

# Full-field feature profile models in process control

Terrence E. Zavecz\* – TEA Systems

## ABSTRACT

Most process window analysis applications are capable of deriving the functional focus-dose workspace available to any set of device specifications. Previous work in this area has concentrated on calculating the superpositioned optimum operating points of various combinations of feature orientations or feature types. These studies invariably result in an average performance calculation that is biased by the impact of the substrate, reticle and exposure tool contributed perturbations.

Many SEM's and optical metrology tools now provide full-feature profile information for multiple points in the exposure field. The inclusion of field spatial information into the process window analysis results in a calculation of greater accuracy and process understanding because now the capabilities of each exposure tool can be individually modeled and optimized. Such an analysis provides the added benefit that after the exposure tool is characterized, it's process perturbations can be removed from the analysis to provide greater understanding of the true process performance.

Process window variables are shown to vary significantly across the exposure field of the scanner. Evaluating the depth-of-focus and optimum focus-dose at each point in the exposure field yields additional information on the imaging response of the reticle and scan-linearity of the exposure tool's reticle stage. The optimal focus response of the reticle is then removed from a full wafer exposure and the results are modeled to obtain a true process response and performance.

**Keywords:** Lens, aberration, perturbation, process control, SEM, Scatterometry, scanner, semiconductor, model, wafer, process window, profile, Side-Wall-Angle, SWA, BCD, lithography, metrology, Weir PW, BARC, Resist, Slit, k1, Duty Cycle

## 1. INTRODUCTION

The concept of deriving lithographic focus and exposure-dose settings using process windows has been used in the semiconductor industry for many years with the first commercial analysis package arising in 1989[1]. After over a decade, papers still appear with new work implementing or evolving the original methodology. Most recently, Mack & Byers presented a derivation of the algorithm from first principal interactions between one-dimensional aerial images with thin, infinite contrast photoresist using Taylor Series expansions to derive the original empirical construct.[2] The end-point derived relationship is presented in equation <1.

A feedback process control implementation of equation <1 was introduced by Ausschnitt and Cheng by first calibrating the coefficients to a selected isolated feature's top and bottom response. Subsequent sparse sampling of the features in the lot during the production cycle allowed the inverse model to be solved for localized exposure and focus.[3]

Significant work continues in this area with other publications directed toward techniques for reducing the subjective nature involved in selecting the proper coefficients or weighting strategies to include in the fit or by pre-conditioning the equations with simulations.[4]

One of the more recent renditions of the process window algorithm consists of an expansion in focus and dose of the form:

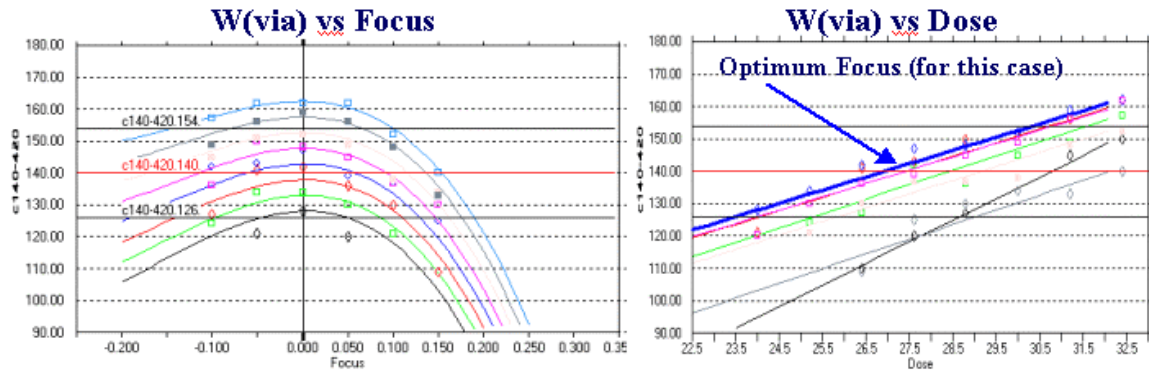
$$W = \sum_{m=0}^M \sum_{n=0}^N a_{nm} \left( 1 - \frac{E_s}{E} \right)^n F^m \quad <1$$

---

\* Contact: [tzavecz@TEAsystems.com](mailto:tzavecz@TEAsystems.com) phone: (+01) 610 682 4146, <http://www.TEAsystems.com>

Where  $W$  is the feature response or Critical Dimension (CD),  $E$  the exposure-dose,  $E_s$  is the dose needed to obtain the desired feature width and  $F$  is the defocus from optimum focus.

The need for intimately involving the skills of the engineer in the application process can be seen from the graphs of figure 1. The focus and dosage relationships of 140 nm Vias are displayed as separate graphs in the figure. The left graph highlights the need to weight and select the surface coefficients of equation <1 in order to best-fit the individual responses of each dose-curve to the measured data. Notice the misfits of the lower curves to the measured data points. The true reason for these misfits will be shown in the last section of this paper.

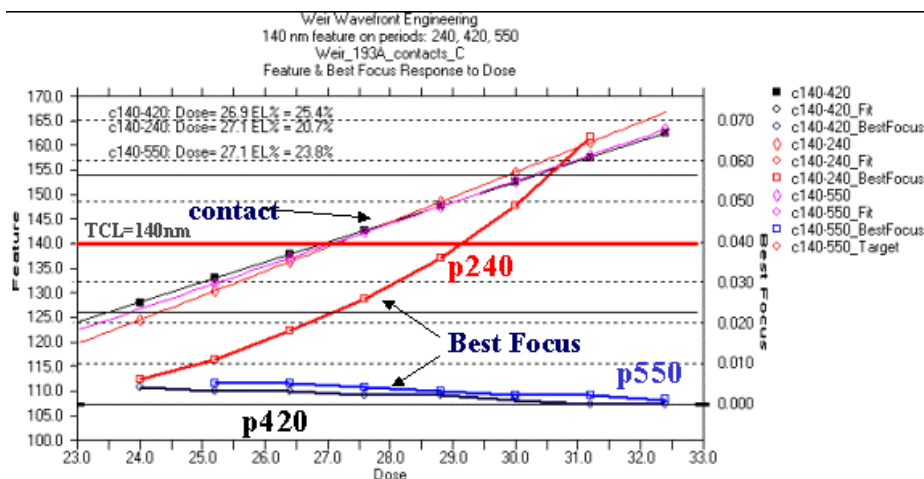


**Figure 1: Process window sub-plots of response to focus and dose**

The right graph of  $W$  vs Dose is less often observed because of the profusion of conflicting information generated by the varied curve responses of the feature at each base defocus value. The line indicated in the graph is the response of the nearest measurement set to best focus for this data and is therefore the most accurate response curve in the presented data. Optimally, the curve corresponding to the measured feature's response to dose from measurements taken exactly at Best Focus should be the curve used in the interpretation but in practice the data is rarely collected exactly at optimum focus values and the selection of nearest focus curve to this value not obvious.

