead
Advanced Surface Microscopy, Inc., 6009 Knyghton Rd., Indianapolis IN 46220-4955, USA
phone: 317-251-1364; fax: 317-254-8690; E-mail: donc@asmicro.com; www.asmicro.com

ABSTRACT

We describe statistical analysis of AFM measurements of bump size, shape and position on DVD stampers. We present statistical concepts that lead to useful measurements of process position (such as bump length) and process noise (bump length variation). These physical measurements are compared with key electrical measurements such as asymmetry and jitter.

Keywords: Atomic force microscopy, jitter, pit geometry, width, angle, height, optical discs, statistics, process control, length analysis, ANOVA, length deviation, length offset, asymmetry.

1. INTRODUCTION

The Atomic Force Microscope (AFM) has long been used for qualitative and semi-quantitative analysis of the topography of compact discs and DVDs. We have previously described a technique for calibrating and automating AFM measurements, so that nm-scale precision and accuracy can be achieved.[1] Because we can now measure the size, shape and position of hundreds of data marks (when only a few were measured previously), further statistical analysis of the results is possible. This paper discusses how to measure various parameters relating to process position (mean values) and process noise (standard deviations). The goal is to show how length analysis of AFM images can help improve process control in optical disc manufacturing. When one compares microstructure with electrical measurements, one can learn which aspects of structure contribute to good or bad performance. In new format development, length analysis can provide key information about the manufacturing process even when suitable disc analyzers are not available.

2. EXPERIMENTAL METHODS

The test specimens included a "power series" DVD-5 stamper containing bands mastered at 11 different laser power levels (ranging from 90% to 112.5% of a nominal value), as well as additional stampers and replicas. The test equipment was a commercial stamper analyzer and an AFM (Digital Instruments NanoScope® IIIA/Dimension 3000 large sample TappingMode™ AFM). We captured AFM images of the surface topography at a scan size of 10x10µm. To assure high accuracy and precision, we captured corresponding images of a calibration reference specimen. For this work, we used a two-dimensional holographic grating (MOXTEK) with a known pitch spacing of 292 nm.

Image analysis.

We analyzed the images using ASM's DiscTrack Plus™ software. As previously described, the software recognizes each data bump and measures its size, shape and position[1]. The *basic* parameters include: height at the summit relative to the land adjacent to the bump, side wall angles on all 4 sides, and the edge positions at half-height on all 4 sides. We then compute various *derived* parameters. For example, from the edge positions, we compute bump width, bump length, and the land length between consecutive bumps on a track.

3. STATISTICAL CONCEPTS

Process position and process noise.

In order to control an optical disc manufacturing process, it is important to establish target values for various measurable parameters, such as bump height. One can then compare the mean height found on a given disc (current *process position*) with the *process target* value. If height varies from one bump to the next, this will add noise to the measured electrical signals when the disc is played. Therefore, it is useful to tabulate the standard deviation, since this is a measure of *process noise.*

Table 1. Some examples of process position and noise parameters.

Position	**Noise**
Track Pitch	Pitch Variation
Bump Length (offset, deviation, asymmetry)	AFM Jitter
Bump Width	Width Variation
Bump Height	Height Variation
Side wall angle (slope)	Slope Variation

A Statistics Tutorial.

Because many scientists and engineers have no formal training in statistics, we give a brief tutorial about two valuable tools: descriptive statistics and analysis of variance. A third tool, linear regression (curve fitting), is used without tutorial explanation because most readers are familiar with it.

Descriptive statistics.

