Measurement System R&R Analysis for Zero-Inflated Correlated Defect Counts in Semiconductor Wafer AOI Inspection

Chih-Chiang Fang and Ming-Nan Chen *

Posted Date: 14 May 2026

doi: 10.20944/preprints202605.0948.v1

Keywords: semiconductor wafer inspection; automated optical inspection (AOI); Gauge R&R; measurement system analysis; zero-inflated bivariate Poisson model; correlated defect counts; EM algorithm; precision manufacturing; defect-count modeling; measurement uncertainty; repeatability and reproducibility; statistical quality control

 doi:10.20944/preprints202605.0948.v1

Article

Measurement System R&R Analysis for Zero-Inflated Correlated Defect Counts in Semiconductor Wafer AOI Inspection

Chih-Chiang Fang [1] and Ming-Nan Chen [2,*]

[1] School of Computer Science and Software, Zhaoqing University, Zhaoqing 526061, China

[2] School of Economics and Management, Zhaoqing University, Zhaoqing 526061, China

* Correspondence: t46111024@gs.ncku.edu.tw

Abstract

This study proposes a novel measurement system repeatability and reproducibility (R&R) framework for zero-inflated correlated defect-count data in semiconductor wafer automated optical inspection (AOI). In advanced semiconductor manufacturing environments, AOI systems are extensively used to detect wafer defects such as particles, scratches, and structural abnormalities. However, conventional Gauge R&R methods are primarily developed for continuous Gaussian-type measurements and are therefore not fully appropriate for high-yield semiconductor inspection data characterized by discrete defect counts, excessive zero observations, and correlated defect categories. To address these limitations, this study develops a zero-inflated bivariate Poisson (ZIBP) measurement system model capable of simultaneously capturing correlated defect-generation mechanisms and structural zero-defect states. A latent-variable representation is introduced to model shared and category-specific defect sources, while a zero-inflation mechanism accounts for defect-free wafer observations commonly encountered in precision manufacturing. An expectation-maximization (EM) algorithm is further developed for parameter estimation, including latent common defect counts and structural-zero probabilities. Based on the fitted model, repeatability variance, reproducibility variance, total measurement variation, and Percent R&R are estimated under the proposed probabilistic framework. In addition, bootstrap resampling is employed to construct confidence intervals for the proposed R&R measures. Theoretical properties of the proposed framework, including covariance structure, identifiability, EM monotonicity, estimator consistency, and asymptotic behavior of the Percent R&R estimator, are analytically established. The proposed framework extends traditional Gauge R&R analysis from continuous Gaussian measurements to zero-inflated correlated count-type defect inspection data and provides a statistically rigorous methodology for evaluating AOI measurement system reliability in semiconductor wafer manufacturing environments.

Keywords: semiconductor wafer inspection; automated optical inspection (AOI); Gauge R&R; measurement system analysis; zero-inflated bivariate Poisson model; correlated defect counts; EM algorithm; precision manufacturing; defect-count modeling; measurement uncertainty; repeatability and reproducibility; statistical quality control

MSC: 62N05; 62C10; 65C20

1. Introduction

Semiconductor manufacturing has become one of the most technologically intensive and quality-sensitive industries in modern production systems. As semiconductor devices continue to evolve toward smaller feature sizes, higher integration density, and increasingly complex process architectures, the need for accurate, stable, and reliable inspection systems has become increasingly

 doi:10.20944/preprints202605.0948.v1

important. Automated optical inspection (AOI) systems are widely adopted in electronics and semiconductor manufacturing because they provide high-throughput, noncontact, and automated defect-detection capabilities for wafer surfaces, chips, packages, and micro-scale manufacturing components [1]. In semiconductor wafer manufacturing, AOI systems are frequently used to detect particle contamination, scratches, pattern abnormalities, structural defects, and surface irregularities. Recent studies have further demonstrated the importance of AOI and machine-learning-based inspection systems in wafer acceptance testing, wafer defect inspection, chip localization, and semiconductor defect detection [2–6].

Although AOI technologies have advanced rapidly, the reliability of the measurement system itself remains a critical issue. In practical semiconductor fabrication lines, multiple AOI tools, recipes, inspection stations, or operators may be used for wafer inspection. Even when these systems operate under nominally identical conditions, differences in illumination, focus calibration, optical sensitivity, detection thresholds, algorithmic settings, and inspection recipes may lead to inconsistent defect-count measurements. Such inconsistency can directly influence yield analysis, process monitoring, defect excursion detection, root-cause analysis, and production decision making. Therefore, a rigorous measurement system analysis framework is required to evaluate whether AOI inspection systems provide repeatable and reproducible defect-count measurements.

Traditional measurement system analysis, particularly Gauge repeatability and reproducibility (Gauge R&R), has been widely used to evaluate measurement-system capability in manufacturing environments. Existing studies have examined Gauge R&R for destructive measurements, micron-level metrology, ANOVA-based analysis, Bayesian inference, operator effects, nominal data, and optimal experimental designs [7–17]. However, most traditional Gauge R&R methods were originally developed for continuous measurement data under approximately normal assumptions. This assumption is often inappropriate for semiconductor AOI inspection data because the observed responses are usually non-negative integer-valued defect counts rather than continuous measurements.

In addition to the count-data nature of AOI inspection results, semiconductor wafer defect data frequently exhibit a high proportion of zero-defect observations. Under high-yield manufacturing conditions, many wafer regions may contain no observable defects. Consequently, the number of zero observations may be much larger than expected under a standard Poisson distribution. This phenomenon is commonly referred to as zero inflation. Zero-inflated Poisson models and their extensions have been extensively studied for sparse count data in statistics, bioinformatics, social sciences, causal learning, manufacturing deterioration modeling, and network science [18–25]. Nevertheless, the integration of zero-inflated count-data modeling into measurement system R&R analysis for semiconductor AOI inspection remains limited.

Another important feature of semiconductor defect data is the correlation among defect categories. Particle contamination, pattern defects, scratches, and structural abnormalities may not occur independently because they can be induced by common process disturbances, such as contamination events, lithography instability, wafer handling variation, plasma etching nonuniformity, and localized process drift. Therefore, a suitable statistical model for semiconductor AOI inspection should not only handle zero inflation and count data but also capture correlation among multiple defect categories. A bivariate Poisson structure is particularly suitable for this purpose because it can represent both category-specific defect sources and shared latent defect-generation mechanisms.

Recent semiconductor manufacturing quality research has increasingly focused on virtual metrology, explainable artificial intelligence, smart inspection, process optimization, and advanced quality analytics [26–32]. These studies demonstrate the growing importance of data-driven and statistical methods in semiconductor production systems. However, most existing studies focus on defect detection, prediction accuracy, process control, or optimization rather than measurement-system variability analysis under zero-inflated correlated defect-count conditions. As a result, a

 doi:10.20944/preprints202605.0948.v1

methodological gap remains between conventional Gauge R&R analysis and the statistical characteristics of modern semiconductor AOI defect inspection data.

To address this gap, this study proposes a measurement system R&R framework for zero-inflated correlated defect counts in semiconductor wafer AOI inspection. The proposed framework extends conventional Gauge R&R analysis by incorporating a zero-inflated bivariate Poisson model. The zero-inflation component represents structural zero-defect states commonly observed in high-yield semiconductor manufacturing, while the bivariate Poisson component captures correlated defect categories through a shared latent defect source. An expectation-maximization (EM) algorithm is developed to estimate latent structural-zero states, shared defect components, and defect-intensity parameters. Based on the fitted model, repeatability variance, reproducibility variance, total measurement variation, and Percent R&R are estimated.

The contributions of this study are fourfold. First, this study extends measurement system R&R analysis from continuous Gaussian-type measurements to zero-inflated correlated count-type defect data. Second, a zero-inflated bivariate Poisson measurement model is formulated for semiconductor wafer AOI inspection, enabling simultaneous modeling of structural zero-defect states and correlated defect categories. Third, an EM-based estimation procedure is developed to estimate latent variables and model parameters within the proposed framework. Fourth, theoretical properties and numerical analyses are provided to demonstrate the validity and practical applicability of the proposed method in semiconductor wafer AOI inspection environments.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature on measurement system analysis, semiconductor wafer AOI inspection, zero-inflated count-data modeling, and semiconductor manufacturing quality analytics. Section 3 presents the proposed mathematical model and EM estimation procedure. Section 4 provides theoretical properties and analytical propositions. Section 5 presents the semiconductor wafer AOI numerical analysis. Section 6 concludes the paper and discusses future research directions.

2. Literature Review

2.1. Measurement System Analysis and Gauge R&R

Measurement system analysis is a fundamental component of statistical quality control because inaccurate or unstable measurement systems can lead to incorrect process evaluation, poor quality decisions, and ineffective improvement actions. Gauge repeatability and reproducibility analysis is one of the most widely used methods for assessing measurement-system variation. The purpose of Gauge R&R is to decompose the observed measurement variation into repeatability variation, which reflects within-system measurement error, and reproducibility variation, which reflects between-system or between-operator measurement differences.

De Mast and Trip [7] investigated Gauge R&R studies for destructive measurements and discussed the methodological difficulty of evaluating measurement systems when repeated measurements on the same item are not possible. Hoffa and Laux [8] examined the use of Gauge R&R for micron-level metrology and emphasized the need for measurement adequacy evaluation in high-precision manufacturing environments. Kazerouni [9] further discussed the design and analysis of Gauge R&R studies based on the ANOVA method and highlighted the importance of statistical decision making in measurement-system evaluation.

Bayesian and advanced statistical approaches have also been developed for Gauge R&R analysis. Weaver et al. [10] proposed a Bayesian approach to Gauge R&R data analysis, allowing uncertainty in variance components to be quantified more flexibly. Sharma et al. [11] applied Gauge R&R to validate destructive measurement systems in an engineering case study. Mircea et al. [12] demonstrated the use of Gauge R&R in agricultural machinery manufacturing, showing that measurement-system evaluation remains important across diverse manufacturing sectors. Petitjean and Musset [13] examined repeatability and reproducibility in an honest Gauge R&R study, while Hamada [14] proposed operator linear analysis for Gauge R&R studies. Malindzakova et al. [15]

 doi:10.20944/preprints202605.0948.v1

further emphasized the use of repeatability and reproducibility methods in measurement quality control. Hamada and Ryan [16] studied Gauge R&R for nominal data, and Petitjean [17] proposed optimal design approaches to reduce the number of required measurements in Gauge R&R studies.

Although these studies have significantly advanced Gauge R&R methodology, most existing approaches are primarily designed for continuous or categorical measurement data. Semiconductor wafer AOI inspection, however, often produces discrete defect counts with excessive zero observations and correlated defect categories. Therefore, applying conventional Gauge R&R directly to semiconductor AOI defect-count data may lead to inappropriate variance decomposition and misleading measurement-system evaluation.

2.2. Semiconductor Wafer AOI Inspection and Defect Analysis

Automated optical inspection has become an essential inspection technology in electronics and semiconductor manufacturing. Abd Al Rahman and Mousavi [1] provided a comprehensive review of AOI and quality-monitoring methods in the electronics industry, showing that AOI systems have become increasingly important for automated defect detection and quality assurance. Wang and Yang [2] proposed a machine-learning approach for improving wafer acceptance testing based on station and equipment combinations, demonstrating the potential of data-driven methods for semiconductor inspection decision support. Wu et al. [3] applied denoising diffusion probabilistic models to wafer defect inspection and showed the increasing role of advanced generative models in semiconductor defect analysis.

Several studies have further focused on AOI systems and semiconductor defect detection. Fu et al. [4] proposed a central array method to locate chips in AOI systems for semiconductor manufacturing. Fu et al. [5] investigated the use of artificial intelligence to improve chip defect detection using semiconductor equipment data. Tabatabaeemoshiri et al. [6] developed a graph-based semi-supervised learning method for test-induced defect detection in semiconductor wafers. Nurmi [33] studied improvements in resolution, repeatability, and detection capability of AOI systems in micro-electro-mechanical systems manufacturing. Li et al. [34] further demonstrated that AOI concepts can be transferred from wafer inspection to multiscale characterization in smart battery manufacturing.

These studies demonstrate the rapid development of AOI technologies, machine learning, and intelligent inspection systems. However, most existing AOI studies primarily focus on detection accuracy, defect classification, image-processing performance, and inspection automation. Comparatively less attention has been paid to the statistical reliability of AOI measurement systems themselves. In particular, the repeatability and reproducibility of AOI defect-count measurements under zero-inflated correlated defect conditions remain insufficiently studied.