The need for user-interactions in the process window surface arises from other feature response perturbation sources not addressed by the two variable surface of equation <1. These "other" sources were recently listed and modeled by Dusa et al.[5] Many of the feature response perturbations listed by Dusa are removed by the averaging that occurs when the process window is modeled to the data. However errors encountered in reticle fabrication, lens-slit aberrations and reticle-scan errors will result in deviations of the process window base fit to the raw data set.

Even with the need for subjective tuning, the process window algorithm has withstood the test of time and continues in general use. Yet, most users are unaware of the additional information that can be extracted from the analysis beyond that of focus, dose, exposure-latitude (EL) and depth-of-focus (DoF). The next section will discuss some of these



**Figure 2: Feature response for best focus (left ordinate), and best focus (right ordinate) as a function of dose**

extensions and show how they can be used to even greater advantage by incorporating full-profile metrology elements into the analysis.

## 2. THE BEST FOCUS FEATURE CONTOUR

### 2.1 Best Focus and optimum feature size determination

In the general nature of the application of equation <1, most feature size responses to focus are near quadratic in form. The observation of empirical data for features such as widths, side-wall-angle (SWA), film thickness and even line-edge roughness confirms this the validity of this form over the ranges used in production and process setup. Quadratic response curves lend themselves to a simple method of determining the optimum focus for each dose as the first order derivative of the response with respect to focus or:

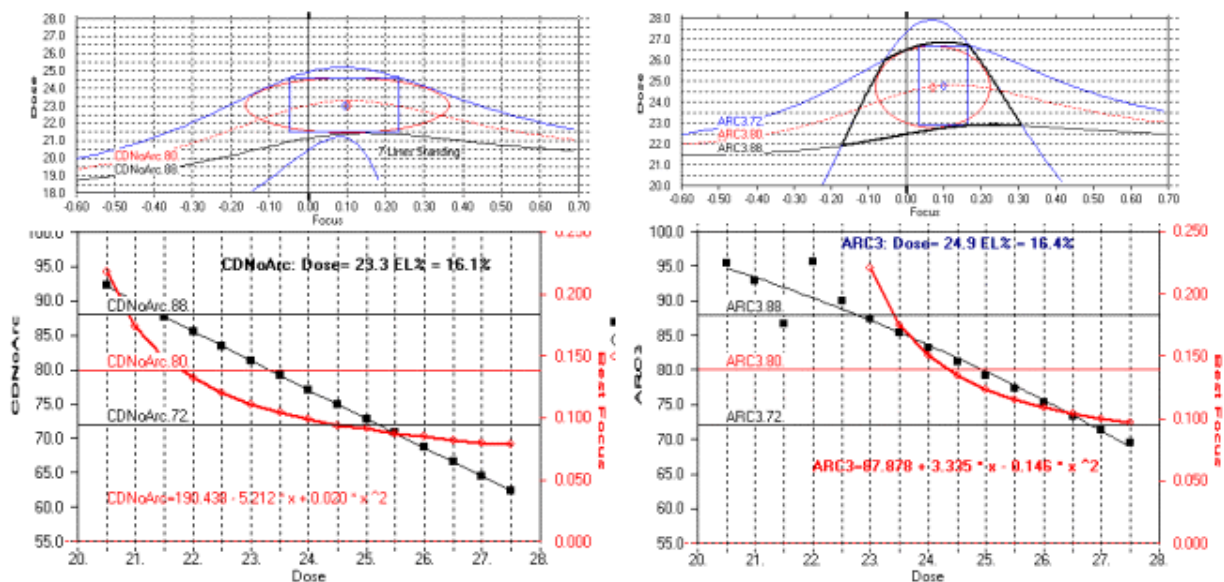
$$\frac{\partial W}{\partial F} = \sum_{m=1}^M \sum_{n=0}^N m * a_{nm} \left( 1 - \frac{E_s}{E} \right)^n F^{m-1} = 0 \quad <2$$

By setting equation <2 to zero and solving for focus the point at which the feature's response to focus changes is maximized and this is defined as the Best Focus for the dose family. Since the dose and best focus are now known, the optimal feature size at best focus and the dose-localized Depth of Focus (DoF) can be accurately calculated.

At this point in the analysis sequence we have determined the optimum focus for each dose. The analysis can now plot the feature response at best focus to dose and obtain meaningful results as will be seen in the discussion of figure 2. The presentation now turns to a discussion of some of the observed characteristics of this extended analysis.

### Duty Cycle Study

The plots shown in figure 2 represent the response of 140 nm contacts (Vias) in a 193 nm resist process. Three duty cycles were studied with periods of 240,420 and 550 nm. The contact response to dose can be seen in the three straight -1<sup>st</sup> order-- lines whose feature sizes correlate to the left ordinate of the graph. Observe the clean response to dose in these curves. The change in slope from one data set to the next reflects only the proximity issues engendered by the duty-cycle sensitivity of the patterns.



**Figure 3: BARC response comparison for 80 nm lines in a 193 nm resist (see text for explanation)**  
The Left Half of the figure contains the response graphs for a wafer with No ARC coating.  
The Right Half contains plots for a substrate with an “ARC3” film.

The Best Focus for each dose is plotted in the three curves located at the bottom of the figure and correlated with the right-hand ordinate of the graph. The two high period curves, 420 and 550 nm respectively, respond as expected with little change in focus across the dose range. Ideally, an aberration free exposure without mechanical distortions will experience little or no change in focus with exposure. However the 240 nm period focus response curve indicate the increase in assymetry in the focus response of the features and the onset of severe proximity effects in the imaging.

### BARC Performance Comparison

Wafers with and without a Bottom Anti-Reflective Coating (BARC) were studied for the plots of figure 3. The process window graphs, at the top of the figure, represent Critical Dimension sizes measured for 80 nm, 1:1 lines. Each graph also plots the variable response representing the focus and dose range for the number of fully-standing or “collapsed” line structures. During exposures that are near the imaging limits of the process, lines may be resolved during imaging only to collapse during subsequent processing because of chemical forces or profile undercut. Standing waves from substrate interface reflections are a major cause of line collapse.