A typical image of a stamper contains more than 90 bumps. We can summarize the results for height, width, length, and slope angles using descriptive statistics such as mean, standard deviation (SD), and standard error of mean (SEM). In order to explain the meaning of these statistics, let us step back and consider the measurement process in general. When we make measurements of a parameter (for example, bump height), we usually find that there is some variation between measurements: some numbers are higher and some numbers are lower. In a normal distribution, the Mean value (also called 'average') is the most probable measurement value. The formula is:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \text{ , where the } x_i \text{ are the individual observations of parameter } x. \qquad (1)$$

SD describes the spread (variation) of individual observations. In a normal distribution, approximately 95% of the individual measurements fall within 2 x SD of the mean value, that is, in the range $\bar{x} - 2 \times SD$ to $\bar{x} + 2 \times SD$.
The formula for SD is:

$$SD = \sqrt{(Variance)} \text{ , where} \qquad (2)$$

$$Variance = \frac{1}{(n-1)} \sum_{i=1}^{n} (\bar{x} - x_i)^2 \qquad (3)$$

If we make a number of experiments (independent sets of measurements) and compute the mean values from each experiment, we usually find that there is some variation between the means. However, the variation of the means is much less than the variation of the individual measurements in each set (Some of the individual values are greater than the mean and some are less than the mean; these values tend to balance each other). SEM describes the variation of mean values. The formula is

$$SEM = SD / \sqrt{n} \qquad (4)$$

We often use SEM values to judge whether an apparent difference in mean value between two sets of measurements (for example, bump height on two stampers) is "statistically significant". This test, called the "t-test", proceeds as follows. Given mean values M_1 and M_2 and SEM values SEM_1 and SEM_2, we compute

$$t = \frac{(M_1 - M_2)}{\sqrt{(SEM_1^2 + SEM_2^2)}} \qquad (5)$$

If $t > 2$ then we say that the difference in mean values (M_1 - M_2) is statistically significant. The value 2 corresponds approximately to 95% confidence that the difference is systematic (not due to a random process).

A separate question is whether the difference is practically significant. This requires a judgment by the scientist or engineer responsible for the process. Sometimes we find that the minimum difference that is statistically significant is larger than the practically significant difference we would like to detect. This means that the precision (or sensitivity) of the mean value measurement must be increased. In order to see how to do this, note that the SEM value depends on a combination of factors:

- the intrinsic variability of the parameter (say, bump height) from bump to bump
- the measurement precision of the overall system (for example, AFM plus analysis software)
- the data capture protocol (how many bumps we measured)

The first two factors relate to the observed SD. The third factor relates to n, the number of observations. It is often possible to increase n so that the desired precision can be obtained. Increasing n by a factor of 4 improves precision by a factor of 2.

These basic descriptive statistics turn out to be useful for some parameters but not for others. For example, an AFM image of one disc yielded a mean bump height of 124 nm with SD of 2.9 nm and SEM of 0.28 nm. Because the SD and SEM values are small and because a histogram of height values (figure 1A) indicated that there was a single group, the basic statistics are useful measures of process position and noise. In contrast, for length we found a mean of 561 nm, SD of 227 nm, and SEM of 22 nm. But these numbers are almost useless, because the histogram of length values shows many groups (figure 1B). Of course, we know that length varies a tremendous amount: a T14 bump is almost 5 times as long as a T3 bump, so the mean value measured from a small sample (one AFM image) can change depending on the random occurrence of different T-numbers in the data stream. A different approach is needed.