2.3. Zero-Inflated and Correlated Count-Data Modeling

Zero-inflated count-data models have been widely studied for datasets in which the number of zero observations exceeds the level expected under standard count distributions. Wagh and Kamalja [18] reviewed zero-inflated models and estimation methods for zero-inflated Poisson distributions. Zhou et al. [19] applied a zero-inflated Poisson model to classify next-generation sequencing data, showing its usefulness in sparse biological count-data environments. Kamalja and Wagh [20] further developed estimation methods for zero-inflated generalized Poisson distributions.

Zero-inflated models have also been applied across broader scientific domains. Truong et al. [21] discussed zero-inflated Poisson regression models and their applications in the sciences and social sciences. Choi et al. [22] developed Bayesian causal structural learning with zero-inflated Poisson Bayesian networks, demonstrating the usefulness of zero-inflated structures in probabilistic graphical modeling. Zhang and Yi [23] studied zero-inflated Poisson models with measurement error in the response, which is highly relevant to measurement-related count-data analysis. Fang and Chen [24] integrated zero-inflated non-homogeneous Poisson deterioration into production run optimization for imperfect production systems, showing the value of zero-inflated modeling in

 doi:10.20944/preprints202605.0948.v1

manufacturing-quality problems. Lu et al. [25] proposed a zero-inflated Poisson latent position cluster model for network science, further demonstrating the flexibility of zero-inflated latent-variable modeling.

Although the zero-inflated modeling literature is extensive, its integration with measurement system analysis remains limited. In particular, existing studies rarely consider the simultaneous presence of zero inflation, correlated defect categories, and measurement-system repeatability and reproducibility. This gap is critical for semiconductor wafer AOI inspection because defect-count data often exhibit all three characteristics simultaneously.

2.4. Semiconductor Manufacturing Quality and Intelligent Inspection Systems

Semiconductor manufacturing quality research has increasingly incorporated statistical modeling, artificial intelligence, virtual metrology, and optimization methods. Chien et al. [26] proposed decision-based virtual metrology for advanced process control in semiconductor manufacturing, demonstrating how predictive measurement models can support smart production. Senoner et al. [27] showed that explainable artificial intelligence can improve semiconductor process quality and support manufacturing decision making. Maitra et al. [28] reviewed virtual metrology in semiconductor manufacturing and discussed future prospects for intelligent process monitoring.

Recent research has further expanded semiconductor manufacturing analytics into statistical modeling, smart inspection, defect propagation, and production optimization. Kim et al. [29] applied statistical models for optimal packaging in semiconductor manufacturing processes. Suwattananuruk and Chien [30] developed denoising variational autoencoders for smart inspection of wafer probe card PCB channels, contributing to advanced semiconductor quality control. Fiorino et al. [31] investigated defect propagation in 4H-SiC from substrate to epitaxial layers, emphasizing the importance of understanding defect-generation mechanisms. Jiang et al. [32] proposed a Q-learning-enhanced multi-objective particle swarm optimization method for photolithography-etching scheduling in semiconductor manufacturing.

These studies demonstrate that semiconductor quality control is increasingly moving toward intelligent, data-driven, and statistically advanced manufacturing systems. However, most existing semiconductor manufacturing quality studies focus on process prediction, virtual metrology, smart inspection, scheduling, and defect propagation. Few studies explicitly investigate AOI measurement-system variability under zero-inflated correlated defect-count conditions. Therefore, a rigorous R&R framework for semiconductor wafer AOI inspection remains necessary.

2.5. Research Gap and Motivation

The literature reviewed above suggests several important gaps. First, although Gauge R&R has been widely studied, most existing approaches assume continuous, normal, or categorical measurement data and are not specifically designed for zero-inflated defect-count data. Second, although AOI inspection technologies have advanced rapidly in semiconductor manufacturing, most studies focus on detection algorithms, classification accuracy, or image-processing methods rather than measurement-system variability. Third, although zero-inflated Poisson and related count-data models have been widely studied, they have not been sufficiently integrated with Gauge R&R analysis. Fourth, semiconductor manufacturing quality research has increasingly adopted artificial intelligence and virtual metrology, but relatively little work has addressed repeatability and reproducibility evaluation for AOI defect-count inspection systems.

To address these limitations, this study develops a zero-inflated bivariate Poisson R&R framework for semiconductor wafer AOI inspection. The proposed model captures excessive zero-defect observations through a structural-zero mechanism and represents correlated defect categories through a shared latent Poisson component. By integrating EM estimation, variance decomposition, and bootstrap inference, the proposed methodology provides a statistically rigorous framework for evaluating measurement-system reliability in high-yield semiconductor AOI inspection environments.

 doi:10.20944/preprints202605.0948.v1

3. Research Problem and Mathematical Model Development

3.1. Research Problem

In semiconductor wafer manufacturing, automated optical inspection (AOI) systems are widely used to detect surface defects, particle contamination, scratches, pattern abnormalities, and other defect types that may influence yield and downstream reliability. As the semiconductor industry moves toward high-yield and high-precision manufacturing environments, the majority of inspected wafers, dies, or inspection regions may exhibit no observable defects. Consequently, the inspection data generated by AOI systems are often characterized by two important statistical features. First, the response variables are not continuous measurements but non-negative integer-valued defect counts. Second, the number of zero observations is often much larger than would be expected under a conventional Poisson distribution. This phenomenon is known as zero inflation.

Traditional measurement system analysis, especially Gauge Repeatability and Reproducibility analysis, was originally developed for continuous measurement data under approximately normal assumptions. In a classical Gauge R&R framework, the observed measurement is decomposed into part-to-part variation, repeatability variation, and reproducibility variation. However, when the observed response is a defect count, such as the number of particles or scratches detected by an AOI system, the Gaussian assumption becomes theoretically inappropriate. Moreover, in semiconductor inspection, different defect categories may be statistically correlated because they can be caused by common process conditions, such as contamination, lithography instability, etching non-uniformity, wafer handling, or localized process disturbance. Therefore, a measurement system model for wafer AOI inspection should simultaneously account for count-data characteristics, zero inflation, and correlation between defect categories.

This study addresses this research gap by developing a zero-inflated correlated defect-count R&R model for semiconductor wafer AOI inspection. The purpose is to evaluate whether multiple AOI inspection systems, inspection recipes, operators, or measurement stations produce consistent and reliable defect-count measurements under high-yield manufacturing conditions. The proposed model extends conventional R&R analysis by replacing the normal measurement error assumption with a zero-inflated bivariate Poisson structure. The bivariate Poisson component captures the dependence between two related defect categories, while the zero-inflation component captures the excess probability of observing defect-free wafers or regions. The resulting framework provides a more suitable basis for estimating repeatability, reproducibility, total measurement variation, and the percentage contribution of measurement system variation. Figure 1 illustrates the research framework of the study.

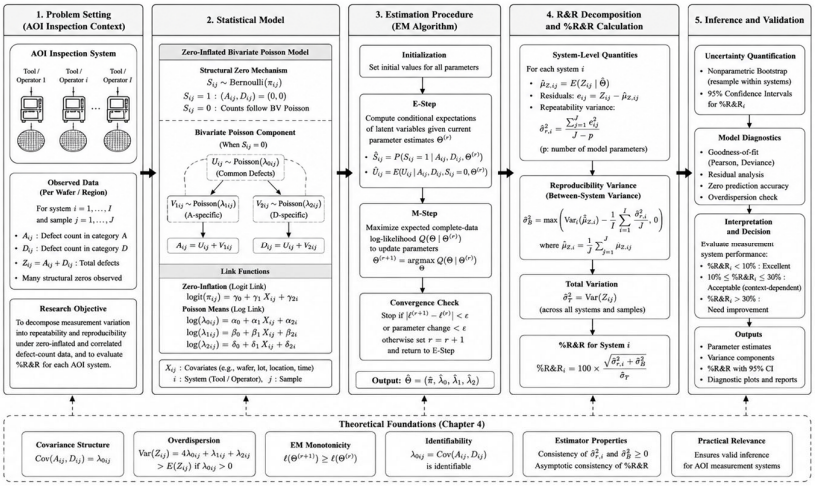

Figure 1. The Research Framework of the Measurement System R&R Analysis for Zero-Inflated Correlated Defect Counts in Semiconductor Wafer AOI Inspection.

 doi:10.20944/preprints202605.0948.v1

Let $i = 1, \ldots, I$ denote the index of the AOI measurement system, inspection tool, inspection recipe, or operator. Let $j = 1, \ldots, J$ denote the index of the inspected wafer, die, or inspection region. For each system i and item j, two defect-count responses are observed. The first defect count is denoted by A_{ij}, representing one defect category such as particle contamination. The second defect count is denoted by D_{ij}, representing another defect category such as scratch defects, pattern defects, or structural anomalies. The total observed defect count is defined as

$$Z_{ij} = A_{ij} + D_{ij}. \tag{1}$$

The main research problem is to determine how much of the observed variation in Z_{ij} is attributable to within-system measurement variability and how much is attributable to between-system measurement variability. In the language of measurement system analysis, the within-system variation corresponds to repeatability, whereas the between-system variation corresponds to reproducibility. A reliable AOI measurement system should exhibit low repeatability variation and low reproducibility variation relative to the total observed variation.

3.2. Classical R&R Model and Its Limitation

The classical measurement system model can be written as

$$Z_{ij} = \mu + B_i + \varepsilon_{ij}. \tag{2}$$

where Z_{ij} is the measurement obtained from system i on item j, μ is the overall mean measurement level, B_i is the systematic deviation associated with measurement system i, and ε_{ij} is the random within-system measurement error. In the classical setting, it is usually assumed that

$$B_i \sim N\left(0, \sigma_B^2\right). \tag{3}$$

and

$$\varepsilon_{ij} \sim N(0, \sigma_r^2). \tag{4}$$

where σ_B^2 represents the reproducibility variance and σ_r^2 represents the repeatability variance. The total measurement system variance is therefore

$$\sigma_R^2 = \sigma_B^2 + \sigma_r^2. \tag{5}$$

If σ_T^2 denotes the total observed variance, the classical percentage R&R is defined as

$$\%R\&R = 100 \times \frac{\sqrt{\sigma_B^2 + \sigma_r^2}}{\sigma_T}. \tag{6}$$

Although this model is simple and widely used, it is not fully appropriate for semiconductor AOI defect-count data. The observed response Z_{ij} is discrete and non-negative, while the normal model assumes a continuous response over the entire real line. Furthermore, the probability of observing $Z_{ij} = 0$ in high-yield wafer inspection can be substantially larger than the probability predicted by a standard Poisson or normal model. A conventional R&R model may therefore misrepresent the true measurement uncertainty, particularly when most inspected wafers or regions are defect-free. To overcome this limitation, this study formulates the AOI measurement process using a zero-inflated bivariate Poisson model.

3.3. Bivariate Poisson Representation of Correlated Defect Counts

To describe the correlation between two defect categories, this study introduces three latent Poisson components. Let U_{ij}, V_{1ij}, and V_{2ij} be independent latent Poisson random variables conditional on their intensity parameters. Specifically,

$$\%R\&R = 100 \times \frac{\sqrt{\sigma_B^2 + \sigma_r^2}}{\sigma_T}, \tag{7}$$

$$U_{ij} \sim \text{Poisson}\left(\lambda_{0ij}\right), \text{ and} \tag{8}$$

$$V_{2ij} \sim \text{Poisson}\left(\lambda_{2ij}\right). \tag{9}$$

 doi:10.20944/preprints202605.0948.v1

The variable U_{ij} represents a common latent defect source that contributes simultaneously to both observed defect categories. In a semiconductor AOI context, this common component may be caused by shared process disturbance, local contamination, wafer handling instability, or inspection sensitivity. The variable V_{1ij} represents the defect source specific to category A, while V_{2ij} represents the defect source specific to category D.

The observed defect counts are defined by

$$A_{ij} = U_{ij} + V_{1ij} \tag{10}$$

and

$$D_{ij} = U_{ij} + V_{2ij}. \tag{11}$$

This construction implies that both A_{ij} and D_{ij} contain the shared component U_{ij}. Therefore, the covariance between the two defect categories is

$$\mathrm{Cov}(A_{ij}, D_{ij}) = \mathrm{Var}(U_{ij}) = \lambda_{0ij}. \tag{12}$$

This result provides a clear interpretation of λ_{0ij}. A larger λ_{0ij} indicates stronger correlation between the two observed defect categories, while $\lambda_{0ij} = 0$ implies that the two categories are conditionally independent. Under this structure, the marginal means are

$$E(A_{ij}) = \lambda_{0ij} + \lambda_{1ij}, \tag{13}$$

and

$$E(D_{ij}) = \lambda_{0ij} + \lambda_{2ij}. \tag{14}$$

The marginal variances are

$$\mathrm{Var}(A_{ij}) = \lambda_{0ij} + \lambda_{1ij} \tag{15}$$

and

$$\mathrm{Var}(D_{ij}) = \lambda_{0ij} + \lambda_{2ij}. \tag{16}$$

The expected total defect count is

$$E(Z_{ij}) = E(A_{ij} + D_{ij}) = 2\lambda_{0ij} + \lambda_{1ij} + \lambda_{2ij}. \tag{17}$$

This bivariate Poisson structure is useful for semiconductor wafer AOI inspection because different defect classes are rarely fully independent. For example, particles may increase the probability of surface irregularities, and process non-uniformity may simultaneously generate multiple defect categories in the same wafer region. Therefore, the shared latent component U_{ij} provides a mathematically interpretable mechanism for correlated defect counts.

