Spatial Variance Spectrum Analysis and Its Application to Unsupervised Detection of Systematic Wafer Spatial Variations

Jakey Blue and Argon Chen*

Abstract — Investigation of wafer spatial variations is critical for semiconductor process/equipment optimization and circuit design. The objective of spatial variation study is to differentiate the systematic variation component from the random component. This is usually done by contrasting with a set of known systematic patterns based on engineering knowledge. However, there could exist unknown systematic components remaining in the unexplained residuals and overlooked by the conventional spatial variation study. In this paper, we develop a novel spatial variance spectrum (SVS) to analyze the systematic variations without any priori information of the systematic patterns. The SVS is a series of spatial variations over a range of spatial moving window sizes from the smallest spatial moving window consisting of only two metrology sites to the largest one covering all metrology sites of the entire wafer. The SVS can be used to characterize the wafer spatial variations and to detect existence of systematic variations by a proposed hypothesis test. We also propose an index to summarize from the SVS the systematic proportion of the spatial variation. The proposed test and index of systematic variations will be demonstrated and validated through both hypothetical examples and actual cases of wafer critical dimension (CD) metrology data.

Note to Practitioners — Wafer spatial variations study is critical to within-wafer control and optimization. Only through investigation of the spatial variations on the wafer surface do we get to learn more about the nature of the semiconductor process/equipment and the manufacturability of certain circuit design patterns. In this paper, we propose a novel analysis tool to help engineers characterize the wafer spatial variations and distinguish the systematic variation from the random variation. With the spatial systematic variations identified and described, the engineers can discover the opportunities to further improve the processing quality and thus the final yield.

Index Terms—spatial variation, wafer topography, systematic variation, random variation, CD metrology

Manuscript received. (Write the date on which you submitted your paper for review.)

Jakey Blue is with National Taiwan University, 1 Roosevelt Rd. Sec. 4, Taipei, 106, ROC (Taiwan). He is a Ph.D. candidate in the Department of Mechanical Engineering. (e-mail: d93522042@ntu.edu.tw)

*Argon Chen, corresponding author, is with National Taiwan University, 1 Roosevelt Rd. Sec. 4, Taipei, 106, ROC (Taiwan). He is now a professor in the Graduate Institute of Industrial Engineering and Department of Mechanical Engineering. (e-mail: achen@ntu.edu.tw).

I. INTRODUCTION

As the metrology technology continues to advance, more and more wafer data must be analyzed quickly and efficiently for monitoring and controlling the fabrication processes. The wafer spatial variations study is a key to yield enhancement especially when the fabrication technology enters the 32nm node. Therefore, the study of spatial variation becomes critical to both process control and circuit design [1], [2]. In doing this, the systematic and random components, both of which contribute to the spatial variations, must be identified before the root cause of yield loss can be found and removed.

Kibarian et al. [3] examine the spatial dependencies (referred to as spatial correlations) of the process parameters, such as polysilicon line width and film thickness, on circuit testing data. Mozumder and Lowenstein [4], and Guo and Sachs [5] model the within-wafer variation based on multiple response surface methods while Smith et al. [6] compare it with the single response surface methods. Boning and Chung [2] describe the concept of statistical metrology and the decomposition of spatial variations into wafer-to-wafer, die-to-die, site-to-site variations, and residuals. Stine et al. [1] characterize the wafer-level, die-level, and wafer-die spatial variations for correlation studies. The literatures extracting and characterizing the spatial variations can be mainly categorized into two groups [7]. In one group, the variation is separated into variance components by employing methods such as analysis of variance (ANOVA) or Fourier transform. The other group identifies the distinct exemplar-based variation patterns and analyzes the impact of these factors through proper decomposition of the wafer and yield data.

To characterize the wafer-level or die-level variance components, ANOVA methods are widely applied [8]-[10]. Significance of the components can be then ranked for further causal analysis. At the process level, Steele et al. [11] assume the total critical dimension (CD) variation to be the combination of independent variance components from coating, developing, and baking processes and use design of experiments to model and understand the CD uniformity issues. Yu et al., [12] employ the fast Fourier transform (FFT) to decompose the wafer CD’s spatial variations into wafer spatial patterns or residual variations. Ye et al. [13] and Han et al. [14] analyze the pattern generator-induced mask CD errors in the spatial frequency domain and identify error contributors using the Fourier transform. Ouyang et al. [15], [16] identify the amplitude
excursions in the spatial frequency domain of CD’s using the spatial Fourier transform (SFT). By applying an inverse SFT, the variance components can be separated and used to explain the systematic and random errors in spatial variations. The idea of transforming observations in the space domain to a spectrum in the frequency domain is very useful for understanding the natures of the spatial variations. In particular, the low-frequency parts of the spectra are usually considered caused by the systematic patterns while the high-frequency part is believed to be a result of the noise in the space domain [12], [15], [16]. However, the FFT/SFT is not able to reduce the domain dimension. For example, the two-dimensional spatial wafer metrology data remains two-dimensional after transforming to the frequency domain by the SFT. Moreover, the statistical properties of the SFT spectrum in the frequency domain are not clear enough to construct a hypothesis test for identifying the existence of systematic variations.