If a nest of twelve lines-and-spaces, located near the edge of the process window, is observed under high magnification, some or all may be found leaning against their neighbors. However, in cases where the entire structure is not collapsed, an automated CD-SEM’s targeting software will seek out and measure features matching the target profile. As a result valid measurements are gathered from structures outside of the true process window.

Upon first examination of the process window graphs located in the upper half of figure 3, there is not much difference between the two response contours for CD. However, to incorporate the information of the standing/collapsed structures, each measured image was manually inspected and the number of free-standing lines for each focus/dose was reported and is graphed in the second set of curves on the figure.

The “No Arc” plot on the left shows that the collapsed-line surface does not intersect and share a common process window with the 80 nm target specification for the process. However, the “ARC3” component on the right does show a sizeable window that, in turn, can be used for process setup.

The CD vs Dose plots in the lower section present additional the BARC characteristics. For the “No-BARC” plot, lower-left quadrant of figure 3, we see a well-behaved CD vs dose response curve with the Best Focus curve changing slowly until dose values are lowered below the target dosage of 23.5 mJ/cm<sup>2</sup>.

The “ARC3” curve on the right side of figure 3 shows that, as expected, an increased dosage is needed to expose the BARC coated wafers to size. The CD vs Dose curve slope of the ARC3 substrate however has a lower slope suggesting a more stable process than that of the no-ARC wafers. It’s interesting to note the characteristic instability of CD size seen for the very-low dose values indicating a limited bias range for the film structure.

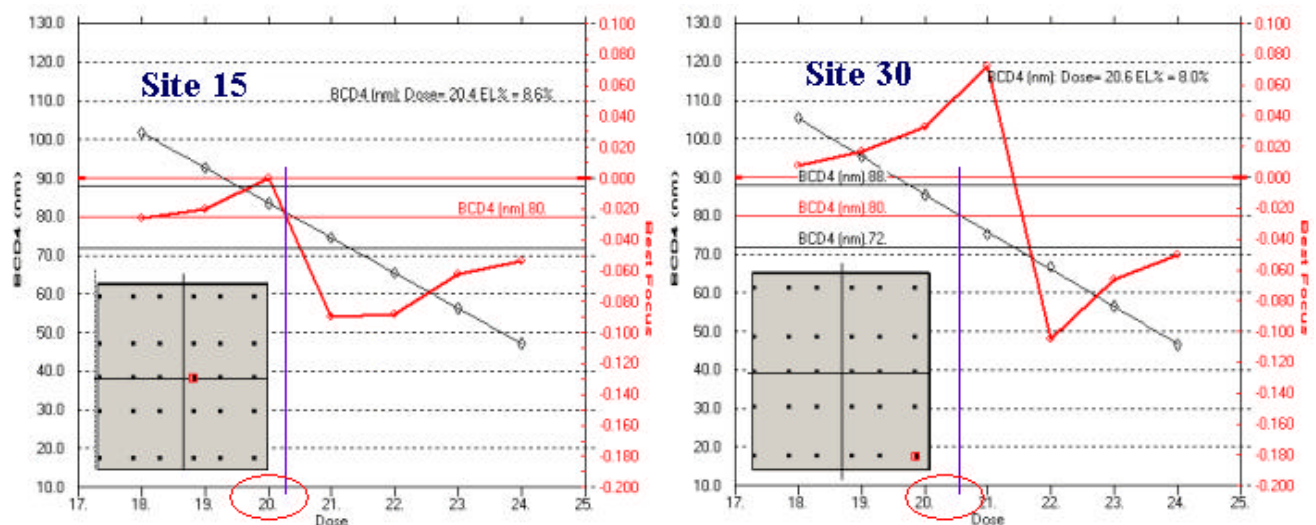


Figure 4: Analysis of 80 nm, 1:1 features in 193 nm resist for two sites on a 6x5 field (inset)

Best focus for the ARC is very sensitive to dose and rapidly increases in value with exposures below the target dosage again showing a limitation in the process's ability to be biased to higher feature sizes.

### High Resolution Applications

The process's resolution capabilities are being challenged in this example as 80 nm features on 1:1 duty cycles are resolved. The left-side of figure 4 shows the dose response measured on a Scatterometer for site #15, located near the center of the lens and scan as shown in the field inset. A Best Focus calculation always loses some accuracy near the isofocal dose, at about 20.5 mJ/cm<sup>2</sup> in this example, because of the lack of curvature in the CD vs Focus curve. However, what is significant is the offset in best focus seen before and after the isofocal dose. This is a common characteristic in systems that operate near the resolution limits and is an early indicator of the onset of significant response to aberrations and other feature response perturbations such as flare and scatter.

A measured feature near the extreme corner of the exposure, Site 30 as plotted in the right-hand graph of figure 4, shows an even greater best focus offset for doses before and after the isofocal position. Note also that the isofocal dose has changed for this site and that the Best Focus response slope to dose is greater than for sites located near the less-stressed imaging at the field center.

The response curves for site 30, right side of figure 4, show an offset in the dose needed to achieve line size. This can be the result of lens aberrations but in this analysis it is not the caused by local focal plane defocus because the procedure plots response at best focus. Studies have shown that such offsets are most frequently due to differences in the size of features on the reticle and suggest a method for which reticle sizes could be inferred.[6] We will address this behavior in the next section.

## 3. FULL-FIELD PROCESS WINDOW

### 3.1 Optimum Field Response

Process windows can be viewed as spatial contour plots. The plot on the left side of figure 5 presents data taken from a CD-SEM measured Side-Wall-Angle (SWA) of 80 nm features, one field per exposure, on a focus-dose matrix.