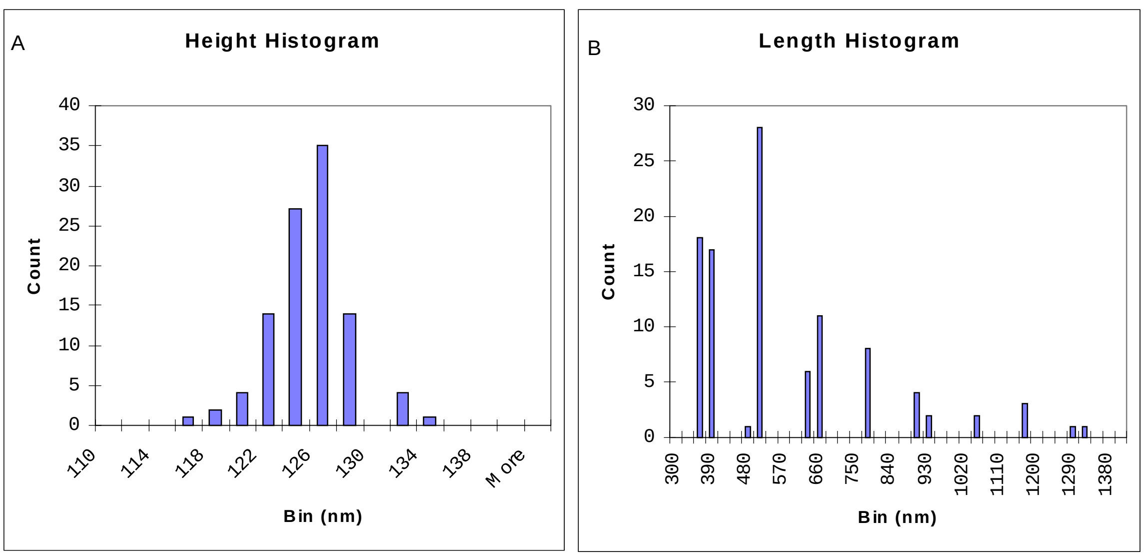

Figure 1. Histograms of bump heights (A) and lengths (B) for a DVD stamper.

ANOVA (analysis of variance).

When we think about bump length on DVDs, we can get sharper measures of process position and noise by looking at the mean and standard deviation of length for long and short bumps, considered separately. In a graph of bump width vs. length, we can see the natural grouping of the data by T-number.

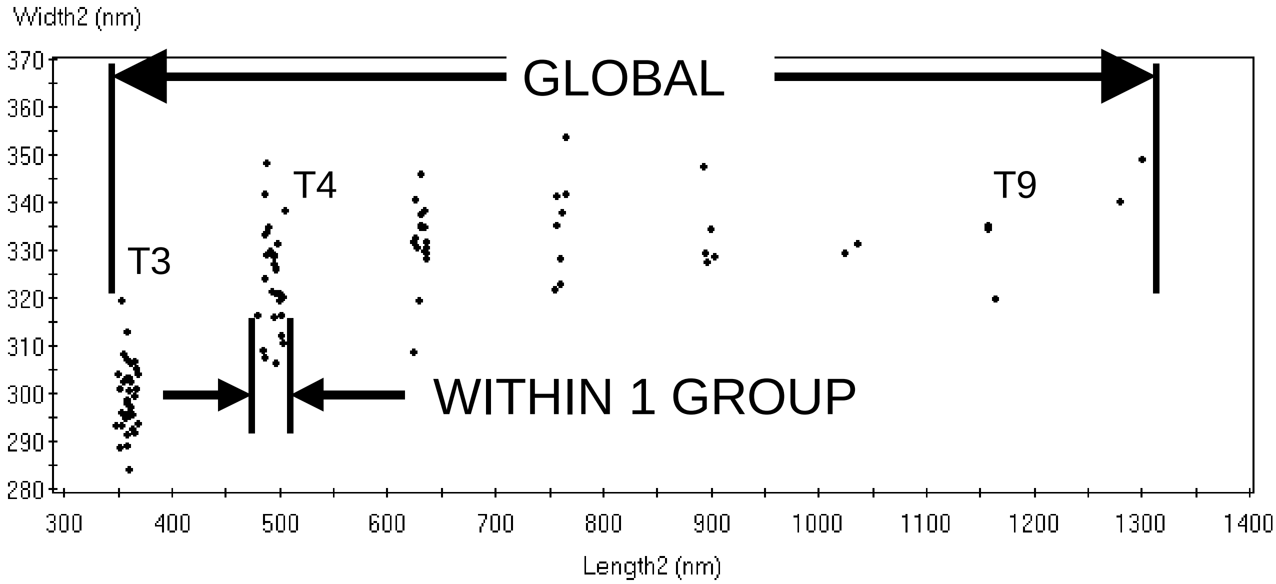

Figure 2. Width vs. Length for bumps on a DVD stamper. Arrows indicate the Global (overall) range of the length observations and the range just within the T4 group. Whereas the global SD was 227 nm, the within group SD (computed by ANOVA, see below) was 5.8 nm.