Systematic wafer spatial variations usually form patterns on the two-dimensional wafer map. The advancement of information technology has helped enhance the 2-D/3-D visualization of the wafer metrology data and thus facilitate the causal analysis when studying spatial variations. Wong et al. [17] propose a three-step methodology to characterize the line-width variation. Spatial analysis first decomposes the CD’s metrology data into several variance components. Causes with similar spatial signatures, defined based on engineering knowledge, are then classified by contributor-specific measurements. Unanticipated components are finally classified as residuals. Vanoppen et al. [7] apply the methodology for breaking down and ranking of the systematic sources of line-width variations. Evaluation of the exposure tool performance in relation to the contributors of line-width variation is also presented. Burch et al. recently propose a fault signal detection algorithm (FSDA) which serves as a yield fault detection and diagnosis solution integrating several practical data mining techniques and engineering data analysis methods [18], [19]. FSDA first identifies the known failure metrics and spatial/reticle zones in preparation of wafer data. Wafers with similar patterns are then clustered together. Characteristics of the patterned clusters are used for the drilldown yield analysis to identify the root causes of yield loss.

Most existing methods discussed above, no matter pre-assuming the variance components or pre-defining the exemplar variation patterns, require knowing related issues/faults of the process/tool in advance and categorize the unexplained parts as residuals. Engineering knowledge is definitely helpful in analyzing the variation components for specific failure types. However, engineering knowledge is usually acquired through a high-price learning process where faults or yield losses are found in the later stages of fabrication with corresponding engineering causes learned to locate in the much earlier stages. In fact, any systematic pattern must be associated with certain physical issues. The engineering knowledge can thus be established through relating the data excursion to the out-of-control processes. It is our belief that we should have the data reveals itself as much as possible so that the corresponding engineering knowledge can be learned as early as possible. Therefore, a model-free methodology without priori knowledge to reveal systematic patterns is proposed in this paper so that possible problems and respective knowledge can be learned immediately after they occur.

The novel spatial variance spectrum (SVS) to characterize the spatial variations is proposed in Section II. The SVS, which manifests the significance of systematic patterns, will be then summarized into an overall systematic pattern index based on the formation of the spectrum in Section III. The index is expected to provide a quick understanding of the spatial patterns and can be drilled down to the high/middle/low frequency portions for further analysis. Actual wafer metrology data with certain systematic variations are collected from a local semiconductor fab and analyzed for validation of the proposed methodologies in Section IV. Finally, some concluding remarks are made in Section V.

II. CHARACTERIZATION OF SPATIAL VARIATION

Assume that \( n \) observations, such as the thickness or linewidth, are taken from sampled metrology sites. \( m_i \) denotes the \( i^{th} \) observation at the metrology site with an Euclidean coordinate \((x_i, y_i)\) on the wafer (the origin is referred to the center of the wafer). As can be seen in Fig. 1, a basic understanding of this kind of dataset could be done by drawing a 2-D contour map (Fig. 1a) or a 3-D response surface (Fig. 1b).

To analyze the wafer spatial variations, the sample variance (1) is often used to characterize the spatial variation:

\[
\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} (m_i - \bar{m})^2}{n-1},
\]

where \( \bar{m} \) is the average of the \( n \) observations.

However, the sample variance is significantly biased if the observations comprise systematic components, which usually result in systematic patterns on the wafer surface as shown in Fig. 1. Chen and Blue [20] suggest calculating the moving variance for a temporal series of observations to reduce the estimate bias. A similar idea may be used for the spatial data. Instead of the moving variance of a temporal series, the sample variance is calculated for observations from a spatial area. The spatial sample variances calculated based on different sizes of spatial moving windows then form a spatial variance spectrum (SVS).

![Fig. 1. The visualization of a hypothetical wafer metrology data with dome pattern: (a) 2-D contour map; (b) 3-D response surface.](image)
A. Spatial Sample Variance

The spatial sample variance utilizes the spatial information, i.e. the Euclidean coordinates, to help decide the size and constituents of a spatial moving window. Given the observations \(m_i, i=1, \ldots, n\), we first calculate the Euclidean distances for all pairs of observations and get a symmetric distance matrix \(D\) with all 0’s on its diagonal:

\[
D = \begin{bmatrix}
    d_{11} & d_{12} & \cdots & d_{1n} \\
    d_{21} & d_{22} & \cdots & \vdots \\
    \vdots & \vdots & \ddots & \vdots \\
    d_{n1} & \cdots & \cdots & d_{nn}
\end{bmatrix}
\]

(2)

where \(d_{ij}\) is the distance between \(i^{th}\) and \(j^{th}\) observation.

For the \(i^{th}\) row in \(D\), we can sort the distances, \(d_{i1}, \ldots, d_{in}\), in an ascending order and find the \(k^{th}\)-nearest observation for observation \(i\), denoted as \(m_{i(k)}\). One special case is that \(m_{i(k)}\) is actually \(m_i\) itself because the nearest observation for \(i\) would be itself based on the distances in (2). A spatial moving window with size \(p\) (where \(2 \leq p \leq n\)) for observation \(i\) can be defined as:

\[
w_i^p = \{m_{i(1)}, m_{i(2)}, \ldots, m_{i(p)}\}, i=1, 2, \ldots, n \text{ and } 2 \leq p \leq n.
\]

(3)

The sample variance for the observations within the spatial moving window \(w_i^p\) is then calculated and denoted as \(s_{w_i^p}^2\).