The joint probability mass function of the bivariate Poisson model is given by

$$P(A_{ij} = a, D_{ij} = d) = \exp\left(-(\lambda_{0ij} + \lambda_{1ij} + \lambda_{2ij})\right) \sum_{u=0}^{\min(a,d)} \frac{\lambda_{0ij}^{u} \lambda_{1ij}^{a-u} \lambda_{2ij}^{d-u}}{u!\,(a-u)!\,(d-u)!}. \tag{18}$$

In this expression, u denotes the possible value of the latent common component U_{ij}. Since U_{ij} contributes to both A_{ij} and D_{ij}, it cannot exceed either observed count. Thus, the summation range is from 0 to $\min(a, d)$.

3.4. Covariate-Dependent Defect Intensities

In semiconductor manufacturing, defect generation is often influenced by process and inspection conditions. Let x_{ij} denote a process-related or measurementrelated covariate for item j inspected by system i. This covariate may represent wafer-level process variation, inspection sensitivity, local pattern density, environmental variation, or normalized process deviation. To ensure that the intensity parameters are positive, the model uses log-linear intensity functions.

The shared defect intensity is specified as

$$\lambda_{0ij} = \exp(\alpha_0 + \alpha_1 x_{ij}), \tag{19}$$

where α_0 is the baseline log-intensity of the common defect source and α_1 measures the effect of the covariate x_{ij} on the common defect source. The category-specific intensity for defect category A is specified as

 doi:10.20944/preprints202605.0948.v1

$$\lambda_{1ij} = \exp(\beta_0 + b_{1i} + \beta_1 x_{ij}), \tag{20}$$

where β_0 is the baseline log-intensity of category A, β_1 measures the covariate effect on category A, and b_{1i} is the system-specific deviation associated with measurement system i. Similarly, the category-specific intensity for defect category D is specified as

$$\lambda_{2ij} = \exp(\gamma_0 + b_{2i} + \gamma_1 x_{ij}), \tag{21}$$

where γ_0 is the baseline log-intensity of category D, γ_1 measures the covariate effect on category D, and b_{2i} is the system-specific deviation associated with defect category D.

The terms b_{1i} and b_{2i} are important for measurement system analysis. They represent systematic differences among AOI tools, inspection recipes, operators, or measurement stations. In a random-effects formulation, they may be assumed to follow normal distributions:

$$b_{1i} \sim N(0, \sigma_{b1}^2) \tag{22}$$

and

$$b_{2i} \sim N(0, \sigma_{b2}^2). \tag{23}$$

The variances σ_{b1}^2 and σ_{b2}^2 capture between-system variability in the two defect categories. These quantities are directly related to reproducibility because they measure how much the expected defect count changes across measurement systems after accounting for the covariate x_{ij}.

3.5. Zero-Inflation Mechanism

High-yield semiconductor manufacturing often produces many wafers or inspection regions with no observable defects. A standard bivariate Poisson model may underestimate the probability of observing $(A_{ij}, D_{ij}) = (0,0)$. To address this problem, the model introduces a latent zero-inflation state variable S_{ij}, defined as

$$S_{ij} = \begin{cases} 1, & \text{if the observation comes from the structural-zero state} \\ 0, & \text{if the observation comes from the bivariate Poisson state.} \end{cases} \tag{24}$$

The structural-zero state represents observations that are defect-free due to high process quality, inspection region stability, or other conditions that make defects extremely unlikely. Let $S_{ij} \sim \text{Bernoulli}(\pi_{ij})$ where π_{ij} is the probability that the observation belongs to the structural-zero state. When $S_{ij} = 1$, the observed defect counts are forced to be $(A_{ij}, D_{ij}) = (0,0)$

When $S_{ij} = 0$, the observed counts are generated from the bivariate Poisson model described above. The zero-inflation probability may be modeled as a constant π, or it may depend on covariates through a logistic function:

$$\pi_{ij} = \frac{\exp(\eta_0 + \eta_1 x_{ij})}{1 + \exp(\eta_0 + \eta_1 x_{ij})}. \tag{25}$$

Here, η_0 represents the baseline tendency toward structural-zero observations, while η_1 measures how the covariate x_{ij} affects the probability of observing a defect-free wafer or region.

The resulting zero-inflated bivariate Poisson probability is therefore

$$P(A_{ij} = 0, D_{ij} = 0) = \pi_{ij} + (1 - \pi_{ij})\exp\left(-(\lambda_{0ij} + \lambda_{1ij} + \lambda_{2ij})\right). \tag{26}$$

For nonzero observations, where $a + d > 0$, the probability is

$$P(A_{ij} = a, D_{ij} = d) = (1 - \pi_{ij}) f_{BP}(a, d; \lambda_{0ij}, \lambda_{1ij}, \lambda_{2ij}). \tag{27}$$

where $f_{BP}(\cdot)$ denotes the bivariate Poisson probability mass function.

3.6. Observed-Data Likelihood

Let the full parameter vector be denoted by

$$\Theta = (\alpha_0, \alpha_1, \beta_0, \beta_1, \gamma_0, \gamma_1, \eta_0, \eta_1, b_{1i}, b_{2i}). \tag{28}$$

For notational simplicity, define

$$f_{ij}(a, d) = f_{BP}(a, d; \lambda_{0ij}, \lambda_{1ij}, \lambda_{2ij}). \tag{29}$$

The observed-data likelihood is

 doi:10.20944/preprints202605.0948.v1

$$L(\Theta) = \prod_{i=1}^{I}\prod_{j=1}^{J}\left[\pi_{ij} + (1-\pi_{ij})f_{ij}(0,0)\right]^{\mathbb{I}(a_{ij}=0,d_{ij}=0)} \times \left[(1-\pi_{ij})f_{ij}(a_{ij},d_{ij})\right]^{\mathbb{I}(a_{ij}+d_{ij}>0)}. \tag{30}$$

Taking logarithms gives the observed-data log-likelihood

$$\ell(\Theta) = \sum_{i=1}^{I}\sum_{j=1}^{J}\mathbb{I}(a_{ij}=0,d_{ij}=0)\log\left[\pi_{ij} + (1-\pi_{ij})f_{ij}(0,0)\right] + \sum_{i=1}^{I}\sum_{j=1}^{J}\mathbb{I}(a_{ij}+d_{ij}>0)\left[\log(1-\pi_{ij}) + \log f_{ij}(a_{ij},d_{ij})\right]. \tag{31}$$

Direct maximization of this likelihood is difficult because the latent structural-zero indicator S_{ij} and the shared Poisson component U_{ij} are unobserved. Therefore, this study adopts an expectation-maximization algorithm to estimate the model parameters.

3.7. Complete-Data Likelihood

The EM algorithm is constructed by treating S_{ij} and U_{ij} as missing data. If these latent variables were observed, the complete-data likelihood would have a simpler form. Conditional on $S_{ij} = 0$, the latent components satisfy

$$U_{ij} = u_{ij}, \tag{32}$$

$$V_{1ij} = a_{ij} - u_{ij}, \text{ and} \tag{33}$$

$$V_{2ij} = d_{ij} - u_{ij}. \tag{34}$$

The complete-data likelihood contribution for observation (i,j) is

$$\pi_{ij}^{S_{ij}}\left[(1-\pi_{ij})P(U_{ij}=u_{ij})P(V_{1ij}=a_{ij}-u_{ij})P(V_{2ij}=d_{ij}-u_{ij})\right]^{1-S_{ij}}. \tag{35}$$

The corresponding complete-data log-likelihood is

$$\begin{aligned}\ell_c(\Theta) = &\sum_{i,j} S_{ij}\log(\pi_{ij}) + (1-S_{ij})\log(1-\pi_{ij}) \\ &+ \sum_{i,j}(1-S_{ij})\left[-\lambda_{0ij} + u_{ij}\log(\lambda_{0ij}) - \log(u_{ij}!)\right] \\ &+ \sum_{i,j}(1-S_{ij})\left[-\lambda_{1ij} + (a_{ij}-u_{ij})\log(\lambda_{1ij}) - \log\left((a_{ij}-u_{ij})!\right)\right] \\ &+ \sum_{i,j}(1-S_{ij})\left[-\lambda_{2ij} + (d_{ij}-u_{ij})\log(\lambda_{2ij}) - \log\left((d_{ij}-u_{ij})!\right)\right].\end{aligned} \tag{36}$$

Since the factorial terms do not involve the parameters, the parameter-dependent part can be written as

$$\begin{aligned}\ell_c(\Theta) \propto &\sum_{i,j} S_{ij}\log(\pi_{ij}) + (1-S_{ij})\log(1-\pi_{ij}) \\ &+ \sum_{i,j}(1-S_{ij})\left[-\lambda_{0ij} + u_{ij}\log(\lambda_{0ij})\right] \\ &+ \sum_{i,j}(1-S_{ij})\left[-\lambda_{1ij} + (a_{ij}-u_{ij})\log(\lambda_{1ij})\right] \\ &+ \sum_{i,j}(1-S_{ij})\left[-\lambda_{2ij} + (d_{ij}-u_{ij})\log(\lambda_{2ij})\right].\end{aligned} \tag{37}$$

This expression provides the basis for the EM algorithm.

 doi:10.20944/preprints202605.0948.v1

3.8. E-Step of the EM Algorithm

At iteration r, suppose that the current parameter estimate is $\Theta^{(r)}$. The E-step computes the conditional expectation of the complete-data log-likelihood given the observed data and current parameters:

$$Q(\Theta \mid \Theta^{(r)}) = E_{\Theta^{(r)}}[\ell_c(\Theta) \mid A_{ij}, D_{ij}, x_{ij}]$$

The first quantity required in the E-step is the posterior probability that an observed zero is a structural zero. Define

$$\tau_{ij}^{(r)} = E(S_{ij} \mid A_{ij} = a_{ij}, D_{ij} = d_{ij}; \Theta^{(r)})$$

If $(a_{ij}, d_{ij}) \neq (0,0)$, then the observation cannot come from the structural-zero state, so $\tau_{ij}^{(r)} = 0$. However, if $(a_{ij}, d_{ij}) = (0,0)$, then Bayes' theorem gives

$$\tau_{ij}^{(r)} = \frac{\pi_{ij}^{(r)}}{\pi_{ij}^{(r)} + (1 - \pi_{ij}^{(r)}) \exp\left(-\left(\lambda_{0ij}^{(r)} + \lambda_{1ij}^{(r)} + \lambda_{2ij}^{(r)}\right)\right)}. \tag{38}$$

The value $\tau_{ij}^{(r)}$ is the estimated probability that a zero observation is a structural zero. Its complement, $\omega_{ij}^{(r)} = 1 - \tau_{ij}^{(r)}$ is the posterior probability that the observation belongs to the bivariate Poisson state.

The second quantity required in the E -step is the conditional expectation of the shared latent component U_{ij}. Given $A_{ij} = a, D_{ij} = d$, and $S_{ij} = 0$, the possible values of U_{ij} are $u = 0,1, \ldots, \min(a, d)$.