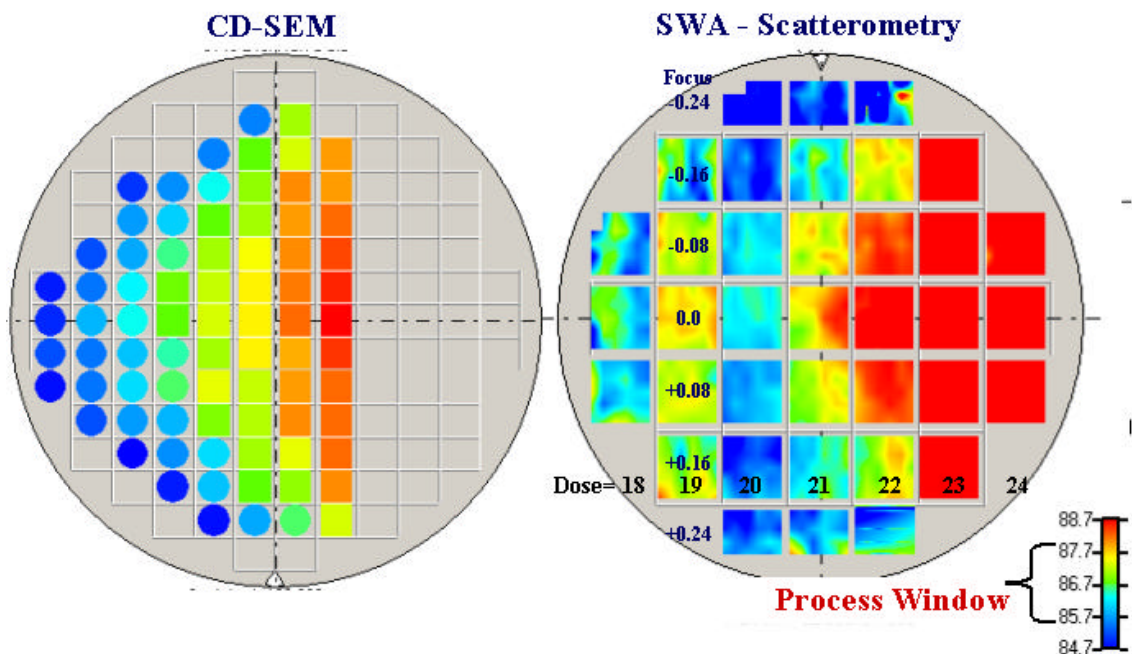


Figure 5: Spatial contour plot of Process Windows from CD-SEM (Left) data and Side-Wall-Angle (Right).

Contrast this with the ellipsometer measured SWA plot on the right side of figure 5.

SWA readings here were taken from 30 sites per exposure in a focus-dose matrix. The focus values changed by row, as indicated in the figure on the right, and by dose for each column. The SWA feature size key was set so that the process window's upper and lower control limits covered the color range from light-blue to light-red as indicated.

From first observations it can be seen that there is much more tool and process information to be seen when the whole exposure field is examined. This methodology allows a rapid visualization of the extent and particulars of the field and the process window sustained by the process. The display also suggests that single-point field analyses of process windows can be significantly biased by improper site selection during exposure. This is especially true in scanners today where the slit is located across an optimized region of the lens and exposure field center may not correspond to a site with minimum image perturbation.

### The Optimal Full-Field Image - validation

Equation <1 can be redefined to address feature response perturbations across by full exposure field as:

$$FR(x, y) = \sum_{m=0}^M \sum_{n=0}^N a_{nm}(x, y) \left(1 - \frac{E_s}{E}\right)^n F^m \quad <3$$

Where FR(x,y) now represents the response for a feature located at some site (x,y) on the reticle image. Using this construct a matrix of process window response curves are created demonstrating the full-field response of the process. Best Focus, Dose response and the optimal feature size across the exposure are then calculated using the procedures described in the previous section.

A scanner reticle with 100 nm 1:1, final size reticle features was exposed in a 193 nm litho process using a limited range focus and dose matrix. The wafers were measured using a normal incidence commercial ellipsometer and analyzed with the Weir PW software.[6] The resulting optimum bottom-feature-width responses for 121 (11x11) sites across the exposure fields for uniformity at each dose are display in table 1.

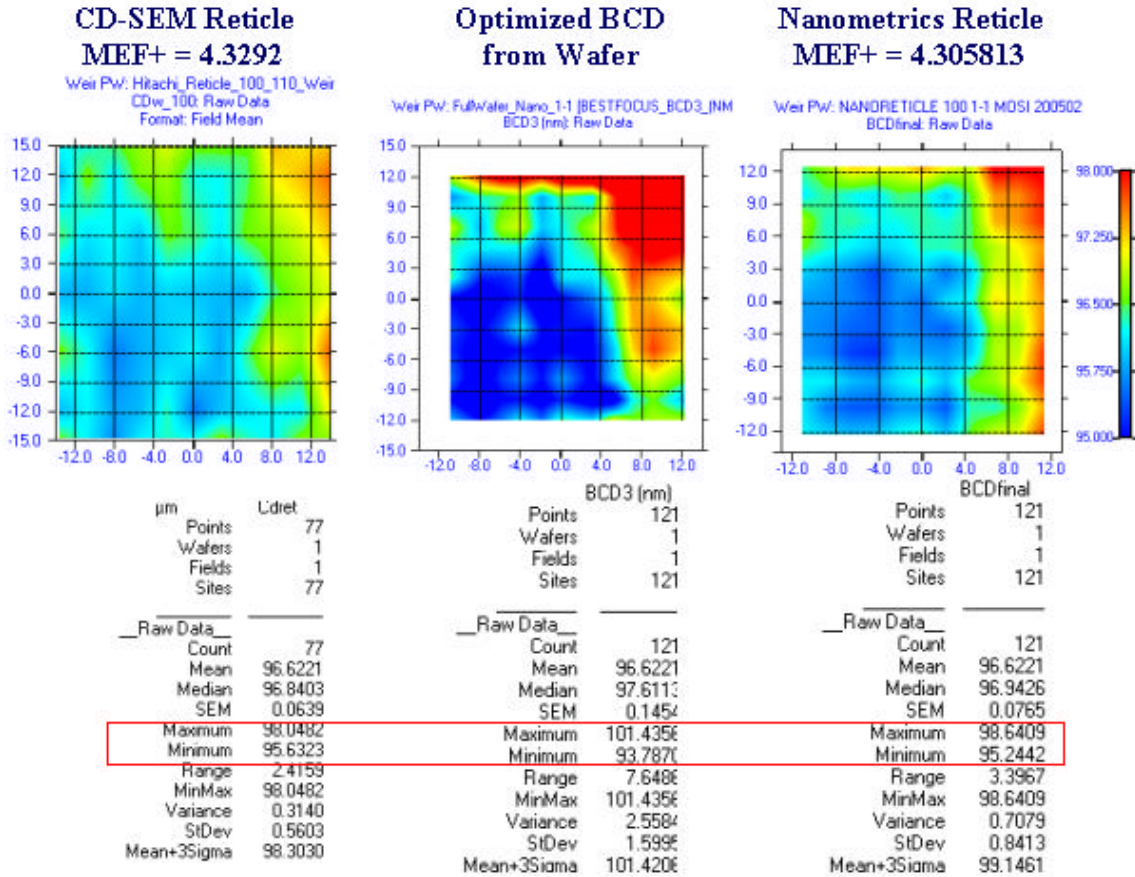
The table summarizes the mean and intra-field-deviation (IFD) for the optimum field obtained at each dose. Depth-of-