Given a size \(p\), a total of \(n\) sample variances will be obtained from the \(n\) spatial moving windows \((w_i^p, i=1, 2, \ldots, n)\). These sample variances are then pooled together to be the spatial sample variance for window size \(p\), that is,

\[
s_p^2 = \frac{1}{n} \sum_{i=1}^{n} s_{w_i^p}^2,
\]

(4)

Furthermore, we can define the expected spatial sample variance for window size \(p\) as the expected value of \(s_p^2\). That is \(\sigma_i^2 = E(s_i^2)\) for \(2 \leq p \leq n\).

When \(p=n\), the moving window becomes the whole dataset, i.e. all \(n\) observations are used to calculate the sample variance. Thus, \(s_n^2\) is exactly equal to the sample variance in (1).

Fig. 2 shows an example of the spatial sample variance calculation given a window size \(p=3\). Assume the sampling locations are equal-spaced vertically or horizontally, and the vertical space between two adjacent observations is slightly larger than the horizontal one. For the first observation \(m_1\), the two nearest observations, i.e. \(m_{1(2)}\) and \(m_{1(3)}\), are \(m_2\) and \(m_3\), respectively. For the observation at each metrology site, there is a corresponding spatial moving window consisting of three observations to generate a sample variance. Then, the \(n\) spatial sample variances for \(p=3\) can be used to calculate \(s_n^2\) in (4).

![Fig. 2. An illustration for calculating the spatial sample variance of window size p=3.](image)

B. Spatial Variance Spectrum (SVS)

By varying the size of spatial moving window from 2 to \(n\), we will have \(n-1\) spatial sample variances:

\[
s_2^2, s_3^2, \ldots, s_{n-1}^2, s_n^2
\]

which are defined as the components of the SVS. Since the total number of metrology sites \(n\) may differ from wafer to wafer, using \(p\), whose range depends on \(n\), as the domain for the variance spectrum is not appropriate. We define the spatial variation frequency \(f = p/n\) (\(0 < f \leq 1\)) for the spectrum domain. A smaller (larger) \(f\) represents a higher (lower) frequency because it covers a smaller (larger) area for calculation of
spatial sample variances. The spatial variation frequency not only defines the frequency domain of the spectrum, but also allows comparing spectra calculated from wafers with different numbers of observations as long as the measurements are rather symmetrically and uniformly distributed over the entire wafer. The spatial sample variances can be plotted against the spatial frequency to illustrate the structure of spatial variations. Fig. 3 shows a spectrum calculated based on the simulated data in Fig. 1.

![Fig. 3. The SVS of a hypothetical wafer metrology data with dome pattern.](image)

The red dashed line in Fig. 3 denotes the sample variance calculated by (1). As can be seen, the spatial sample variances over the range of middle to low frequencies \( f > 0.3 \) are larger than those over the range of higher frequency \( f < 0.3 \). The larger sample variances with spatial windows covering at least 30% portion \((p/n)>0.3\) of the wafer surface reflect the pattern.

Fig. 4 shows a metrology data with observed variations. Because there is no dominant pattern, the SVS appears to be random around the overall sample variance. The entire spectrum is rather stationary as compared to that in Fig. 3 despite there are small fluctuations locally. In particular, the spatial sample variances of high frequencies, i.e. small window sizes, appear to be more unstable. This is because the small windows are extremely sensitive to the local randomly-formed patterns. However, the impact of the local, small patterns will be soon gradually canceled out as the window size increases, and the SVS becomes flat.

![Fig. 4. The contour map of a randomly-distributed wafer metrology data and its SVS.](image)

To further understand the properties of SVS, we assume the wafer spatial variations are purely the random noises. Let the observations \( m_i \) be generated by:

\[
m_i = a + \varepsilon_i \quad \text{for } i = 1, \ldots, n,\]

where \( a \neq 0 \) is the mean level of all metrology observations, and \( \varepsilon_i \sim N(0, \sigma^2) \) for \( i = 1, \ldots, n, \)

are independent and follow an identical normal distribution with zero mean and variance \( \sigma^2 \). And let the spatial systematic variations be variations other than random noises satisfying (6) and (7). Then we have the following theorem.

**Theorem 1.** The spatial variation must consist of variations other than random noises, i.e. systematic variations, if the expected spatial variance spectrum is uneven for \( 0 < f \leq 1 \).

**Proof:** see Appendix A.

Based on Theorem 1, we know that the expected SVS should display as a flat horizontal line if the wafer spatial variations purely consist of random variations. Otherwise, the wafer spatial variations with SVS exhibiting a non-stationary pattern must consist of both the systematic and random variations.

**C. Identification of Systematic Variations**

Our attempt is to detect whether the metrology data is only randomly distributed or contains systematic patterns. To check if the SVS is calculated from the data consisting of random variation only, we study the relation between the conventional sample variance and the proposed spatial sample variance, and develop the following theorem.