The conditional probability of $U_{ij} = u$ is

$$P(U_{ij} = u \mid A_{ij} = a, D_{ij} = d, S_{ij} = 0) = \frac{\frac{\lambda_{0ij}^{u} \lambda_{1ij}^{a-u} \lambda_{2ij}^{d-u}}{u!\,(a-u)!\,(d-u)!}}{\sum_{k=0}^{\min(a,d)} \frac{\lambda_{0ij}^{k} \lambda_{1ij}^{a-k} \lambda_{2ij}^{d-k}}{k!\,(a-k)!\,(d-k)!}}. \tag{39}$$

Using the current parameter estimates, define

$$p_{iju}^{(r)} = P(U_{ij} = u \mid A_{ij} = a_{ij}, D_{ij} = d_{ij}, S_{ij} = 0; \Theta^{(r)}). \tag{40}$$

Then the conditional expectation of U_{ij} is

$$m_{ij}^{(r)} = E(U_{ij} \mid A_{ij}, D_{ij}, S_{ij} = 0; \Theta^{(r)}) = \sum_{u=0}^{\min(a_{ij}, d_{ij})} u p_{iju}^{(r)}. \tag{41}$$

Because the bivariate Poisson component is active only with posterior probability $\omega_{ij}^{(r)}$, the expected latent counts used in the M-step are

$$\tilde{u}_{ij}^{(r)} = \omega_{ij}^{(r)} m_{ij}^{(r)}, \tag{42}$$

$$\tilde{v}_{1ij}^{(r)} = \omega_{ij}^{(r)} (a_{ij} - m_{ij}^{(r)}), \text{ and} \tag{43}$$

$$\tilde{v}_{2ij}^{(r)} = \omega_{ij}^{(r)} (d_{ij} - m_{ij}^{(r)}). \tag{44}$$

Here, $\tilde{u}_{ij}^{(r)}$ is the expected shared defect count, $\tilde{v}_{1ij}^{(r)}$ is the expected category-A specific defect count, and $\tilde{v}_{2ij}^{(r)}$ is the expected category-D-specific defect count.

3.9. M-Step of the EM Algorithm

The M -step updates the parameter vector by maximizing the expected complete data log-likelihood:

$$\Theta^{(r+1)} = \arg\max_{\Theta} Q(\Theta \mid \Theta^{(r)}). \tag{45}$$

The function $Q(\Theta \mid \Theta^{(r)})$ can be decomposed into four parts:

$$Q(\Theta \mid \Theta^{(r)}) = Q_\pi + Q_0 + Q_1 + Q_2, \tag{46}$$

 doi:10.20944/preprints202605.0948.v1

where Q_π corresponds to the zero-inflation model, Q_0 corresponds to the common Poisson intensity λ_{0ij}, Q_1 corresponds to the category-A intensity λ_{1ij}, and Q_2 corresponds to the category-D intensity λ_{2ij}.

The zero-inflation component is

$$Q_\pi = \sum_{i,j} \left[\tau_{ij}^{(r)} \log(\pi_{ij}) + \omega_{ij}^{(r)} \log(1-\pi_{ij})\right]. \tag{47}$$

If $\pi_{ij} = \pi$ is assumed constant, the update has the closed form

$$\pi^{(r+1)} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} \tau_{ij}^{(r)}. \tag{48}$$

If π_{ij} is modeled through a logistic regression, the M-step for η is equivalent to maximizing a weighted Bernoulli log-likelihood with fractional response $\tau_{ij}^{(r)}$.

The common defect intensity component is

$$Q_0 = \sum_{i,j} \left[-\omega_{ij}^{(r)} \lambda_{0ij} + \tilde{u}_{ij}^{(r)} \log\, \lambda_{0ij}\right]. \tag{49}$$

Substituting $\lambda_{0ij} = \exp\left(\alpha_0 + \alpha_1 x_{ij}\right)$ we obtain

$$Q_0 = \sum_{i,j} \left[-\omega_{ij}^{(r)} \exp\left(\alpha_0 + \alpha_1 x_{ij}\right) + \tilde{u}_{ij}^{(r)} \left(\alpha_0 + \alpha_1 x_{ij}\right)\right]. \tag{50}$$

This has the same structure as a weighted Poisson regression. Therefore, α_0 and α_1 can be updated by numerical maximization or by an iteratively reweighted least squares procedure.

Similarly, the category-A intensity component is

$$Q_1 = \sum_{i,j} \left[-\omega_{ij}^{(r)} \lambda_{1ij} + \tilde{v}_{1ij}^{(r)} \log\, \lambda_{1ij}\right]. \tag{51}$$

Using $\lambda_{1ij} = \exp\left(\beta_0 + b_{1i} + \beta_1 x_{ij}\right)$, the corresponding objective becomes

$$Q_1 = \sum_{i,j} \left[-\omega_{ij}^{(r)} \exp\left(\beta_0 + b_{1i} + \beta_1 x_{ij}\right) + \tilde{v}_{1ij}^{(r)} \left(\beta_0 + b_{1i} + \beta_1 x_{ij}\right)\right]. \tag{52}$$

The update for β_0, β_1, and b_{1i} may be obtained through a Poisson regression with system effects or through a generalized linear mixed model if the system effects are treated as random.

The category-D intensity component is

$$Q_2 = \sum_{i,j} \left[-\omega_{ij}^{(r)} \lambda_{2ij} + \tilde{v}_{2ij}^{(r)} \log\, \lambda_{2ij}\right]. \tag{53}$$

Since $\lambda_{2ij} = \exp\left(\gamma_0 + b_{2i} + \gamma_1 x_{ij}\right)$ the objective is

$$Q_2 = \sum_{i,j} \left[-\omega_{ij}^{(r)} \exp\left(\gamma_0 + b_{2i} + \gamma_1 x_{ij}\right) + \tilde{v}_{2ij}^{(r)} \left(\gamma_0 + b_{2i} + \gamma_1 x_{ij}\right)\right]. \tag{53}$$

Thus, γ_0, γ_1, and b_{2i} can be updated in the same manner as the category-A intensity parameters.

3.10. Intercept-Only Closed-Form EM Updates

For theoretical clarity, consider a simplified intercept-only version of the model in which $\lambda_0, \lambda_1, \lambda_2$, and π are constants. In this case, the M-step has closed-form solutions. Using the E-step quantities, the updates are

$$\lambda_0^{(r+1)} = \frac{\Sigma_{i,j}\, \tilde{u}_{ij}^{(r)}}{\Sigma_{i,j}\, \omega_{ij}^{(r)}}, \tag{54}$$

$$\lambda_1^{(r+1)} = \frac{\Sigma_{i,j}\, \tilde{v}_{1ij}^{(r)}}{\Sigma_{i,j}\, \omega_{ij}^{(r)}}, \tag{55}$$

$$\lambda_2^{(r+1)} = \frac{\Sigma_{i,j}\, \tilde{v}_{2ij}^{(r)}}{\Sigma_{i,j}\, \omega_{ij}^{(r)}}, \text{ and} \tag{56}$$

 doi:10.20944/preprints202605.0948.v1

$$\pi^{(r+1)} = \frac{\sum_{i,j} \tau_{ij}^{(r)}}{IJ}. \tag{57}$$

These formulas show the intuition of the EM algorithm. The new estimate of λ_0 is the posterior expected number of common defects divided by the expected number of observations belonging to the bivariate Poisson state. Similarly, λ_1 and λ_2 are updated using the expected category-specific defect counts. The zeroinflation probability π is updated as the average posterior probability of structuralzero membership.

3.11. Connection Between EM Estimation and R&R Variance Decomposition

After the EM algorithm converges, the fitted model provides estimates of $\hat{\lambda}_{0ij}, \hat{\lambda}_{1ij}, \hat{\lambda}_{2ij}, \hat{\pi}_{ij}$. The fitted expected defect counts are

$$\hat{\mu}_{A,ij} = (1 - \hat{\pi}_{ij})(\hat{\lambda}_{0ij} + \hat{\lambda}_{1ij}), \tag{58}$$

$$\hat{\mu}_{D,ij} = (1 - \hat{\pi}_{ij})(\hat{\lambda}_{0ij} + \hat{\lambda}_{2ij}), \text{ and} \tag{59}$$

$$\hat{\mu}_{Z,ij} = \hat{\mu}_{A,ij} + \hat{\mu}_{D,ij}. \tag{60}$$

Therefore,

$$\hat{\mu}_{Z,ij} = (1 - \hat{\pi}_{ij})(2\hat{\lambda}_{0ij} + \hat{\lambda}_{1ij} + \hat{\lambda}_{2ij}). \tag{61}$$

The residual for the total defect count is defined as

$$e_{ij} = Z_{ij} - \hat{\mu}_{Z,ij}. \tag{62}$$

For measurement system i, the repeatability variance is estimated by

$$\hat{\sigma}_{r,i}^2 = \frac{\sum_{j=1}^{J} e_{ij}^2}{J - p}, \tag{63}$$

where p denotes the effective number of estimated parameters involved in the fitted model for the system-level residual calculation. This variance measures the unexplained within-system variation after accounting for zero inflation, defect correlation, and covariate effects.

The reproducibility variance captures systematic differences across AOI measurement systems. Let $\bar{\mu}_{Z,i} = \frac{1}{J}\sum_{j=1}^{J} \hat{\mu}_{Z,ij}$ be the model-adjusted mean defect count for system i. A method-of-moments estimate of between-system variance is

$$\hat{\sigma}_B^2 = \max\left\{\mathrm{Var}_i(\bar{\mu}_{Z,i}) - \frac{1}{I}\sum_{i=1}^{I} \frac{\hat{\sigma}_{r,i}^2}{J}, 0\right\}. \tag{64}$$

Alternatively, a Mandel-type equation can be used. Define

$$W_i(\sigma_B^2) = \frac{\bar{Z}_i}{\sqrt{\sigma_B^2 + \hat{\sigma}_{r,i}^2/J}} \tag{65}$$

where $\bar{Z}_i = \frac{1}{J}\sum_{j=1}^{J} Z_{ij}$.

The between-system variance σ_B^2 is estimated by solving $\mathrm{Var}_i[W_i(\sigma_B^2)] = 1$

Equivalently, one solves

$$g(\sigma_B^2) = \mathrm{Var}_i\left[\frac{\bar{Z}_i}{\sqrt{\sigma_B^2 + \hat{\sigma}_{r,i}^2/J}}\right] - 1 = 0. \tag{66}$$

The resulting solution is denoted by $\hat{\sigma}_B^2$. This quantity represents reproducibility variation among AOI inspection systems.

3.12. Percent R&R Under the Proposed Model

For system i, the total measurement system variance is

$$\hat{\sigma}_{R,i}^2 = \hat{\sigma}_{r,i}^2 + \hat{\sigma}_B^2. \tag{67}$$

The total observed standard deviation of the defect-count response is

 doi:10.20944/preprints202605.0948.v1

$$\hat{\sigma}_T = SD(Z_{ij}). \tag{68}$$

The proposed model-based percent R&R is therefore

$$\%R\&R_i = 100 \times \frac{\sqrt{\hat{\sigma}_{r,i}^2 + \hat{\sigma}_B^2}}{\hat{\sigma}_T}. \tag{69}$$

This index measures the percentage of total observed variation attributable to measurement system variation. A smaller value indicates that the AOI measurement system is more reliable relative to the total variation observed in the defect-count data. A larger value indicates that the inspection process itself contributes substantially to the observed variability and may require calibration, recipe adjustment, tool maintenance, operator training, or measurement system redesign.

3.13. Bootstrap Inference for Percent R&R

Because the proposed percent R&R estimator depends on nonlinear functions of estimated zero-inflated bivariate Poisson parameters, closed-form standard errors may be difficult to obtain. Therefore, this study uses bootstrap resampling to construct confidence intervals. For each bootstrap replication $b = 1, \dots, B$, observations are resampled with replacement within each measurement system. The proposed model is then re-estimated, and a bootstrap estimate of percent R&R is obtained:

$$\%R\&R_i^{(b)}. \tag{70}$$

After B bootstrap replications, the 95% confidence interval is computed as

$$CI_{i,95\%} = \left[q_{0.025}\left(\%R\&R_i^{(1)}, \dots, \%R\&R_i^{(B)}\right), q_{0.975}\left(\%R\&R_i^{(1)}, \dots, \%R\&R_i^{(B)}\right)\right]. \tag{71}$$

The bootstrap confidence interval provides an uncertainty assessment for the measurement system evaluation. If the confidence interval lies entirely below an acceptable threshold, the measurement system may be considered reliable. If the interval overlaps or exceeds a critical threshold, further investigation is required.

3.14. Summary of the Proposed Mathematical Framework

This section develops a mathematical framework for measurement system R&R analysis in semiconductor wafer AOI inspection. The proposed model begins by recognizing that AOI defect observations are discrete count data rather than continuous normal measurements. It then introduces a bivariate Poisson structure to model correlated defect categories and a zero-inflation component to account for the large number of defect-free wafers or regions observed in high-yield manufacturing. The EM algorithm is developed to estimate the latent structural zero probability and the latent common defect component. After model estimation, the fitted defect-count means are used to compute residual-based repeatability variation, while between-system differences are used to estimate reproducibility variation. Finally, percent R&R and bootstrap confidence intervals are calculated to quantify the reliability of the AOI measurement system.