By inspecting the graph, we see that the mean length for bumps in the T3 group was about 360 nm with some spread (about 10-20 nm) in the individual values. Each group has a different mean value and variation. ANOVA-1 (one-way analysis of variance) is a straight-forward statistical technique to handle this situation.[2] The phrase "one-way" means that we group the data using a single factor, here T# (T-number). We begin by labeling each bump with its T-number, and then compute the descriptive statistics for each T-number group separately. See Table 2.

Table 2. A portion of the ANOVA worksheet bump length (nm).

T#	3	4	5	6	7	8	9	10
N	35	29	17	8	6	2	3	2
Mean	359.88	494.40	631.55	760.97	895.50	1,031.34	1,160.23	1,291.13
SD	5.33	6.58	4.52	4.01	6.92	7.91	4.02	14.86

The within group standard deviation SD_w is computed by pooling (averaging) the standard deviations of the individual groups, weighted according to the number of observations in each group.

$$CSS = \sum_{i=1}^{g} \left(SD_i^2 \times (n_i - 1) \right)$$
$$DF = \sum_{i=1}^{g} (n_i - 1) \qquad (6)$$
$$SD_w = \sqrt{CSS / DF}$$

where g is the number of groups.
For the data in table 2, we find SD_w = 5.81 nm.

Definition of "AFM Jitter".

When an optical disc is played, the player detects pulses corresponding to the leading and trailing pit edges – transitions from land to pit and from pit to land. The timing of the pulses is critical. If the disc and player are perfect, the pulses coincide with the pulses of the average clock. If the pulse timing varies ("Jitter"), it is possible to detect a pulse at the wrong time (low level error). As jitter increases, there are more low level errors and eventually there are uncorrectable high level errors. Customers reject discs that have too many errors.

Therefore, one goal in mastering an optical disc is to put the pit edges in the right places. That means to make every pit the right size and to put it in the right place. In order to investigate the contribution of feature placement error to electrical jitter, we compute geometric edge jitter ("AFM Jitter"). Unlike the PLL (phase locked loop) circuit in a disc analyzer, the AFM cannot scan a large enough region to compute the average clock. But we can infer edge jitter by analyzing the variability of the bump lengths and land lengths.

We compute "data to clock" jitter as

$$AFMJitter = \left(\frac{SD_w}{\sqrt{2}}\right) \times \left(\frac{100}{CBL}\right) \qquad (7)$$

where CBL = Channel bit length (see below).

The divisor √2 converts "data to data" jitter to "data to clock". It comes from normal error propagation based on the following ideas:

- We assume that, when measured relative to the clock, the leading edge and trailing edge positions ($Edge_1$ and $Edge_2$) vary normally and independently from one another and have equal variance. ($SD_1 = SD_2$).
- The normal rules of error propagation state that the variance of a sum or difference of two variables equals the sum of the variances of the two variables.

Applying these ideas, we have:

$$SD_w = SD(Length) = SD(Edge_1 - Edge_2) = \sqrt{\left(SD_1^2 + SD_2^2\right)} = \sqrt{2 \times SD_1^2}$$

which leads to (8)

$$SD_1 = \frac{SD_w}{\sqrt{2}}$$

Finally, the factor 100/CBL converts the length jitter to a percentage of the channel bit length, which is the basis for specifying jitter in the DVD standard.

Using the data from Table 2, the AFM Jitter is 3.09%. AFM Jitter is calculated for land lengths independently of bump lengths.

To summarize, AFM jitter is the standard deviation of edge placement relative to the underlying clock, expressed as a percentage of channel bit length.

4. RESULTS

Jitter.

Table 3 shows some of the best and worst jitter results we have seen in our laboratory.

Table 3. AFM Jitter computed from bump length variation in 2 discs

Disc	A	C
SD_w	17.8 nm	4.25 nm
Edge Jitter	12.6 nm	3.01 nm
AFM Jitter	9.5%	2.3%

Length Deviation and Definition of Length Offset.
The effects of bump length on electrical properties are so important that further analysis is worthwhile. Some electrical disc analyzers report deviation by comparing actual mean pulse durations with nominal durations. We now do the same, using physical lengths measured by AFM, tabulated as in Table 2.

$$Deviation = MeanLength - (T\# \times CBL) \quad (9)$$

For three bands on the power series stamper we observed a systematic change in deviation with laser power. See figure 3. At the lowest power, deviation ranged from about –80 nm to about –70 nm.