	Best Focus	BCD4 (nm)	DoF	Dose		Best Focus	BCD4 (nm)	DoF	Dose
Mean:	0.0070	104.1780		18.0	Mean:	-0.0390	65.3970	0.0160	22.0
Max:	0.0810	110.6640		18.0	Max:	0.2880	70.2900	0.0210	22.0
Min:	-0.1570	100.6540		18.0	Min:	-0.2800	61.7610	0.0110	22.0
IFD:	0.2380	10.0100		18.0	IFD:	0.5680	8.5290	0.0100	22.0
Mean:	0.0180	94.3450	0.0530	19.0	Mean:	-0.0840	56.2490		23.0
Max:	0.0830	99.8950	0.0530	19.0	Max:	-0.0320	59.1160		23.0
Min:	-0.1340	91.0560	0.0530	19.0	Min:	-0.1570	53.4410		23.0
IFD:	0.2170	8.8390	0.0000	19.0	IFD:	0.1250	5.6740		23.0
Mean:	0.0330	84.4540	0.3560	20.0	Mean:	-0.0570	46.2440		24.0
Max:	0.0940	88.9940	0.5970	20.0	Max:	0.0210	47.9520		24.0
Min:	-0.0740	81.4100	0.1710	20.0	Min:	-0.1340	43.7280		24.0
IFD:	0.1680	7.5840	0.4260	20.0	IFD:	0.1550	4.2230		24.0
Mean:	0.1080	72.8100	0.6260	21.0					
Max:	0.9820	77.6260	0.6400	21.0					
Min:	-0.1050	30.3330	0.5230	21.0					
IFD:	1.0870	47.2940	0.1170	21.0					

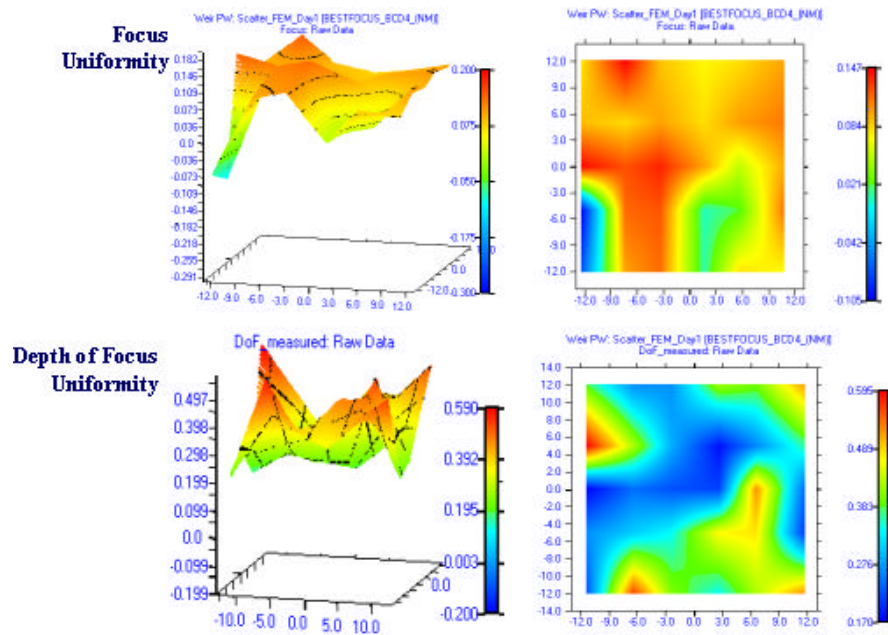
**Table 1: Full-field summary for Bottom Critical Dimension (BCD) optimum feature calculations**

Focus (DOF) values are calculated for those exposures where the calculated sizes fell into the 80nm +/-10% target specification. The IFD dose-response curve is not monotonic because of the film interactions with the aerial image during extreme image-averaging at higher dosage.

Taking the Optimum BCD value at each site, the full-field response contour can be plotted for each dose. Of greatest



**Figure 6: Signatures of Optimized BCD (on wafer) and reticle measurements made by CD-SEM and Scatterometry**



**Figure 7: Optimal field derived focus error and DoF for the exposure system**

### Additional information extracted

In calculating the optimal image the analysis also derived the defocus and DoF for each point in the exposure. These contour surfaces are next plotted in figure 7. The defocus calculation resulting from this method is not as accurate as the specialty patterns common in the industry, such as the Phase Shift Focus Monitor, however defocus estimates typically present an uncertainty of about 30 nm 3-sigma.

In addition, since the Depth of focus is calculated at each point's optimum focus in the exposure field, the calculations typically are better than those that will be experienced full-field during production. However, this is an improved estimate of the lens and film interactions and is therefore a better estimator for film response tests as encountered when setting thickness for Anti-Reflective coatings.

## 4. APPLICATION TO FULL-WAFER RESPONSES

### 4.1 Observations of the Raw Data

The full-wafer contours of figure 8 were taken from the same experiment described in the previous section. Each image contains 121 points per field. Some relationships, such as those between BARC and SWA as well as BCD and TCD are already visible.

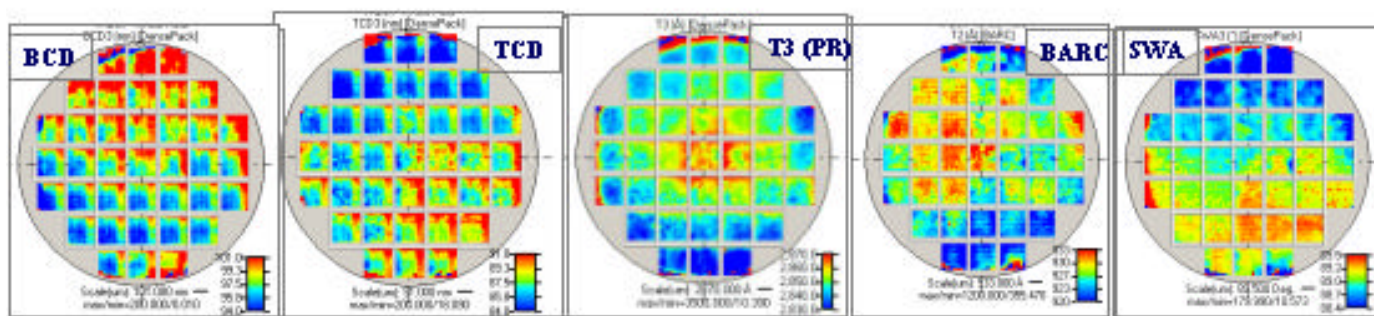


Figure 8: Raw data contour plots from one data set. Notice the characteristic BCD signature also seen in figure 6