The resulting framework provides a statistically coherent extension of classical Gauge R&R analysis to zero-inflated, correlated, count-type defect measurements. It is therefore particularly suitable for semiconductor wafer AOI inspection, where high-yield manufacturing, correlated defect mechanisms, and measurement system consistency are all central concerns.

4. Theoretical Properties and Analytical Foundations of the Proposed Framework

4.1. Introduction

The previous section established the zero-inflated bivariate Poisson measurement system framework for semiconductor wafer AOI inspection and developed the associated expectation-maximization estimation procedure. Although the proposed framework provides a flexible statistical structure for correlated zero-inflated defect counts, a rigorous theoretical foundation is also necessary

 doi:10.20944/preprints202605.0948.v1

in order to justify the validity of the estimation procedure and the interpretation of the resulting R&R measures.

In semiconductor manufacturing environments, AOI defect inspection systems are frequently deployed for high-volume and high-yield process monitoring. Consequently, the statistical reliability of the proposed model is essential because inaccurate estimation of repeatability or reproducibility may lead to incorrect conclusions regarding inspection tool performance, calibration stability, or manufacturing quality control. The purpose of this section is therefore to establish several important theoretical properties of the proposed framework. In particular, this section presents propositions and mathematical proofs related to the covariance structure of the bivariate Poisson model, the identifiability of the latent common defect component, the monotonicity of the EM algorithm, the consistency of the repeatability variance estimator, the non-negativity of the reproducibility estimator, and the asymptotic behavior of the proposed Percent R&R measure.

The propositions developed in this section provide the analytical basis for understanding why the proposed framework is suitable for semiconductor wafer AOI inspection under high-yield manufacturing conditions.

4.2. Covariance Structure of the Bivariate Poisson Defect Model

The proposed framework models two correlated defect categories through the latent Poisson decomposition $A_{ij} = U_{ij} + V_{1ij}$ and $D_{ij} = U_{ij} + V_{2ij}$, where $U_{ij} \sim \text{Poisson}(\lambda_{0ij})$, $V_{1ij} \sim \text{Poisson}(\lambda_{1ij})$, and $V_{2ij} \sim \text{Poisson}(\lambda_{2ij})$ are mutually independent conditional on the model parameters. The following proposition establishes the covariance structure of the observed defect counts.

Proposition 1. *Covariance Structure of the Bivariate Poisson Model Under the proposed bivariate Poisson defect-count model,* $Cov(A_{ij}, D_{ij}) = \lambda_{0ij}$.

Furthermore, $Corr(A_{ij}, D_{ij}) = \frac{\lambda_{0ij}}{\sqrt{(\lambda_{0ij}+\lambda_{1ij})(\lambda_{0ij}+\lambda_{2ij})}}$.

Proof of Proposition 1:

From the model definition, $A_{ij} = U_{ij} + V_{1ij}$ *and* $D_{ij} = U_{ij} + V_{2ij}$.

Therefore, $Cov(A_{ij}, D_{ij}) = Cov(U_{ij} + V_{1ij}, U_{ij} + V_{2ij})$.

Using bi-linearity of covariance,

$Cov(A_{ij}, D_{ij}) = Var(U_{ij}) + Cov(U_{ij}, V_{2ij}) + Cov(V_{1ij}, U_{ij}) + Cov(V_{1ij}, V_{2ij})$.

Because the latent variables are mutually independent, $Cov(U_{ij}, V_{2ij}) = 0$, $Cov(V_{1ij}, U_{ij}) = 0$, *and* $Cov(V_{1ij}, V_{2ij}) = 0$.

Hence, $Cov(A_{ij}, D_{ij}) = Var(U_{ij})$.

Since $U_{ij} \sim Poisson(\lambda_{0ij})$, *the variance of a Poisson random variable equals its mean, implying* $Var(U_{ij}) = \lambda_{0ij}$. *Therefore,* $Cov(A_{ij}, D_{ij}) = \lambda_{0ij}$.

Next, observe that $Var(A_{ij}) = \lambda_{0ij} + \lambda_{1ij}$ *and* $Var(D_{ij}) = \lambda_{0ij} + \lambda_{2ij}$.

Therefore, $Corr(A_{ij}, D_{ij}) = \frac{Cov(A_{ij}, D_{ij})}{\sqrt{Var(A_{ij})Var(D_{ij})}}$,

which yields $Corr(A_{ij}, D_{ij}) = \frac{\lambda_{0ij}}{\sqrt{(\lambda_{0ij}+\lambda_{1ij})(\lambda_{0ij}+\lambda_{2ij})}}$.

The result of Proposition 1 has important engineering implications for semiconductor wafer AOI inspection. The parameter λ_{0ij} directly controls the dependence between different defect categories. In practice, correlated defect behavior may arise from common process disturbances such as contamination, lithography instability, wafer handling variation, or localized process drift. Therefore, the proposed framework allows the dependence structure of defect counts to be explicitly quantified rather than ignored.

4.3. Mean and Variance of the Total Defect Count

The total observed defect count is defined as $Z_{ij} = A_{ij} + D_{ij}$. The following proposition derives the first two moments of the total defect count.

 doi:10.20944/preprints202605.0948.v1

Proposition 2. *Mean and Variance of the Total Defect Count Under the proposed bivariate Poisson framework,* $E(Z_{ij}) = 2\lambda_{0ij} + \lambda_{1ij} + \lambda_{2ij}$ *and* $Var(Z_{ij}) = 4\lambda_{0ij} + \lambda_{1ij} + \lambda_{2ij}$.

Proof of Proposition 2:
Since $Z_{ij} = A_{ij} + D_{ij}$, *the expectation is* $E(Z_{ij}) = E(A_{ij}) + E(D_{ij})$.
Using the marginal expectations, $E(A_{ij}) = \lambda_{0ij} + \lambda_{1ij}$ *and* $E(D_{ij}) = \lambda_{0ij} + \lambda_{2ij}$,
we obtain $E(Z_{ij}) = 2\lambda_{0ij} + \lambda_{1ij} + \lambda_{2ij}$.
Next, $Var(Z_{ij}) = Var(A_{ij} + D_{ij})$.
Using the variance expansion formula, $Var(Z_{ij}) = Var(A_{ij}) + Var(D_{ij}) + 2Cov(A_{ij}, D_{ij})$.
Substituting the previous results, $Var(Z_{ij}) = (\lambda_{0ij} + \lambda_{1ij}) + (\lambda_{0ij} + \lambda_{2ij}) + 2\lambda_{0ij}$.
Thus, $Var(Z_{ij}) = 4\lambda_{0ij} + \lambda_{1ij} + \lambda_{2ij}$.

Proposition 2 shows that the variance of the total defect count exceeds its mean whenever $\lambda_{0ij} > 0$. Therefore, the proposed model naturally accommodates over dispersion induced by correlated defect mechanisms. This property is particularly important in semiconductor manufacturing because AOI defect data frequently exhibit over dispersed behavior relative to a standard Poisson model.

4.4. Monotonicity of the EM Algorithm

The EM algorithm is used to estimate the latent structural-zero states and the latent common defect counts. The following proposition establishes the monotonicity property of the proposed EM estimation procedure.

Proposition 3. *Monotonicity of the EM Algorithm*
Let $\Theta^{(r)}$ *denote the parameter vector at iteration* r.
If the M -step satisfies $Q(\Theta^{(r+1)} \mid \Theta^{(r)}) \geq Q(\Theta^{(r)} \mid \Theta^{(r)})$, *then the observed-data log-likelihood satisfies* $\ell(\Theta^{(r+1)}) \geq \ell(\Theta^{(r)})$.

Proof of Proposition 3:
Let Y *denote the latent variables* $Y = (S_{ij}, U_{ij})$, *and let* X *denote the observed defect counts. The observed-data likelihood is* $p(X \mid \Theta) = \sum_Y p(X, Y \mid \Theta)$.
Therefore, $\ell(\Theta) = log(p(X \mid \Theta)) = log(\sum_Y p(X, Y \mid \Theta))$.
Introduce the conditional distribution $p(Y \mid X, \Theta^{(r)})$.
Then, $\ell(\Theta) = log\left(\sum_Y p(Y \mid X, \Theta^{(r)}) \frac{p(X,Y|\Theta)}{p(Y|X,\Theta^{(r)})}\right)$.
Applying Jensen's inequality to the logarithm function yields

$$\ell(\Theta) \geq \sum_Y p(Y \mid X, \Theta^{(r)}) log\left(\frac{p(X, Y \mid \Theta)}{p(Y \mid X, \Theta^{(r)})}\right).$$

Hence, $\ell(\Theta) - \ell(\Theta^{(r)}) \geq Q(\Theta \mid \Theta^{(r)}) - Q(\Theta^{(r)} \mid \Theta^{(r)})$.
If the M-step satisfies $Q(\Theta^{(r+1)} \mid \Theta^{(r)}) \geq Q(\Theta^{(r)} \mid \Theta^{(r)})$, *then* $\ell(\Theta^{(r+1)}) \geq \ell(\Theta^{(r)})$.
Thus, the observed-data log-likelihood is non-decreasing across EM iterations.

Proposition 3 guarantees the numerical stability of the proposed estimation procedure. In semiconductor AOI applications involving large-scale wafer inspection data, stable likelihood improvement is essential because the estimation process may involve thousands or millions of inspection regions.

4.5. Identifiability of the Shared Defect Component

The common latent component U_{ij} plays a central role in the proposed framework because it represents the shared source of correlated defects. The following proposition establishes its identifiability.

Proposition 4. *Identifiability of the Common Defect Parameter*

 doi:10.20944/preprints202605.0948.v1

Suppose that the joint distribution of (A_{ij}, D_{ij}) *follows the proposed bivariate Poisson model. Then the common defect intensity parameter* λ_{0ij} *is identifiable from the covariance structure of the observed data.*

Proof of Proposition 4:
From Proposition 1, $Cov(A_{ij}, D_{ij}) = \lambda_{0ij}$. *The covariance of the observed variables is uniquely determined by the joint distribution of* (A_{ij}, D_{ij}).
Therefore, $\lambda_{0ij} = Cov(A_{ij}, D_{ij})$, *which is uniquely identifiable from the observed second-order moment structure. Hence,* λ_{0ij} *is identifiable.*

This proposition is important because it demonstrates that the shared latent defect mechanism is not merely a modeling artifact. Instead, it corresponds directly to observable covariance behavior in semiconductor defect-count data.

4.6. Non-Negativity of the Reproducibility Variance Estimator

The proposed framework estimates reproducibility variation using a method of moments or Mandel-type variance decomposition. The following proposition guarantees that the estimated between-system variance remains nonnegative.

Proposition 5. *Non-negativity of the Reproducibility Variance Estimator*
Define the between-system variance estimator as $\hat{\sigma}_B^2 = max\left\{Var_i(\bar{\mu}_{Z,i}) - \frac{1}{I}\sum_{i=1}^{I}\frac{\hat{\sigma}_{r,i}^2}{J}, 0\right\}$. *Then* $\hat{\sigma}_B^2 \geq 0$.

Proof of Proposition 5:
By definition, $\hat{\sigma}_B^2 = \max\{a, 0\}$, *where* $a = \mathrm{Var}_i(\bar{\mu}_{Z,i}) - \frac{1}{I}\sum_{i=1}^{I}\frac{\hat{\sigma}_{r,i}^2}{J}$.
The maximum of any real number and zero is always nonnegative. Therefore, $\hat{\sigma}_B^2 \geq 0$.

The importance of Proposition 5 lies in ensuring physical interpretability. Since variance cannot be negative in practical measurement system analysis, the proposed estimator guarantees that the estimated reproducibility component remains meaningful in semiconductor AOI applications.

4.7. Consistency of the Repeatability Variance Estimator

The repeatability variance is estimated from the residuals of the fitted zero-inflated bivariate Poisson model.

Proposition 6. *Consistency of the Repeatability Variance Estimator*
Suppose the proposed zero-inflated bivariate Poisson model is correctly specified and the parameter estimates converge to the true parameter values. Then the repeatability variance estimator $\hat{\sigma}_{r,i}^2 = \frac{\sum_{j=1}^{J} e_{ij}^2}{J-p}$ *is a consistent estimator of the true within-system variance.*

Proof of Proposition 6:
Under correct model specification, $\hat{\mu}_{Z,ij} \xrightarrow{p} \mu_{Z,ij}$, *as* $J \to \infty$
Therefore, $e_{ij} = Z_{ij} - \hat{\mu}_{Z,ij} \xrightarrow{p} Z_{ij} - \mu_{Z,ij}$.
By the law of large numbers, $\frac{1}{J}\sum_{j=1}^{J} e_{ij}^2 \xrightarrow{p} E\left[\left(Z_{ij} - \mu_{Z,ij}\right)^2\right]$.
Since $E\left[\left(Z_{ij} - \mu_{Z,ij}\right)^2\right] = \sigma_{r,i}^2$, *it follows that* $\hat{\sigma}_{r,i}^2 \xrightarrow{p} \sigma_{r,i}^2$.
Hence, the estimator is consistent.