**Theorem 2.** If one takes a sample of \( p \) observations from a total of \( n \) observations, there will be \( C_p^n \) combinatorial possibilities. The average of the sample variances of the \( C_p^n \) combinatorial samples is then equal to the sample variance of the \( n \) observations (1). That is:

\[
\frac{1}{C_p^n} \sum_{k=1}^{C_p^n} s_{p,k}^2 = s_n^2, \quad \text{for } p = 2, \ldots, n,\]

where \( s_{p,k}^2 \) is the sample variance of the \( k \)th sample among the
combinatorial subsets. Based on this theorem, \( s_{n-1}^2 \) is equal to \( s_2^2 \) if the metrology sites are symmetrically distributed such that the \( n \) samples of \( s_{n-1}^2 \), i.e. \( w_i^{n-1}, i=1, \ldots, n \), are the same as those in the \( C_{n-1}^n \) combinatorial samples. From Theorem 1, to detect the existence of systematic variations, we need to detect whether the spatial variances are uneven. A hypothesis test of uneven spatial variances can be developed if the probability distribution of \( s_2^2 \) is known under (6) and (7). If \( s_{n-1}^2 = s_2^2 \), it is clear that

\[
\frac{(n-1)s_{n-1}^2}{\sigma^2} \sim \chi^2_{n-1} \quad \text{and} \quad \frac{(n-1)s_{n-1}^2}{\sigma^2} \sim \chi^2_{n-1},
\]

where \( \chi^2_{n-1} \) is a chi-square distribution with the degree of freedom \( n-1 \). For \( 2 < p < n-1, C_p^n > n \). The \( n \) samples used for calculating the proposed spatial sample variance, \( s_p^2 \), are only the subset of the \( C_p^n \) combinatorial samples. It is thus unclear what distributions \( s_p^2 \) will follow. However, Theorem 2 provides some clues that lead to the following conjecture.

**Conjecture 1.** If \( s_p^2, 2 \leq p \leq n-1 \), is calculated from a wafer metrology data defined in (6) and (7), then

\[
\frac{v_p s_p^2}{E(s_p^2)} = \frac{v_p s_p^2}{\sigma^2} = \frac{v_p s_p^2}{\sigma^2} \sim \chi^2_{v_p}, 2 \leq p \leq n-1,
\]

where \( v_p \) denotes the degree of freedom for the \( \chi^2 \) distribution and is no greater than \( n-1 \).

**Reasoning:** see Appendix C.

To determine \( v_p \) for \( 2 \leq p \leq n-1 \), we perform a Monte-Carlo simulation study and obtain the results in Fig. 5. As \( p \) decreases from \( n \), \( \lfloor n/2 \rfloor < p < n-1 \) (1/2<\( \epsilon < 1 \)), the variance of \( s_p^2 \) (blue solid line) increases because the \( n \) samples for calculating the \( s_p^2 \) is becoming a smaller subset of the \( C_p^n \) combinatorial samples. For \( p \leq \lfloor n/2 \rfloor \), the \( n \) samples is becoming a larger subset of \( C_p^n \) combinatorial samples as \( p \) continues to decrease from \( \lfloor n/2 \rfloor \). However, the variance of \( s_p^2 \) sharply turns up when \( p \) is decreased to be less than 30 if \( \epsilon < 0.05 \) because these particularly small window sizes are sensitive to both local random patterns and unbalanced sampling of observations for calculation of \( s_p^2 \).

With the Monte-Carlo estimate of the variances of \( s_p^2 \) and the assumption of the \( \chi^2 \) distribution (Conjecture 1), \( v_p \) in (9) can be then estimated as:

\[
v_p = \frac{2(E(s_p^2))}{V(s_p^2)} = \frac{2\sigma^2}{V(s_p^2)}, 2 \leq p \leq n
\]

where \( V(s_p^2) \) is the sample variance of 100,000 Monte-Carlo simulations \( s_p^2 \)'s. The red dashed line in Fig. 5 indicates the estimated degrees of freedom. Based on the estimated degrees of freedom, we can now construct the following hypothesis test.

\( H_0: \sigma_2^2 = \sigma_1^2 = \ldots = \sigma_{n-1}^2 = \sigma_0^2 \) versus \( H_1: \) there exist systematic variations.

To reject \( H_0 \), we choose the smallest spatial sample variance:

\[
s_{min}^2 = \min\{s_2^2, s_3^2, \ldots, s_{n-1}^2, s_n^2\},
\]

as the test base and test if all the rest of the spatial sample variances are equal to \( s_{min}^2 \):

\[
H_0: \sigma_p^2 = s_{min}^2, \text{ for } p \neq \arg\min\{s_i^2, k=1, \ldots, n\}.
\]

Since \( s_{min}^2 \) is used as the comparison base, only the one-sided test is required. Let \( \chi^2_{v_p, 1-\alpha} \) denote the critical value of the Chi-square distribution with \( v_p \) degrees of freedom and the type I error probability \( \alpha \). An upper control limit (UCL) can be constructed as:

\[
\text{UCL} = \sigma_p^2 \leq \frac{\chi^2_{v_p, 1-\alpha} s_{min}^2}{v_p} \text{ for } 2 \leq p \leq n-1.
\]

With the upper control limit, if there is at least one spatial sample variance exceeds the control limit, \( H_0 \) is rejected and systematic variations are said to exist; otherwise, there is no evidence to say that the wafer spatial variations consist of any systematic variation. Fig. 6 shows an example of using the upper control limit with \( \alpha = 0.05 \) to check if there exists any systematic variation for the pure-noise metrology data in Fig. 4. As can be seen, the whole spectrum lies within the control limit and there is no sufficient statistical evidence to deny that the metrology data is only randomly distributed.
III. SPATIAL PATTERN INDEX (SPI)

Even with the hypothesis test proposed above, it would be useful to provide simple indices summarizing the SVS to evaluate the significance of the systematic variations. A spatial pattern index (SPI) and three variation ratios of high, middle and low frequencies are proposed. The root causes associated with an identified systematic pattern can be then investigated based on the index and the three variation ratios.