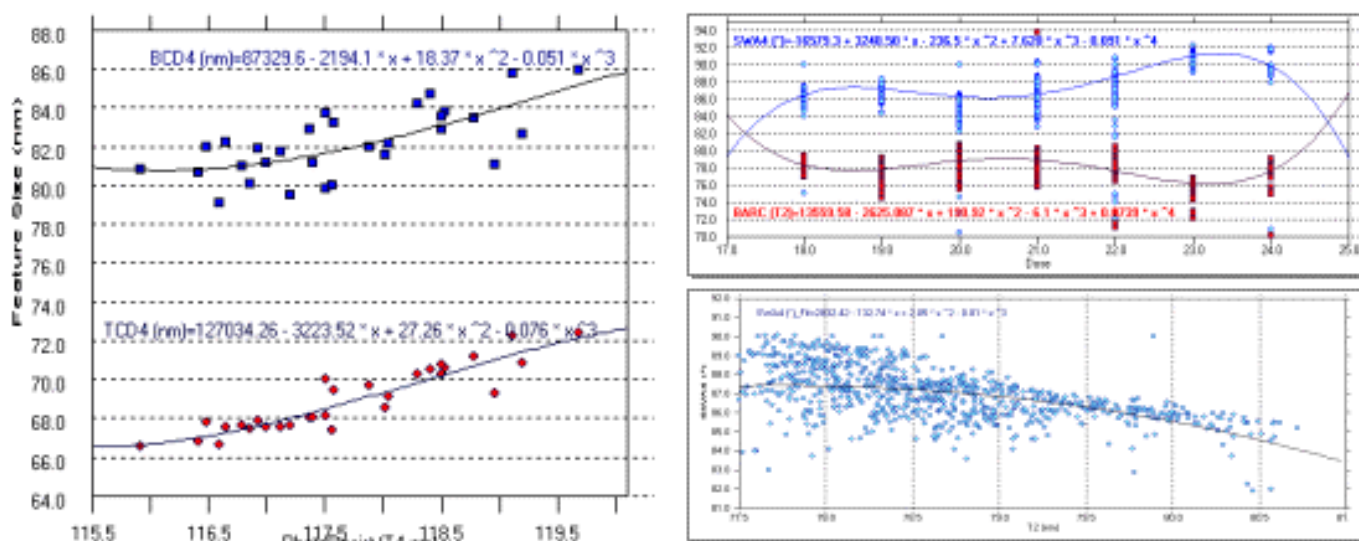


Figure 9: Left side: BCD and Top CD (TCD) variation with Photoresist thickness  
Right Side: Normal wafer variation of SWA and BARC thickness

The process variable relations are now easily derived in the near-swing curve variance seen for BCD and TCD value response to measured thickness of the photoresist. Recall that this wafer was exposed with a focus and dose range that is no more than typically seen in production. Feature size swing-curve response is apparent in both feature widths as plotted on the left side of figure 9.

The BARC thickness and measured SWA distribution across the field is plotting in the upper right graph of the figure as a function of the exposure dose. The same data set is again plotted in the graph below it to show the direct relationship between SWA and BARC thickness.

The slowly-changing full wafer variations of the process variables can be measured and observed. These larger perturbations will have to be systematically modeled and removed in order to see the finer variations inherent in the IntraField exposure .

## 4.2 Formalization

In a previously cited reference, Dusa et al presented a formalization for describing the feature response to process perturbations across a full wafer as:[7]

$$FR_a(x, y) = IF_p(x, y) + W_p(x, y) + DD(x, y) + r \quad <4$$

In equation <4, the feature response( $FR_a$ ) at any location on the wafer is a function of the slowly changing wafer signature  $W_p$ , the die-to-die variability (DD) and the IntraField perturbations ( $IF_p$ ).

A scanner image in photoresist is generated by two independent systems, the lens slit and the reticle scan-stage. Perturbations introduced by the slit consist primarily of those from the lens and any scatter introduced by the slit-aperture edges. Since the slit covers approximately an 11 mm height on the field, lens perturbations will be replicated on each row of the image.

The reticle scanning stage moves the object-reticle across the slit to exposure the entire height of the exposure field. Perturbations here consist of focus changes resulting from reticle scan-stage travel – pitch, yaw etc – and dose changes that result from small variations in the velocity of the scan-slit.

The IntraField perturbation can therefore be described by:

$$IF_p(x, y) = IF_{Slit}(x) + IF_{Scan}(y) + r \quad <5$$

$$IF_p(x, y) = \sum_{Rows} \sum_{n=0}^4 a_n x^n + \sum_{Columns} \sum_{j=0}^4 a_j y^j + r \quad <6$$

Where  $r$  represents the random residual contributions. The individual contributions of slit and scan are adequately described by a simple 4<sup>th</sup> order polynomial because of the one-dimensional nature of the analysis. The coefficients in turn describe the physical response of the slit and scan at the X-column and Y-row locations on the field with the  $a_0$  term representing the offset,  $a_1$  the tilt etc.

## 4.3 Addressing IntraField systematic errors

IntraField errors can be systematic functions of the lens, scan and optomechanical assemblies of the exposure tool. By applying equations 6 to a focus matrix, the influence of the major perturbations from film variations are modeled and removed from the analysis to provide the field sensitive data shown in figure 10.

The scanner data in figure 10 consists of a focus matrix that has been set up to expose alternate fields in opposite scan directions. The “+” nomenclature indicates an “Up” scan of the reticle stage and, conversely, the “-” nomenclature is a “Down” scan field. In this manner adjacent fields can be compared not only for focus offsets but also for scan-direction behavior.

One last very important item about the data is the fact that the data plots on the right and left were exposed at the same dose and focus values. The left side of figures 10 and 11 contains lens data from after the lens was in continuous production for an extended period. The right-side of the figures exhibits contours from the same lens after cleaning. Lens cleaning consisted of the removal of built-up of material on the optic surface just above the wafer which was removed by gently wiping with lens tissue soaked with water and then is wiped again with methanol.