This proposition implies that the repeatability variance estimate becomes increasingly reliable as the amount of AOI inspection data grows, which is especially important in large-scale semiconductor manufacturing environments.

 doi:10.20944/preprints202605.0948.v1

4.8. Asymptotic Behavior of the Percent R&R Estimator

The final proposition establishes the asymptotic convergence of the proposed Percent R&R estimator.

Proposition 7. *Asymptotic Consistency of the Percent R&R Estimator*
Suppose that $\hat{\sigma}_{r,i}^2 \xrightarrow{p} \sigma_{r,i}^2$ *and* $\hat{\sigma}_T \xrightarrow{p} \sigma_T$.

Then the proposed Percent R&R estimator satisfies $\widehat{\%R\&R}_i \xrightarrow{p} 100 \times \frac{\sqrt{\sigma_{r,i}^2+\sigma_B^2}}{\sigma_T}$.

Proof of Proposition 7:

The proposed estimator is $\widehat{\%R\&R}_i = 100 \times \frac{\sqrt{\hat{\sigma}_{r,i}^2+\hat{\sigma}_B^2}}{\hat{\sigma}_T}$
The square-root function and division function are continuous whenever $\sigma_T > 0$.
By the continuous mapping theorem, $\sqrt{\hat{\sigma}_{r,i}^2 + \hat{\sigma}_B^2} \xrightarrow{p} \sqrt{\sigma_{r,i}^2 + \sigma_B^2}$, *and* $\frac{\sqrt{\hat{\sigma}_{r,i}^2+\hat{\sigma}_B^2}}{\hat{\sigma}_T} \xrightarrow{p} \frac{\sqrt{\sigma_{r,i}^2+\sigma_B^2}}{\sigma_T}$.
Multiplying by 100 yields $\widehat{\%R\&R}_i \xrightarrow{p} 100 \times \frac{\sqrt{\sigma_{r,i}^2+\sigma_B^2}}{\sigma_T}$.

4.9. Summary

This section established several theoretical properties of the proposed zero inflated correlated defect-count R&R framework for semiconductor wafer AOI inspection. The propositions demonstrated that the proposed bivariate Poisson structure naturally models correlated defect categories and over dispersed defect count behavior. The theoretical analysis further established that the latent common defect component is identifiable, the EM algorithm possesses the monotonic likelihood property, the repeatability variance estimator is consistent, the reproducibility estimator remains nonnegative, and the proposed Percent R&R estimator converges asymptotically to the true measurement system variation ratio.

These theoretical properties provide strong analytical support for applying the proposed framework to semiconductor AOI inspection environments characterized by high-yield manufacturing, correlated defect mechanisms, and large-scale defect count measurement data.

5. Application and Numerical Analysis

5.1. Semiconductor Wafer AOI Inspection Scenario

To demonstrate the applicability of the proposed zero-inflated correlated defect-count R&R framework, this study considers a semiconductor wafer automated optical inspection (AOI) environment in an advanced high-yield manufacturing facility. In modern semiconductor fabrication, AOI systems are extensively deployed for inline inspection and process monitoring in order to identify particle contamination, pattern abnormalities, scratches, and structural defects on wafer surfaces. As semiconductor process technology continues to move toward smaller feature sizes and higher integration density, maintaining inspection reliability and measurement consistency across AOI systems has become increasingly important for yield management and quality assurance.

In practical semiconductor manufacturing environments, multiple AOI inspection tools are often simultaneously deployed within the same fabrication line. Although these inspection systems may operate under nominally identical inspection conditions, differences in optical sensitivity, illumination configuration, focus calibration, inspection recipes, and algorithmic thresholds may lead to systematic differences in defect-count measurements. Consequently, the same wafer region inspected by different AOI systems may produce different defect-count results. Such measurement inconsistency may directly influence process monitoring decisions, yield analysis, root-cause diagnosis, and defect excursion detection.

 doi:10.20944/preprints202605.0948.v1

Another important characteristic of semiconductor AOI data is the prevalence of zero-defect observations. Under high-yield manufacturing conditions, a large proportion of wafer regions may contain no observable defects. Therefore, the resulting inspection data typically exhibit substantial zero inflation. Moreover, multiple defect categories are frequently correlated because they may originate from common process disturbances such as contamination events, lithography instability, wafer handling variation, or localized process fluctuations. Conventional Gaussian-type Gauge R&R models are generally unable to appropriately capture these statistical characteristics.

To address these issues, this study applies the proposed zero-inflated bivariate Poisson measurement system framework to a simulated semiconductor wafer AOI inspection scenario. The numerical experiment was designed to emulate a realistic high-yield semiconductor manufacturing environment in which correlated defect mechanisms and excessive zero-defect observations coexist. Four AOI inspection systems were considered, representing different inspection sensitivities and recipe configurations. The simulated AOI systems were denoted as AOI_Tool_A_Reference, AOI_Tool_B_Calibrated, AOI_Tool_C_LowSensitivity, and AOI_Tool_D_HighSensitivity.

The inspection dataset consisted of 30 wafers, each containing 12 inspection regions, resulting in 360 observations for each AOI inspection system. Each wafer region was characterized by several process-related covariates, including process risk index, pattern density, and wafer-zone location. The wafer-zone effect was introduced because edge regions in semiconductor wafers often exhibit higher defect susceptibility due to process nonuniformity and edge-process instability. Pattern density was included to represent inspection complexity and local layout characteristics that may influence AOI sensitivity and defect-generation behavior.

To reflect high-yield manufacturing characteristics, a structural-zero probability of 0.58 was adopted in the numerical experiment. This setting implies that a substantial proportion of wafer regions belong to an underlying defect-free state. The structural-zero probability was further modeled as a function of process risk, such that higher process risk reduced the probability of observing defect-free wafer regions. This assumption is consistent with semiconductor manufacturing behavior, where unstable process conditions are more likely to generate detectable defects.

The latent defect-generation process was modeled using a zero-inflated correlated bivariate Poisson structure. Two defect categories were simultaneously considered in the AOI inspection process. The first category represented particle contamination defects, while the second category represented pattern or scratch-related defects. A shared latent Poisson component was introduced to model common defect-generation mechanisms affecting both categories simultaneously. This shared component reflects the fact that multiple defect types may originate from the same underlying process disturbance.

The latent shared defect intensity was specified as $\lambda_{0ij} = \exp(\alpha_0 + \alpha_{risk}\, x_{ij} + \alpha_{density}\, d_{ij} + \alpha_{edge}\, e_{ij})$ where x_{ij} denotes the process risk index, d_{ij} denotes pattern density, and e_{ij} denotes the wafer-edge indicator variable. The category-specific defect intensities were modeled separately for particle defects and pattern/scratch defects in order to capture different defect sensitivities across AOI systems.

The numerical experiment also introduced AOI-system-specific sensitivity offsets to emulate differences in inspection capability across measurement systems. Positive offsets represented AOI systems with higher defect-detection sensitivity, whereas negative offsets represented systems with lower sensitivity. In addition, repeatability noise was incorporated to model within-system inspection fluctuation caused by sensor noise, illumination variation, autofocus instability, and local inspection uncertainty.

The proposed EM-based estimation procedure was subsequently applied to estimate the latent defect-generation structure, repeatability variance, reproducibility variance, and Percent R&R for each AOI inspection system. Bootstrap resampling was further employed to construct confidence intervals for the proposed R&R measures. The resulting numerical analysis provides a realistic demonstration of how the proposed framework can be applied to evaluate measurement-system

 doi:10.20944/preprints202605.0948.v1

reliability in semiconductor wafer AOI inspection environments characterized by correlated defect counts and excessive zero-defect observations. The parameter values for the case are presented in Table 1.

Table 1. Parameter Values for the Case.

Parameter	Value	Interpretation
I	4	Number of AOI inspection systems
W	30	Number of wafers
R	12	Number of inspection regions per wafer
$J = W \times R$	360	Number of observations per AOI system
π_0	0.58	Baseline structural-zero probability
η_1	-1.25	Effect of process risk on zero-defect probability
α_0	-2.65	Baseline shared latent defect intensity
α_{risk}	0.85	Effect of process risk on shared latent defects
α_{density}	0.35	Effect of pattern density on shared latent defects
α_{edge}	0.30	Edge-zone effect on shared latent defects
β_0	-2.25	Baseline particle-defect intensity
β_{risk}	0.70	Effect of process risk on particle defects
β_{density}	0.25	Effect of pattern density on particle defects
β_{edge}	0.45	Edge-zone effect on particle defects
γ_0	-2.45	Baseline pattern/scratch defect intensity
γ_{risk}	0.55	Effect of process risk on pattern/scratch defects
γ_{density}	0.60	Effect of pattern density on pattern/scratch defects
γ_{edge}	0.20	Edge-zone effect on pattern/scratch defects
b_{1i}	[0.00, 0.05, -0.18, 0.22]	AOI-system sensitivity offsets for particle defects
b_{2i}	[0.00, 0.03, -0.15, 0.20]	AOI-system sensitivity offsets for pattern/scratch defects
σ_r	0.06	Repeatability noise standard deviation

5.2. Numerical Analysis Results and Discussion

This section presents the numerical analysis results of the proposed zero-inflated correlated defect-count measurement system framework under the semiconductor wafer AOI inspection scenario. The analysis focuses on the statistical characteristics of the generated defect-count data, the correlated defect behavior among defect categories, the structural-zero properties associated with high-yield manufacturing, the spatial wafer-zone effect, and the resulting measurement system repeatability and reproducibility performance.

The descriptive statistics of total defect counts across the four AOI inspection systems are summarized in Table 2. The results indicate that the average defect count per wafer region remained relatively low for all AOI systems, which is consistent with the characteristics of high-yield semiconductor manufacturing. The mean total defect counts ranged from 0.331 to 0.453 defects per wafer region, while the median defect count remained zero for all AOI systems. These results indicate that the majority of wafer regions contained no detectable defects, thereby confirming the existence of substantial zero-defect observations in the AOI inspection environment.

 doi:10.20944/preprints202605.0948.v1

Table 2. Descriptive Statistics of Total Defect Counts Across AOI Inspection Systems.

AOI_System	Sample _Size	Mean_Total _Defects	Std_Total_ Defects	Min_Total_ Defects	Median_Total _Defects	Max_Total_ Defects
AOI_Tool_A_Reference	360	0.338889	0.868283	0	0.0	6
AOI_Tool_B_Calibrated	360	0.330556	0.952440	0	0.0	7
AOI_Tool_C_LowSensitivity	360	0.388889	0.913464	0	0.0	6
AOI_Tool_D_HighSensitivity	360	0.452778	1.088298	0	0.0	8

The standard deviations were substantially larger than the corresponding means for all AOI systems, suggesting the presence of overdispersion in the defect-count data. Such behavior cannot be adequately represented by conventional Gaussian-type Gauge R&R models and therefore supports the use of the proposed zero-inflated correlated count-data framework. Among the four AOI systems, AOI_Tool_D_HighSensitivity exhibited both the highest average defect count and the largest maximum observed defect count, indicating that this inspection system possessed the strongest defect-detection sensitivity. In contrast, AOI_Tool_A_Reference exhibited relatively lower defect-count measurements and narrower variation.

The correlation structure between particle defects and pattern/scratch defects is presented in Table 3. The estimated correlations ranged from approximately 0.637 to 0.723, indicating moderate-to-strong positive dependence between the two defect categories. These results strongly support the fundamental assumption of the proposed bivariate Poisson framework that different defect categories may originate from common latent defect-generation mechanisms.

Table 3. Correlation Between Particle Defects and Pattern/Scratch Defects.

System_ID	AOI_System	Correlation_Particle_vs_PatternScratch
0	AOI_Tool_A_Reference	0.637395
1	AOI_Tool_B_Calibrated	0.722552
2	AOI_Tool_C_LowSensitivity	0.657682
3	AOI_Tool_D_HighSensitivity	0.642012

From a semiconductor manufacturing perspective, correlated defect behavior is physically reasonable because particle contamination, pattern abnormalities, and structural defects may simultaneously arise from common process disturbances such as contamination events, lithography instability, wafer handling variation, or localized process nonuniformity. Consequently, the proposed correlated defect-count framework provides a more realistic representation of AOI inspection behavior than independent Poisson models.