Based on **Theorem 1**, the spectrum is expected to be a horizontal line if the wafer spatial variations purely consist of random variations and thus the spatial pattern index for a pure noise should be near zero accordingly. For the wafers consisting of systematic patterns, we expect to see a larger SPI value indicating a more significant systematic variation. The spatial pattern index (SPI) is proposed:

\[
SPI = \frac{\sum s_i^2 - s_{\text{min}}^2}{\sum s_i^2} ,
\]

(13)

where \( s_{\text{min}}^2 \) defined in (11) is used as an estimate of the random variation.

The denominator in (13) represents the total spatial variation of the spectrum and the numerator calculates the total systematic variation by removing the random variation from the spatial variation. Therefore, SPI is ranged between 0 and 1. Fig. 7 shows how the SPI is calculated for the wafer metrology data in Fig. 3. The SPI, the red area divided by the blue area, is calculated to be 0.6978.

For a randomly-distributed wafer metrology data as in Fig. 4, its SPI=0.0187 is very close to zero because the systematic variation portion is relatively small compared to the total spatial variation. To further explain the SPI, three variation ratios are calculated by dividing the spectrum into three parts, i.e. high, middle, and low frequencies, which are:

High-freq. variation ratio = \( \frac{\sum_{j=2}^{Nk} s_i^2}{\sum_{i=2}^{N} s_i^2} \times 100\% \),

(14)

Mid-freq. variation ratio = \( \frac{\sum_{j=3}^{Nk-1} s_i^2}{\sum_{i=2}^{N} s_i^2} \times 100\% \),

(15)

Low-freq. variation ratio = \( \frac{\sum_{j=1}^{Nk-2} s_i^2}{\sum_{i=2}^{N} s_i^2} \times 100\% \).

(16)

The SPI's and variation ratios for the hypothetical examples in Figs. 3 and 4 are listed in Table I. The three ratios for the randomly distributed metrology data are almost the same because the data is pattern-free. However, the middle-frequency variation ratio for the wafer with the dome pattern is the largest. This implies that the large SPI = 0.6978 is majorly contributed by the medium-size pattern.

With these indices summarized from the SVS, one can
quickly grasp the significance of the systematic variations. These indices provide simple but effective information about the wafer spatial variations and can be further used to find the root causes.

| Table I |
|-----------------|-----------------|-----------------|
| SPI and Variation Ratios of the Noise and Dome Pattern | Noise | Dome |
| Spatial Pattern Index (SPI) | 0.0187 | 0.6978 |
| High-freq. Variation Ratio | 33.19% | 24.60% |
| Mid-freq. Variation Ratio | 33.32% | 38.82% |
| Low-freq. Variation Ratio | 33.49% | 36.58% |

IV. Case Study

To validate the proposed methodology which identifies systematic patterns resulted from systematic variations within wafer metrology data, hypothetical wafer metrology data are generated and two sets of real wafer CD metrology data are also collected and analyzed.

A. Hypothetical Wafer Metrology Data with Common Patterns

First, we simulate four common systematic patterns with noise disturbance: x-direction drift, y direction drift, dome, and donut to validate the proposed methodology. The contour maps for the four wafers are plotted in Fig. 8 and their SVS’s are calculated as well. In Fig. 8, the x- and y-direction drifting patterns share very similar spectra where the slight difference is resulted from the random variation. Therefore, the SVS’s of the two patterns are similar. The spectrum of the donut pattern reaches a stably high variation level of mid-to-low frequencies \((f > 0.2)\) because the rise and fall of the donut pattern covers 4/5 portion of the wafer.

To test if the four spectra are significantly distinct from that of a randomly distributed wafer metrology data, the upper control limit described in Section II is then built to identify the existence of systematic variations. To save the pages, we only demonstrate the hypothesis testing result of the wafer with dome pattern in Fig. 9 and the other three datasets can follow the same fashion. As can be seen in Fig. 9, the SVS goes out of the control.
limits quickly as $p$ increases. Therefore, the wafer spatial variations are said to contain not only random but also systematic variations.

The spectra are further summarized into the SPI’s and the variation ratios in Table II. The four wafers with systematic patterns all have significantly-high indices as compared to that with pure random variation. In particular, the low-frequency variation ratios of $x$- and $y$-direction drifting patterns are very high because the drifting pattern actually covers up the whole wafer.

Due to the measuring precision, the systematic variations of OCD are smaller than that of ACD explains ACD actually consists of more random noise than OCD. The high frequency variation ratios which decrease from SEM CD to OCD as the precision of those measurements becomes more and more precise also tell the same story. Moreover, the larger low-frequency variation ratio of OCD indicates a clearer pattern with larger coverage of the wafer.