Equation 6 was applied to the data to model every row of every field for the  $IF_{\text{slit}}$  data and then every column of every field for the  $IF_{\text{scan}}$ . The  $IF_{\text{scan}}$  coefficients were then averaged for each field and the averaged scan model subtracted. The plot is essentially the residuals to the  $IF_{\text{scan}}$  model. This display there contains only the lens-slit perturbations to the features where even the defocus errors of the matrix have been subtracted by removal of the  $a_0$  offset term of the model.

Figure 10 contains simultaneous contour plots of the len's perturbations to the Top, Bottom and side-wall-angle (Slope or SWA) components of the profile. The apparent offset in BCD width between up and down scans is apparent in the

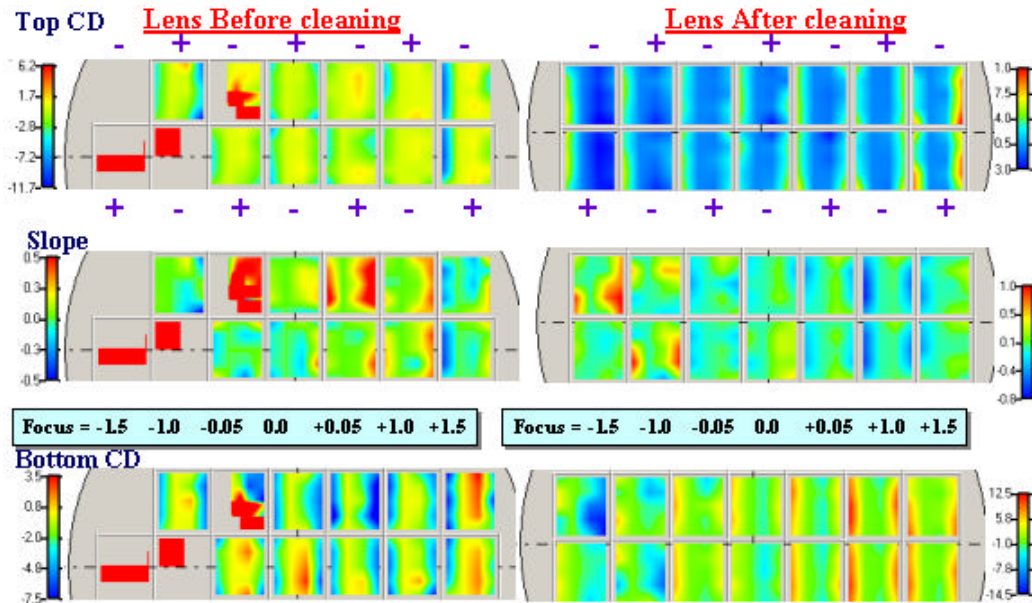


Figure 10: Lens degradation study showing imaging response thru focus for the lens slit