The structural-zero characteristics of the AOI inspection data are summarized in Table 4. The estimated structural-zero rates ranged from approximately 51% to 56% across the four AOI systems, indicating that more than half of the inspected wafer regions belonged to the defect-free state. Such behavior is commonly observed in advanced semiconductor manufacturing environments characterized by highly stable production processes and high wafer yield.

 doi:10.20944/preprints202605.0948.v1

Table 4. Structural-Zero Characteristics Across AOI Inspection Systems.

AOI_System	Structural_Zero_Count	Structural_Zero_Rate	Structural_Zero_Rate_Percent
AOI_Tool_A_Reference	189	0.525000	52.500000
AOI_Tool_B_Calibrated	201	0.558333	55.833333
AOI_Tool_C_LowSensitivity	185	0.513889	51.388889
AOI_Tool_D_HighSensitivity	196	0.544444	54.444444

The large proportion of structural-zero observations demonstrates that traditional Gaussian-based Gauge R&R methods may be statistically inappropriate for semiconductor AOI inspection applications. The proposed zero-inflated framework explicitly captures the excessive zero-defect behavior while simultaneously preserving the correlated structure among defect categories.

Figure 2 illustrates the distribution of total defect counts across the AOI inspection systems. The distributions exhibit strong right-skewed behavior with a dominant concentration at zero defects. This observation is fully consistent with the previously reported structural-zero rates. Although most wafer regions contained no detectable defects, a small number of regions exhibited multiple simultaneous defects, thereby generating long right tails in the defect-count distributions.

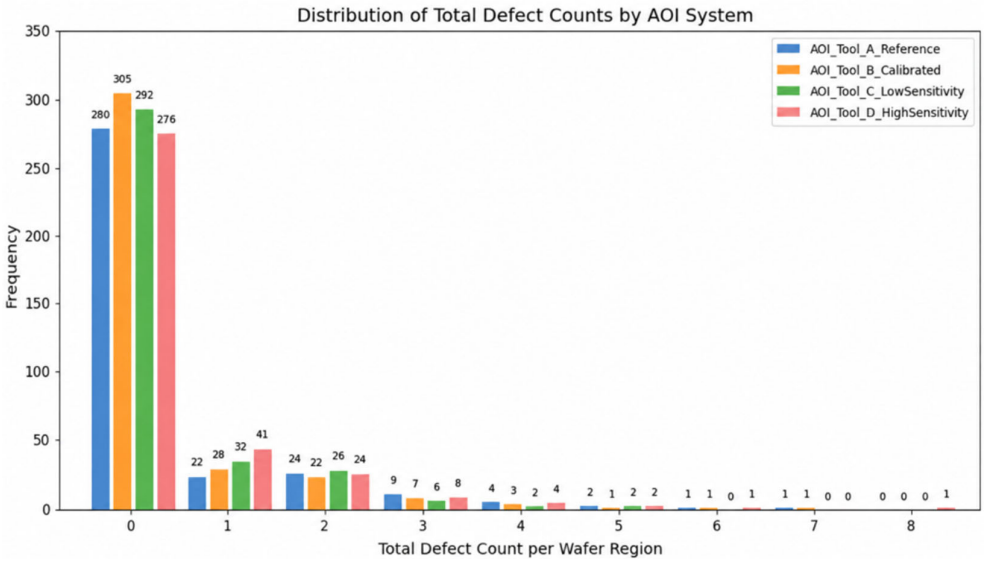

Figure 2. Distribution of Total Defect Counts Across AOI Inspection Systems.

The figure also reveals systematic differences among AOI systems in terms of inspection sensitivity. AOI_Tool_D_HighSensitivity produced a wider distribution with a larger frequency of high-defect observations, whereas AOI_Tool_A_Reference exhibited a more concentrated low-defect distribution. These results indicate that AOI system sensitivity substantially influences observed defect-count variation.

The mean defect counts for different defect categories are compared in Figure 3. The results indicate that particle defects and pattern/scratch defects do not exhibit identical detection behavior across AOI systems. AOI_Tool_D_HighSensitivity consistently produced the highest mean defect counts for both defect categories, whereas AOI_Tool_A_Reference exhibited relatively lower average defect levels

 doi:10.20944/preprints202605.0948.v1

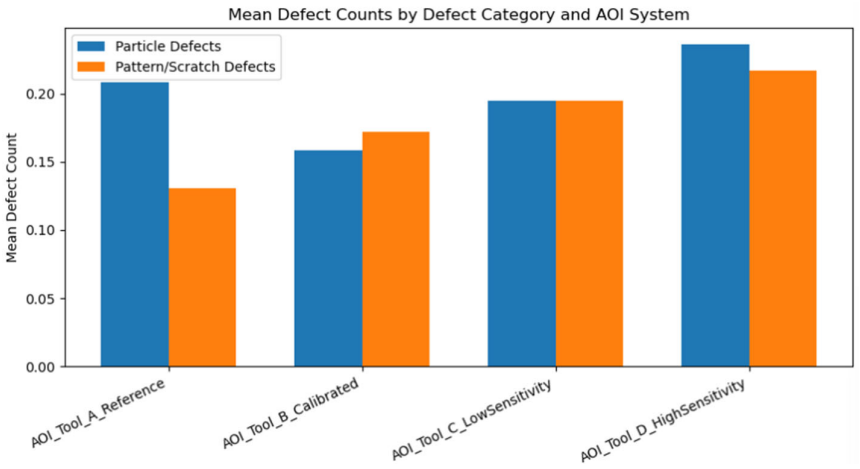

Figure 3. Mean Defect Counts by Defect Category and AOI Inspection System.

The figure further demonstrates that different AOI systems possess varying sensitivities toward different defect-generation mechanisms. This observation provides additional justification for the proposed multivariate correlated defect-count modeling structure instead of a single-response measurement framework.

Figure 4 compares the structural-zero rates generated by the latent zero-inflation mechanism and the observed zero-defect rates in the AOI inspection dataset. The observed zero-defect rates were substantially higher than the structural-zero rates because a portion of observations generated from the bivariate Poisson state also resulted in zero observed defects.

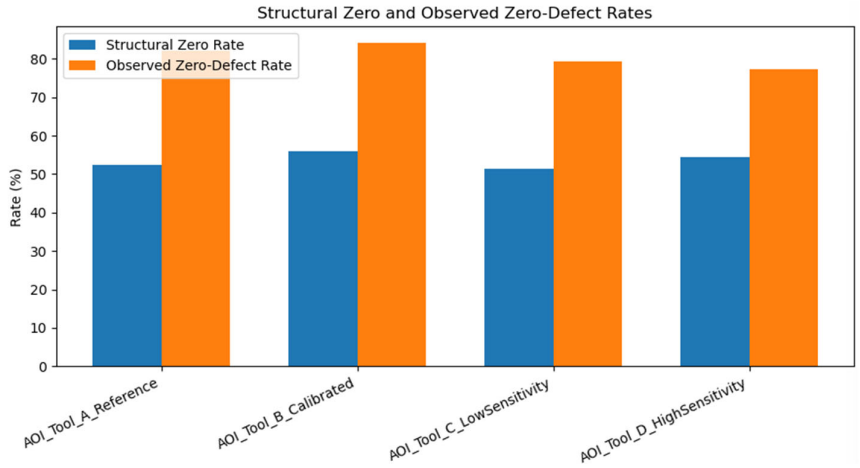

Figure 4. Structural-Zero Rates and Observed Zero-Defect Rates Across AOI Systems.

The consistency between structural-zero behavior and observed zero-defect frequencies demonstrates that the proposed framework successfully reproduces the excessive zero-defect behavior commonly encountered in high-yield semiconductor wafer inspection environments.

The spatial wafer-zone effect is illustrated in Figure 5. The results clearly indicate that edge regions exhibited substantially higher mean defect counts than center and middle wafer regions across all AOI systems. This phenomenon is consistent with practical semiconductor manufacturing behavior because wafer-edge regions are typically more susceptible to process nonuniformity, edge-bead effects, etching instability, and localized contamination.

 doi:10.20944/preprints202605.0948.v1

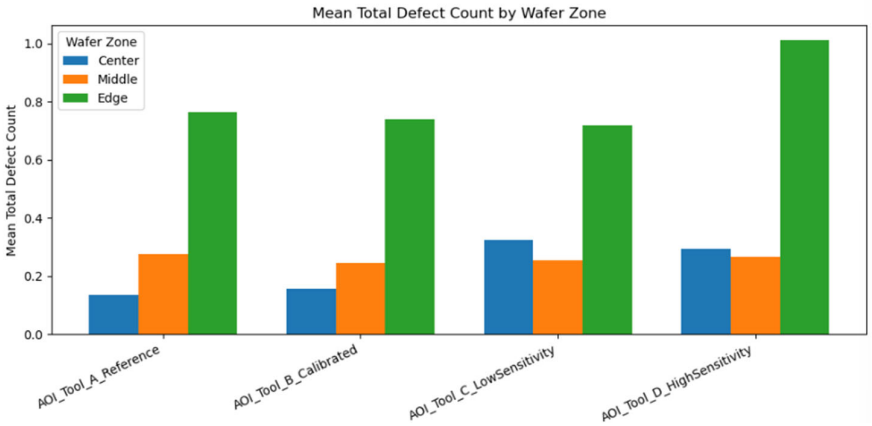

Figure 5. Mean Total Defect Counts Across Wafer Zones.

The figure demonstrates that the proposed framework successfully incorporates spatial process variation into the defect-generation mechanism. Such spatial characteristics are highly relevant for wafer-level process monitoring and semiconductor yield management.

The residual distributions of the fitted zero-inflated correlated defect-count model are presented in Figure 6. Most residual observations were concentrated near zero, indicating that the proposed framework adequately captured the primary defect-generation behavior of the AOI inspection process. Nevertheless, several positive residual outliers remained, corresponding to wafer regions with unusually high defect counts.

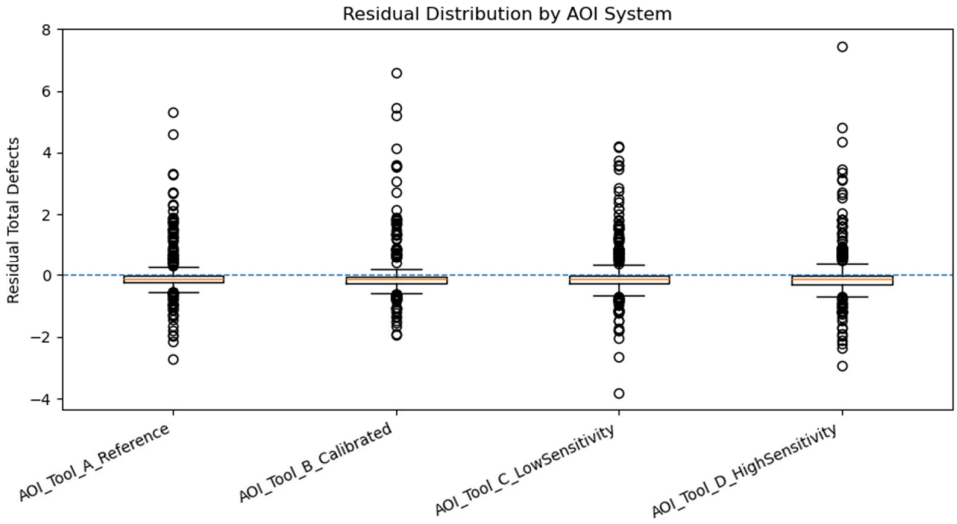

Figure 6. Distribution Across AOI Inspection Systems.

Among the four systems, AOI_Tool_D_HighSensitivity exhibited slightly wider residual dispersion, which may be attributed to its stronger defect-detection capability and increased sensitivity toward local defect fluctuations. Overall, the residual analysis suggests that the proposed framework provides a statistically reasonable representation of semiconductor wafer AOI inspection behavior.

The variance decomposition and Percent R&R results are summarized in Table 5. The results indicate that repeatability variance contributed substantially more to total measurement-system variation than between-system variance. The estimated between-system variance remained relatively small across all AOI systems, indicating that systematic differences among AOI tools were less influential than within-system inspection fluctuation.

 doi:10.20944/preprints202605.0948.v1

Table 5. Variance Decomposition and Percent R&R Results.