As a result, the SVS and its summarized indices can be used to examine the precision levels of the wafer metrology data from different measuring methods. The systematic and random variations of a wafer can be truthfully captured as the precision of measuring methods changes.

### Table II

SPI AND VARIATION RATIOS FOR THE HYPOTHETICAL WAFER DATA

<table>
<thead>
<tr>
<th>Noise</th>
<th>X</th>
<th>Y</th>
<th>Dome</th>
<th>Donut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Pattern Index (SPI)</td>
<td>0.0372</td>
<td>0.5420</td>
<td>0.4882</td>
<td>0.6978</td>
</tr>
<tr>
<td>High-freq. Variation Ratio</td>
<td>33.11%</td>
<td>23.78%</td>
<td>23.03%</td>
<td>24.60%</td>
</tr>
<tr>
<td>Mid-freq. Variation Ratio</td>
<td>33.23%</td>
<td>33.39%</td>
<td>33.40%</td>
<td>38.82%</td>
</tr>
<tr>
<td>Low-freq. Variation Ratio</td>
<td>33.67%</td>
<td>44.84%</td>
<td>43.57%</td>
<td>36.58%</td>
</tr>
</tbody>
</table>

### Table III

SPI AND VARIATION RATIOS OF REAL WAFER METROLOGY DATA WITH DIFFERENT PRECISION LEVELS

<table>
<thead>
<tr>
<th>Noise</th>
<th>X</th>
<th>Y</th>
<th>Dome</th>
<th>Donut</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Pattern Index (SPI)</td>
<td>0.3384</td>
<td>0.7581</td>
<td>0.9465</td>
<td></td>
</tr>
<tr>
<td>High-freq. Variation Ratio</td>
<td>29.55%</td>
<td>25.60%</td>
<td>23.81%</td>
<td></td>
</tr>
<tr>
<td>Mid-freq. Variation Ratio</td>
<td>34.20%</td>
<td>35.16%</td>
<td>36.17%</td>
<td></td>
</tr>
<tr>
<td>Low-freq. Variation Ratio</td>
<td>36.25%</td>
<td>39.24%</td>
<td>40.02%</td>
<td></td>
</tr>
</tbody>
</table>

B. Wafer Metrology Data with Different Precision Levels

Real wafer metrology data from a local semiconductor fab is collected for methodology validation as well. Wafers after the post-exposure bake (PEB) process are usually measured by optical critical dimension (OCD) and scanning electron microscope (SEM), respectively. Although the measurement precision of OCD is better than SEM for the 65nm technology node, the recipe and library generation of OCD is more complicated and time consuming. Ke et al. [21] proposed a concept on OCD-like CD SEM measurement which is said to be the average line width (ALW) or contact hole diameter (ACD) measurement at high resolution and low magnification CD SEM. These measurements, i.e. SEM CD, ALW/ACD, and OCD, manifest different significances of systematic variations of the same wafer due to the measuring precision. Therefore, the proposed SVS and SPI would be an appropriate way to analyze the systematic variations of wafers, and thus examine the precision of those measurements.

The SVS’s of the three kinds of measurements and the control limit for identifying the systematic variations of SEM CD are calculated and plotted in Fig. 10 and 11. The SPI’s and variation ratios are summarized in Table III. As shown in Fig. 10, the overall variation for the less-precise measurements, SEM CD, is much higher than those of ACD and OCD. This is because SEM CD consists of much noise, and thus its $SPI = 0.3384$ is much lower than the other two. The OCD which provides measurements with high precision exhibits a clear pattern on the contour map and a very high $SPI = 0.9465$. The phenomenon that the high-frequency spatial sample variances of OCD are smaller than that of ACD explains ACD actually consists of more random noise than OCD. The high frequency variation ratios which decrease from SEM CD to OCD as the measurements become more and more precise also tell the same story. Moreover, the larger low-frequency variation ratio of OCD indicates a clearer pattern with larger coverage of the wafer.

C. Wafer Metrology Data with Checkerboard Pattern

Another set of wafer metrology data, CD’s from the step-and-scan system, are also analyzed. Checkerboard patterns are sometimes formed after the scanning process because the scanning direction of the tool which scans downward when exposing the odd-numbered chips (or fields) and upward while...
performing exposures on even-numbered ones [17]. As can be seen in Fig. 12, the two wafers not only consist of systematic patterns such as bowl (S1) and dome (S2), but also contain checkerboard patterns. The spectrum of S1 is tested by its upper control limit in Fig. 13. The spatial pattern indices and variation ratios are listed in Table IV.

The checkerboard pattern induced by the alternative scanning effect is considered to be characterized by the high-frequency spatial sample variances because the small-size spatial moving windows would cover the observations with alternative effect and show large sample variances. As can be seen in Fig. 13, the SVS of S1 starts higher than the upper control limit when the spatial variation frequency is very high ($f \to 0$), and immediately goes lower than the control limit as $f$ increases. However, it again goes out of the control limit because the bowl pattern soon takes over and impacts on the spectrum. A checkerboard pattern that behaves very similar to the randomly distributed wafer metrology data will distort the systematic patterns and result in lower SPI’s (0.3 and 0.5244). However, it still can be identified by the high-frequency spatial sample variances through the proposed hypothesis test.