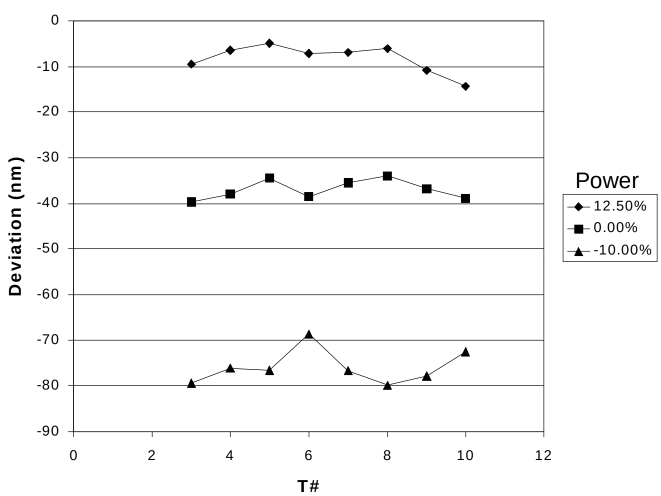

Figure 3. Length Deviation vs. T-number for a power series DVD stamper.

We introduce the parameter "Length Offset" to summarize the length deviations for all T-number groups. Length Offset is obtained from the graph shown in figure 4. Here we have plotted all the length measurements vs. the assigned T-number. We did a linear least squares fit to the data set. The slope of the line is the channel bit length. The intercept at T=0 is the Length Offset (by definition).

To summarize, Length Offset measures the average difference between all groups of bump or land lengths and their nominal lengths. Length Offset is found as the T=0 intercept of a plot of length vs. T-number.

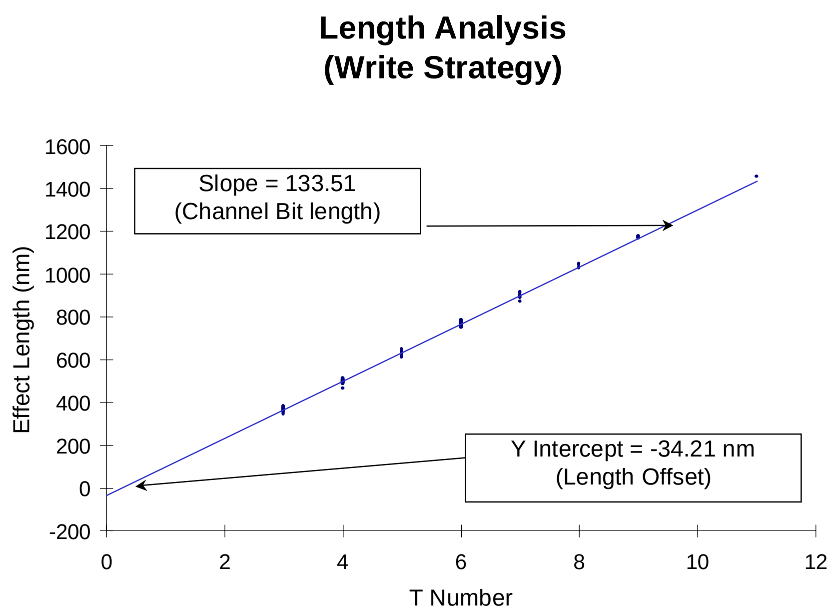

Figure 4. Length analysis of stamper bump lengths, showing the Channel Bit Length and Length Offset.

In a power series stamper, bump length increases with laser power. Likewise, length offset becomes more positive. See figure 5.