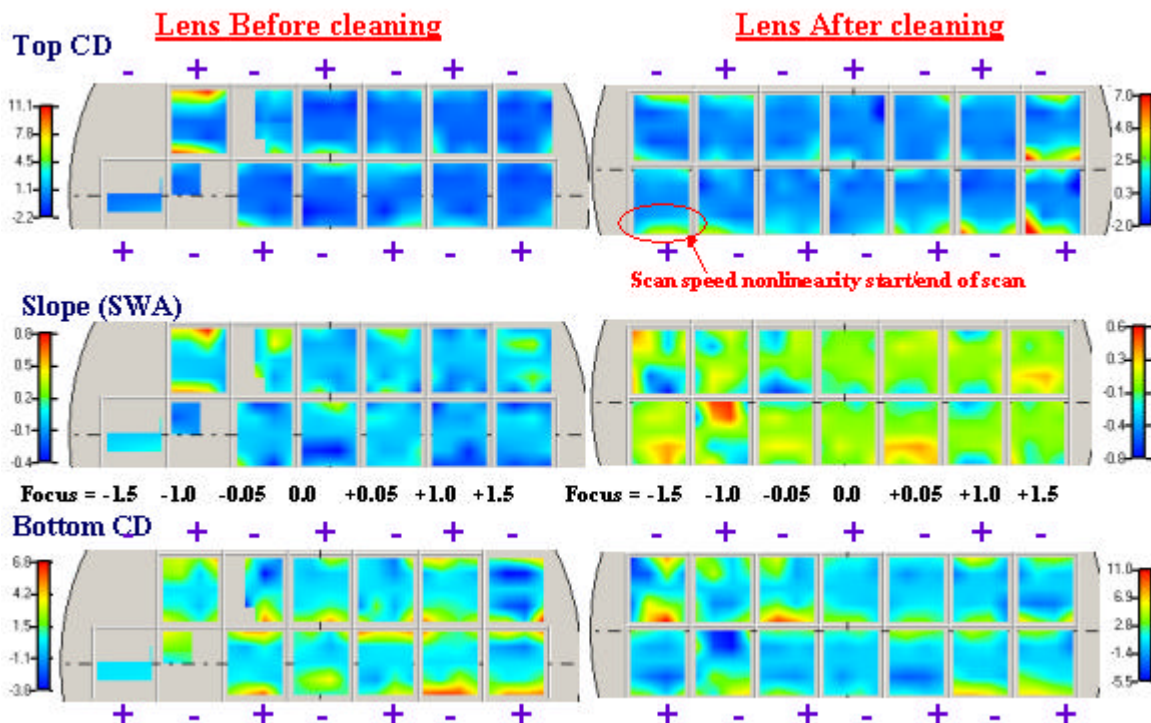


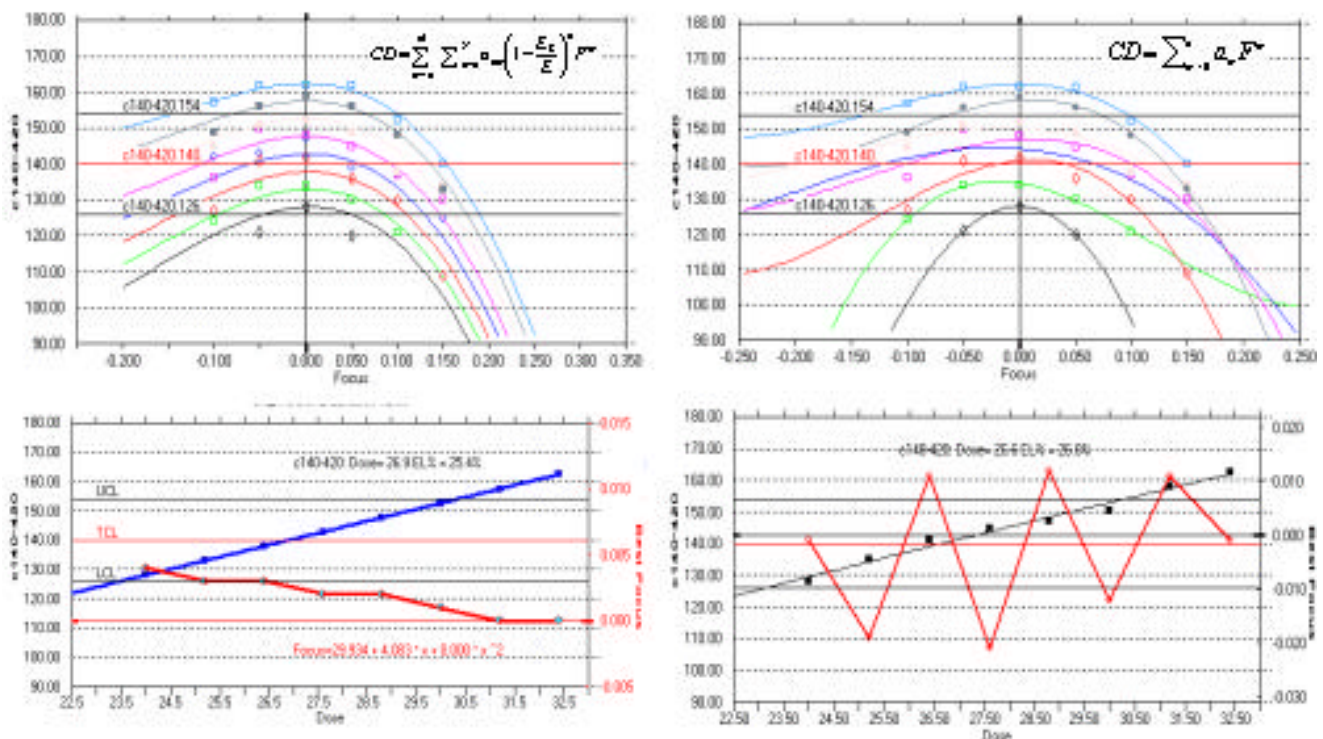
Figure 11: Lens contour plots showing image response for the reticle-scan-stage.

BCD measurement set. Also notice the apparent difference in slit-bow, or 2<sup>nd</sup> order coefficient, behavior particularly around the +0.05  $\mu\text{m}$  defocus field. The reticle may in fact be flexing as the scan reverses however the true value of this may be evident in the contour plots of figure 11.

For the analysis shown in figure 11 the  $IF_{\text{slit}}$  modeled data was removed and the residuals now show individual scan behavior. The effects of the cleaning on the process window are obvious with many pre-clean fields not even capable of being measured below  $-0.05 \mu\text{m}$  defocus. However also notice the increase in BCD or TCD size at the end of every scan. This is the result of scan velocity errors that occur when the slit begins accelerating to move to the next exposure field location, this is also the most likely contributor to the apparent lens-slit bow changes at the ends of the fields in figure 10.

## 5. FINAL NOTE ON PROCESS WINDOW FIT

Back in the discussion of figure 1, the topics turned to the subjective nature of fitting the process window. The curves of figure 1 are repeated on the left side of figure 12 along with the equation (i.e. <1) that was used for the fit. Notice that the data points at the bottom of the CD vs Focus curve do not quite match the surface.



**Figure 12: Process window misfit due to reticle-scan direction artifacts**

Contrast this behavior with the curves on the right side of the figure that fit the data to a simple polynomial in Focus (F). The fit is better in that every data point now has a curve running directly through it however the best focus response to dose shown on the lower right graphic is a saw-tooth shape. This data, different from that of figure 11, was exposed on a scanner and SEM measured. Similar to the previous sections discussion, the scanner alternated scan direction between each scan. A closer observation of the Best Focus curve on the left graphic also shows the same behavior although it has been strongly dampened by the process window algorithm used for the fit.

## 6. CONCLUSIONS

The current state of development of the process window algorithm was reviewed and a method of extending the analysis to examine dose response in a meaningful fashion presented. Along with the new graphic, the Best Focus verses Dose plot can be created and several examples were shown of it's application to BARC resist analysis and the investigation of

the onset of lens aberration dominance.

The methodology was then extended to examine full-field process window response for feature responses now measured by scatterometry and ellipsometric methods. The full-field analysis was shown to be traceable to direct reticle measurements with range errors on the order of 6 nm or less. These range errors were then shown to be the result of lens and scan-system distortions.

An IntraField feature perturbation model was next presented that can separately account for lens-slit and reticle stage-scan errors. Examples were shown for applying the models to lens imaging quality and slit scan-spped analyses of production systems.

## **7. ACKNOWLEDGEMENTS**

As a final note I want to thank Mircea Dusa of ASML and Raymond Hoobler of Nanometrics for their help and support on some of these projects. I want to be quick to mention that the aberrations and errors shown did not necessarily reflect product samples from equipment sold by their respective companies. The exception to this disclaimer being the wafer-to-reticle measurement comparison data that was setup by Raymond and the support and critique given to the project by Mircea.

## **REFERENCES**

1. C. Ausschnitt, "Rapid Optimization of the lithographic process window", SPIE vol 1088 (1989) pp. 115-123
2. C.Mack, J. Byers, "Improved model for focus-exposure data analysis", Proc of SPIE (2003) 5038-39, pp396-405
3. C. Ausschnitt, S. Cheng, "Modeling for Profile-Based Process-Window Metrology", SPIE (2004) vol. 5378, pp 38 – 47
4. W. Zhou, J. Yu, J. Lo, J. Liu, "Lithography process window analysis with calibrated model", Proc of SPIE (2004) 5375-29, pp. 721-726
5. Mircea Dusa, Richard Moerman, Bhanwar Singh, Paul Friedberg, Ray Hoobler, Terrence Zavecz, "Intra-wafer CDU characterization to determine process and focus contributions based on Scatterometry Metrology", Proc. SPIE (2004), Vol. 5378-11
6. T. Zavecz, R. Hoobler, M. Dusa, "Models for reticle performance and comparison of direct measurement", Proc. of SPIE (2005) 5754 – 110.
7. op.cit. 5