![Fig. 12. Analysis of SVS’s for the real wafer metrology data with checkerboard patterns.](image)

V. CONCLUSION

Investigation of wafer spatial variations is critical to process optimization and yield improvement in semiconductor manufacturing. Various methods have been proposed in the literature to identify the systematic components based on priori engineering knowledge about the process or equipment. In this paper, we have proposed a model-free spatial variance spectrum (SVS) to analyze the spatial metrology data so that any engineering problems and issues can be learned as early as possible through identification of systematic patterns. The SVS is generated by the spatial moving sample variances with different spatial window sizes. A hypothesis test has been developed based on the statistical properties of SVS to detect the existence of systematic variations. We have also developed spatial pattern index (SPI) and three variation ratios to provide engineers with a quick look at the systematic variations. Both hypothetical and actual metrology cases are used to validate the proposed methodology. Results show that systematic patterns can be truthfully characterized by the proposed SVS even for metrology with low precision levels and for the unusual checkerboard patterns.

It is worth noting that the metrology sites sampled on a wafer during production may not be uniformly distributed. For the proposed spatial sample variances to work in this case, the moving window $w_i^p$, which consists of $p$ observations including $m_i$ and its $p-1$ neighboring observations, should be now redefined as a circular area with the center at $m_i$ and a radius $r$. Such an area-based moving window may cover indefinite numbers of observations. Let this area-based moving window be denoted as $c_i^r$. By increasing $r$ from 0 to the diameter of the wafer, $R$, we can still calculate the spatial sample

![Fig. 13. The identification of systematic variations for the SVS of S1. (The round shape magnifies the high-frequency part of the spectrum.)](image)

| Table IV: SPI and Variation Ratios of Real Wafer Metrology Data with Checkerboard Patterns |
|----------------------------------|------------------|------------------|
| **Spatial Pattern Index (SPI)**  | 0.3000           | 0.5244           |
| **High-freq. Variation Ratio**   | 28.60%           | 24.97%           |
| **Mid-freq. Variation Ratio**    | 35.09%           | 38.08%           |
| **Low-freq. Variation Ratio**    | 36.31%           | 36.95%           |
variances for different frequencies $f$, which is now calculated as $r^2 R^2$. Using $c_1'$ as the moving window is to uphold the meaning of the variance spectrum where the spatial variation for smaller $r'$s indicate the high-frequency variations and are usually generated by the noises while larger $r'$s represent the low to mid range of frequency and are typically produced by systematic patterns. With a limited number of sampling sites (usually five or nine sites sampled per wafer) during production, the SVS resolution becomes lower and the spatial variance analysis may not be as effective. However, compared to conventional analysis based on just the overall sample variance, the SVS analysis based on just the overall sample variance, the spatial variance analysis may not be as effective. However, compared to conventional analysis based on just the overall sample variance, the SVS still reveals more information about the spatial variations and its corresponding pattern indices still provide a quick look at the significance of systematic patterns, but are, of course, less truthful due to the low sampling resolution.

APPENDIX

A. Proof of Theorem 1

Since the wafer spatial variations only consist of random variation, the sample variance within a moving window, i.e. $s^2_{wp}$, can be expressed as:

$$s^2_{wp} = \frac{\sum_{j=1}^{p} (m_{i,j} - \bar{m}_{iwp})^2}{p-1},$$

where $\bar{m}_{iwp} = \frac{\sum_{j=1}^{p} m_{i,j}}{p}$ is the average of observations within the moving window $w_{wp}$. It is known that the sample variance is an unbiased estimator of the variance, i.e.

$$E(s^2_{wp}) = E\left(\frac{\sum_{j=1}^{p} (m_{i,j} - \bar{m}_{iwp})^2}{p-1}\right) = \sigma^2.$$

Therefore, the expected value of spatial variance for window size $p$ can be simply derived as:

$$E(s^2_p) = E\left(\frac{1}{n} \sum_{i=1}^{n} s^2_{wp}\right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} E(s^2_{wp})$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sigma^2$$

$$= \sigma^2.$$

As a result, if the spatial variations are only random noises as in (6) and (7), the expected values of the spatial sample variances for $2 \leq p \leq n$ must be all equal. That is, if the expected spatial sample variances are not all equal, then the spatial variations must consist of variations other than random noises.

B. Proof of Theorem 2

Let $\bar{m}_{i,1-1(n)}$ and $s^2_{i,1-1(n)}$ denote the sample mean and variance based on $n-1$ observations wherein $m_i$ is excluded. It is then straightforward to derive

$$(n-1)s^2 = (n-2)s^2_{n-1,(n)} + \frac{n-1}{n} (m_n - \bar{m}_{n-1,(n)})^2$$

by replacing the overall sample mean $\bar{m}$ in (1) with $\bar{m} = m_n + (n-1)\bar{m}_{n-1,(n)}$ [22].