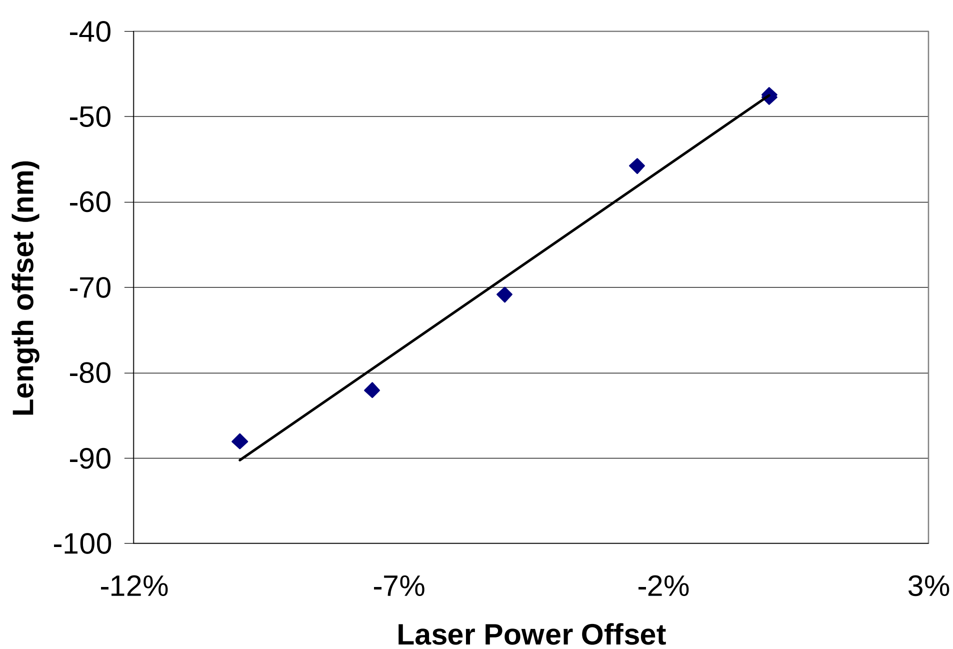

Figure 5. Graph of Length Offset vs. Laser Power Offset. Two points at power offset = 0% coincided almost exactly.

We also found a strong correlation of bump length offset with measured electrical asymmetry, as shown in figure 6. The linear fit had r^2 = 0.994, the slope was 0.248% Asymmetry/nm, and the standard deviation of the fit was 0.38%.

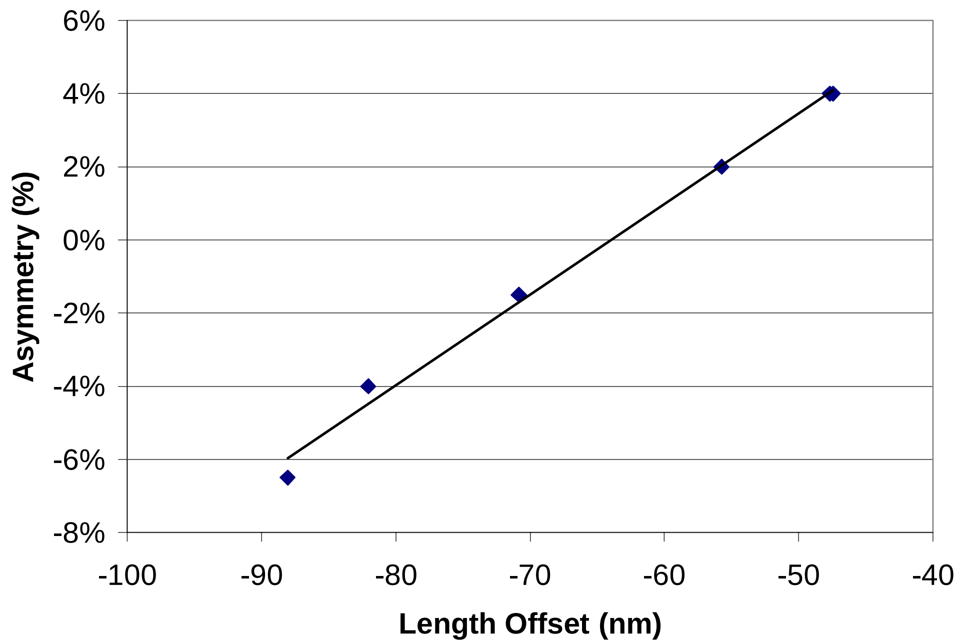

Figure 6. Graph of Asymmetry vs. Length Offset for a power series stamper. Two points at Asymmetry = 4% coincided almost exactly.

Table 4 summarizes some comparisons of AFM and electrical measurements. For a wide range of laser power (from –10% to +12.5%), the AFM jitter and the channel bit length did not change significantly. The AFM jitter for both bumps and lands was in the range 3.0% to 3.5%. The channel bit length was in the range 132.6 to 135.1 nm. In contrast, the length offsets (for both bumps and lands) and the mean T3 width did change significantly. As power decreased, the bump lengths and widths decreased by 68 and 62 nm, respectively, and the land lengths increased by 63 nm. Looking now at the electrical characteristics, we see that asymmetry decreased monotonically (by 24%) as the laser power decreased and the jitter decreased to a minimum and then increased.

Table 4. Microstructure and electrical measurements for a power series stamper.

		Bumps			**Lands**		
	Laser Power Offset	12.5%	0.0%	-10.0%	12.5%	0.0%	-10.0%
AFM + Length Analysis	AFM Jitter	3.20%	3.29%	3.34%	3.00%	3.35%	3.54%
	Channel Bit Length (nm)	132.94	133.05	132.63	132.66	133.26	135.05
	Length Offset (nm)	-7.09	-37.61	-74.91	1.90	34.20	65.32
	T3 Width (nm)	322.27	299.13	259.80			
Stamper Player	Jitter	10.10%	7.10%	15.20%			
	Asymmetry	17.90%	6.10%	-5.90%			

5. DISCUSSION

Why did we not find a correlation between AFM jitter and jitter measured by the stamper analyzer? We should not necessarily expect to see a correlation here. The reason for this is that the AFM is able to measure the position of feature edges without influence from optical effects such as crosstalk and inter-symbol interference. The stamper analyzer is sensitive to these effects and therefore reports a higher jitter value. In addition, it is likely that the range of asymmetry changes can exceed the capacity of the equalizer circuit to compensate for them. It is clear that for this power series stamper, 'bad' values of asymmetry (extremely high or low values) caused bad jitter (high values). The value of the AFM jitter results is that they prove that the precision of edge placement during mastering was not affected by the laser power.

How can AFM results be used to correct a jitter problem? A disc analyzer can only provide a jitter number, which alerts you to the existence of a problem. In contrast, the statistical analysis of AFM results separates this complex problem into several component causes, which are treated separately:

- Edge placement variation (usually a result of mastering) is reported as AFM jitter. We suggest that "AFM jitter < 4%" for a DVD stamper may be a useful internal specification in a disc manufacturing plant.
- Electrical asymmetry correlates with the AFM measurement of "Length Offset". AFM provides a very early warning if one measures the master.
- Length bias is detected by the AFM measurement of channel bit length.
- Other causes involve pit geometry and track pitch variation. In these cases, it is a good idea to compare the microstructure of good and bad stampers.

6. CONCLUSIONS

Statistical analysis of AFM images yields important information about optical disc microstructure and in turn provides information about the performance of the manufacturing process. Some statistics describe process position and others describe process noise. We have given a tutorial presentation of several statistical tools, beginning with descriptive statistics and continuing with analysis of variance and linear regression (curve fitting). It is very useful to analyze geometric parameters by considering the fundamental length groups of the data marks. We define "AFM jitter" and "Length Offset" as new parameters and compare actual results with corresponding electrical measurements, jitter and asymmetry. We describe how to use several distinct AFM results to investigate the many factors that can affect jitter.

7. REFERENCES

[1] Donald A. Chernoff and David L. Burkhead, "Automated, high precision measurement of critical dimensions using the Atomic Force Microscope", J. Vac. Sci. Technol. A 17, 1457 (1999).

[2] G. Barrie Wetherill, *Elementary Statistical Methods*, (Methuen & Co., London, 1967), pp.263 ff.