Similarly, the excluded observation in (17) can be replaced by the other $n-1$ observations, and we have the following equations:

$$(n-1)s^2 = (n-2)s^2_{n-1,(n)} + \frac{n-1}{n} (m_{n-1} - \bar{m}_{n-1,(n-1)})^2$$

$$= (n-2)s^2_{n-1,(n-1)} + \frac{n-1}{n} (m_{n-1} - \bar{m}_{n-1,(n-1)})^2$$

$$= \vdots$$

$$= (n-2)s^2_{n-1,(n)} + \frac{n-1}{n} (m_1 - \bar{m}_{n-1,(1)})^2$$

We can then derive the following relation

$$\frac{n-1}{n} (m_n - \bar{m}_{n-1,(n)})^2 + \frac{n-1}{n} (m_{n-1} - \bar{m}_{n-1,(n-1)})^2$$

$$+ \cdots + \frac{n-1}{n} (m_1 - \bar{m}_{n-1,(1)})^2$$

$$= s^2_{n-1,(n)} + s^2_{n-1,(n-1)} + \cdots + s^2_{1,(1)}$$

$$\Rightarrow \sum_{i=1}^{n} \frac{n-1}{n} (m_i - \bar{m}_{n-1,(1)})^2 = \sum_{i=1}^{n} s^2_{i,1-1(i)}$$

(19)

by expanding the sample variances in (18) with the conventional calculation of sample variances, that is:

$$s^2_{i,1-1(i)} = \frac{n-2}{n-1} \left( \sum_{j=1}^{n} m_j^2 - \left( \sum_{j=1}^{n} m_j \right)^2 \right), \quad \text{for } i=1, \ldots, n.$$

If all the RHS’s in (18) are summed up, we get:

$$\frac{n(n-1)}{n-2} s^2 = (n-2)s^2_{n-1,(n)} + \frac{n-1}{n} (m_n - \bar{m}_{n-1,(n)})^2$$

$$+ (n-2)s^2_{n-1,(n-1)} + \frac{n-1}{n} (m_{n-1} - \bar{m}_{n-1,(n-1)})^2$$

$$+ \vdots$$

$$+ (n-2)s^2_{n-1,(1)} + \frac{n-1}{n} (m_1 - \bar{m}_{n-1,(1)})^2$$

$$= (n-2)\sum_{i=1}^{n} s^2_{i,1-1(i)} + \sum_{i=1}^{n-1} \frac{n-1}{n} (m_i - \bar{m}_{n-1,(i)})^2$$

$$= (n-2)\sum_{i=1}^{n} s^2_{i,1-1(i)} + \sum_{i=1}^{n} s^2_{i,1-1(i)}$$

$$= (n-1)\sum_{i=1}^{n} s^2_{i,1-1(i)}$$

(20)

Given $p=n-1$ in (8), we have

$$s^2 = \frac{1}{C_{n-1}} \sum_{k=1}^{C_{n-1}} s^2_{i,1-1(k)} = \frac{1}{n} \sum_{k=1}^{n} s^2_{i,1-1(k)},$$

which is identical to (20).
Each \( s^2_{n-2(i,j)} \) can be further expressed in terms of the sample variances of the samples with \( n-2 \) observations, and (20) becomes:

\[
s^2 = \frac{1}{n} \sum_{i=1}^{n} s^2_{n-2(i)} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{n-1} \sum_{j=1, j \neq i}^{n} s^2_{n-2(i,j)},
\]

where \( s^2_{n-2(i,j)} \) denotes the sample variance based on \( n-2 \) observations wherein the \( i^{th} \) and \( j^{th} \) observations are both excluded. Each \( s^2_{n-2(i,j)} \) in the RHS of (21) is repeatedly calculated, ex: \( s^2_{n-2(1,2)} \) and \( s^2_{n-2(2,3)} \) are identical and counted. Therefore, (21) can be revised as:

\[
s^2 = \frac{1}{n} \sum_{i=1}^{n} s^2_{n-2(i)}
\]

\[
= \frac{1}{n} \sum_{i=1}^{n} \frac{2}{n-1} \sum_{j=1, j \neq i}^{n} s^2_{n-2(i,j)}
\]

\[
= \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} s^2_{n-2(i,j)}
\]

\[
= \frac{1}{C_{n-2}} \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} s^2_{n-2(i,j)},
\]

which is also identical to (8) with \( p=n-2 \). By applying the result in (20) recursively, Theorem 2 is proved.

C. Reasoning for Conjecture 1

To reason Conjecture 1, in addition to the inspiration by Theorem 2, Monte-Carlo simulations are performed to study the distributions of the spatial sample variances. 100,000 randomly distributed wafers are generated from \( N(0, 1) \), and their spectra are calculated accordingly. One can draw a histogram and perform goodness of fit test to see how the distribution of the 100,000 spatial sample variances fit to the Chi-square distributions. Here, we use \( p=2 \) and 308 (\( n=616 \)) as two examples to show how they resemble the chi-square distribution. Fig. 14(a) and (b) show the histograms fit to the chi-square distributions and the \( p \)-values of the chi-square goodness of fit tests. The \( p \)-values are all near zero and the histograms look perfectly fit to the chi-square distribution curves for both \( p=2 \) and 308 (\( f=0.0032 \) and 0.5).

![Histograms of the simulated spatial sample variances under spatial variation frequencies: (a) \( p=2 \); (b) \( p=308 \).](image)

**ACKNOWLEDGEMENT**

We are grateful to a local IC maker who provides necessary funding and, most importantly, the valuable tool data for our methodology development and validation. This research is also funded in part by NSC98-2221-E-002-132.

**REFERENCES**