AOI_System	Repeatability_Variance	Between_System_Variance	Reproducibility_Variance	Total_Std_Dev	Percent_RR	Bootstrap_Median_Percent_RR	CI_Lower_95	CI_Upper_95
AOI_Tool_A_Reference	0.610712	0.001122	0.611835	0.959387	81.53109	79.559986	63.67681	93.52767
AOI_Tool_B_Calibrated	0.782544	0.001122	0.783666	0.959387	92.27241	91.645465	76.01221	109.34358
AOI_Tool_C_LowSensitivity	0.717524	0.001122	0.718646	0.959387	88.36165	86.157747	73.81676	100.46189
AOI_Tool_D_HighSensitivity	0.828845	0.001122	0.829967	0.959387	94.95913	94.780549	77.97023	111.23211

The estimated Percent R&R values ranged from approximately 81.5% to 95.0%. AOI_Tool_D_HighSensitivity exhibited the largest Percent R&R value, whereas AOI_Tool_A_Reference exhibited the lowest value. These relatively large Percent R&R values indicate that measurement-system-related variation accounted for a substantial portion of the total observed defect-count variation.

From an engineering perspective, these results suggest that high-sensitivity semiconductor AOI systems may simultaneously improve defect detectability while also increasing measurement variability. Consequently, additional calibration procedures, inspection recipe refinement, sensitivity normalization, or algorithmic optimization may be required to improve measurement-system stability.

Figure 7 presents the estimated Percent R&R values together with their corresponding 95% bootstrap confidence intervals. The bootstrap intervals demonstrate that the proposed inference framework provides stable uncertainty quantification for the R&R estimates.

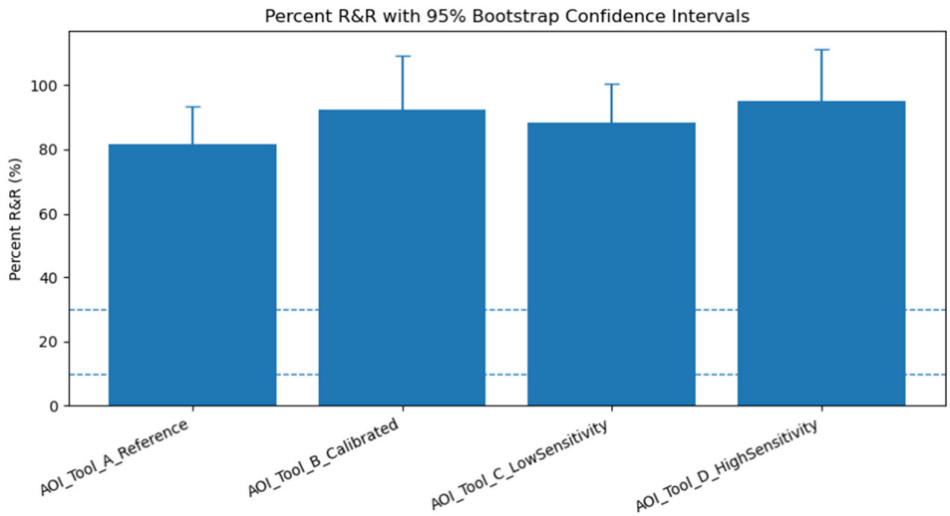

Figure 7. Percent R&R with 95% Bootstrap Confidence Intervals.

AOI_Tool_D_HighSensitivity exhibited both the highest Percent R&R value and the widest confidence interval, indicating that highly sensitive inspection systems may also introduce greater measurement uncertainty. In contrast, AOI_Tool_A_Reference exhibited comparatively narrower confidence intervals and lower Percent R&R values, suggesting more stable measurement behavior.

 doi:10.20944/preprints202605.0948.v1

Finally, Figure 8 illustrates the bootstrap distributions of Percent R&R across the four AOI systems. The distributions were approximately unimodal and reasonably symmetric, indicating that the bootstrap estimation procedure produced numerically stable and statistically consistent R&R estimates.

The centers of the bootstrap distributions closely aligned with the estimated Percent R&R values reported previously, providing additional evidence regarding the robustness of the proposed zero-inflated correlated defect-count R&R framework. Overall, the numerical analysis demonstrates that the proposed methodology is capable of effectively modeling semiconductor wafer AOI inspection data characterized by correlated defect counts, excessive zero-defect observations, and measurement-system variability.

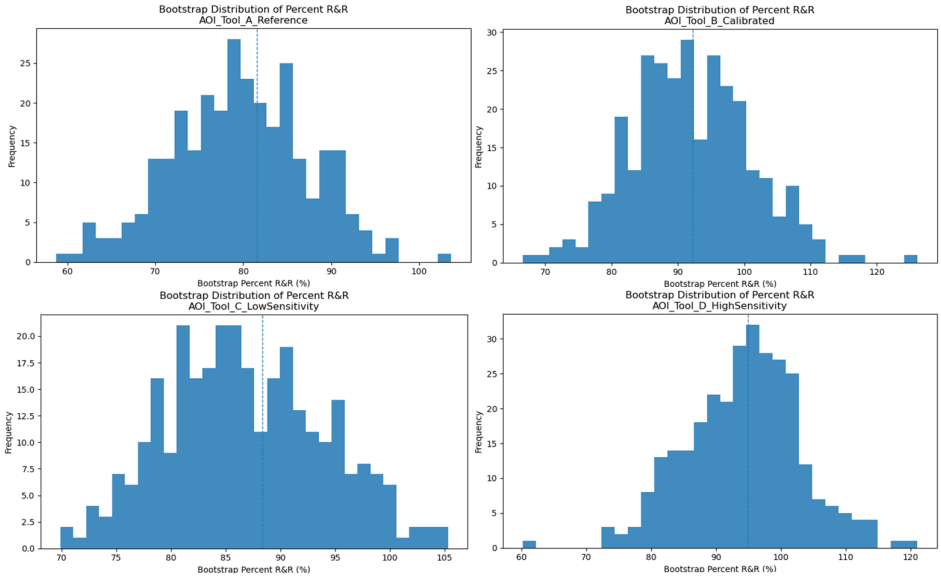

Figure 8. Bootstrap Distributions of Percent R&R Across AOI Systems.

Table 6 summarizes the spatial wafer-zone defect characteristics across the four AOI inspection systems. The results clearly indicate that wafer-edge regions consistently exhibited substantially higher mean defect counts than center and middle wafer regions. For example, AOI_Tool_D_HighSensitivity produced an average defect count exceeding 1.0 defects per wafer region in edge areas, whereas the corresponding center-region mean remained below 0.30 defects.

This phenomenon is highly consistent with practical semiconductor manufacturing behavior. Wafer-edge regions are typically more susceptible to process nonuniformity, edge-bead effects, lithography distortion, plasma etching instability, and localized contamination accumulation. Consequently, edge regions often exhibit higher defect-generation probability and stronger inspection variability.

The results further demonstrate that the proposed framework successfully incorporates spatial process variation into the defect-generation mechanism. Moreover, the variability observed in edge regions was substantially larger than that observed in center and middle regions, as indicated by the larger standard deviations. This observation suggests that wafer-edge inspection may contribute disproportionately to measurement-system variability in semiconductor AOI environments.

Among the four AOI systems, AOI_Tool_D_HighSensitivity exhibited the strongest edge-region defect response, indicating that highly sensitive AOI systems may amplify both defect detectability and measurement variation simultaneously. Such behavior is practically important because semiconductor manufacturers often face a tradeoff between inspection sensitivity and measurement-system stability.

 doi:10.20944/preprints202605.0948.v1

Table 6. Spatial Wafer-Zone Defect Characteristics Across AOI Inspection Systems.

AOI_System	Wafer_Zone	Sample_Size	Mean_Total_Defects	Std_Total_Defects
AOI_Tool_A_Reference	Center	133	0.1353	0.4401
AOI_Tool_A_Reference	Edge	85	0.7647	1.3509
AOI_Tool_A_Reference	Middle	142	0.2746	0.7060
AOI_Tool_B_Calibrated	Center	133	0.1579	0.7054
AOI_Tool_B_Calibrated	Edge	85	0.7412	1.3107
AOI_Tool_B_Calibrated	Middle	142	0.2465	0.8267
AOI_Tool_C_LowSensitivity	Center	133	0.3233	0.8838
AOI_Tool_C_LowSensitivity	Edge	85	0.7176	1.2307
AOI_Tool_C_LowSensitivity	Middle	142	0.2535	0.6351
AOI_Tool_D_HighSensitivity	Center	133	0.2932	0.9027
AOI_Tool_D_HighSensitivity	Edge	85	1.011765	1.451552
AOI_Tool_D_HighSensitivity	Middle	142	0.267606	0.866357

Table 7 presents the variance decomposition and Percent R&R performance evaluation results for the four AOI inspection systems. The results indicate that repeatability variance constituted the dominant source of measurement-system variation across all AOI systems, whereas the estimated between-system variance remained relatively small.

The relatively small between-system variance suggests that systematic differences among AOI tools were limited under the simulated calibration environment. In contrast, within-system inspection fluctuation contributed substantially more strongly to overall measurement uncertainty. Such repeatability variation may originate from illumination instability, autofocus fluctuation, sensor noise, local wafer reflectivity variation, and stochastic inspection uncertainty.

The estimated Percent R&R values ranged from approximately 81.5% to 95.0%, indicating that measurement-system-related variation accounted for a substantial proportion of the total observed defect-count variation. Among the four AOI systems, AOI_Tool_D_HighSensitivity exhibited the highest Percent R&R value, while AOI_Tool_A_Reference exhibited the lowest value.

From an engineering perspective, these results imply that highly sensitive AOI inspection systems may improve defect detectability while simultaneously increasing measurement variability. Consequently, semiconductor manufacturers may need to perform additional inspection calibration, threshold optimization, recipe refinement, or sensitivity normalization in order to reduce measurement-system uncertainty.

The bootstrap median Percent R&R values were highly consistent with the primary Percent R&R estimates, indicating that the proposed bootstrap inference procedure produced numerically stable estimation results. Furthermore, the relatively wide confidence intervals observed for AOI_Tool_D_HighSensitivity suggest that high-sensitivity inspection systems may also introduce greater uncertainty into the measurement-system evaluation process.

Overall, the results demonstrate that the proposed zero-inflated correlated defect-count R&R framework is capable of effectively quantifying measurement-system variation in semiconductor wafer AOI inspection environments characterized by correlated defects, excessive zero-defect observations, and spatial process heterogeneity.

 doi:10.20944/preprints202605.0948.v1

Table 7. Measurement System Variance Decomposition and R&R Performance Evaluation.

AOI_System	Repeatability_Variance	Between_System_Variance	Reproducibility_Variance	Total_Std_Dev	Percent_RR	Bootstrap_Median_Percent_RR	CI_Lower_95	CI_Upper_95	RR_Interpretation
AOI_Tool_A_Reference	0.6107	0.0011	0.6118	0.9594	81.5311	79.5600	63.6768	93.5277	Needs improvement
AOI_Tool_B_Calibrated	0.7825	0.0011	0.7837	0.9594	92.2724	91.6455	76.0122	109.3436	Needs improvement
AOI_Tool_C_LowSensitivity	0.7175	0.0011	0.7186	0.9594	88.3617	86.1577	73.8168	100.4619	Needs improvement
AOI_Tool_D_HighSensitivity	0.8288	0.0011	0.8300	0.9594	94.9591	94.7805	77.9702	111.2321	Needs improvement

6. Conclusions

This study proposed a novel measurement system repeatability and reproducibility (R&R) framework for zero-inflated correlated defect-count data in semiconductor wafer automated optical inspection (AOI) environments. The proposed methodology was developed to address several important limitations of conventional Gauge R&R approaches when applied to modern semiconductor manufacturing systems characterized by discrete defect counts, excessive zero-defect observations, correlated defect-generation mechanisms, and spatial wafer-level process heterogeneity.

Unlike traditional Gaussian-based measurement system models, the proposed framework incorporated a zero-inflated bivariate Poisson structure capable of simultaneously modeling structural zero-defect states and correlated defect categories. The proposed model further introduced a latent shared defect component to represent common process disturbances affecting multiple defect mechanisms simultaneously. Such modeling characteristics are particularly important in semiconductor wafer inspection environments because particle contamination, pattern abnormalities, and structural defects frequently originate from shared process-related factors such as contamination events, lithography instability, wafer handling variation, and localized process nonuniformity.

To estimate the proposed model parameters, this study developed an expectation-maximization (EM) algorithm capable of estimating latent structural-zero states and correlated latent defect-generation components. The proposed estimation procedure further enabled the decomposition of total measurement variation into repeatability variance and reproducibility variance under a zero-inflated correlated defect-count framework. In addition, bootstrap inference was incorporated to construct confidence intervals for the proposed Percent R&R measures and to evaluate the statistical stability of the estimation procedure.

From a theoretical perspective, this study established several important analytical properties of the proposed framework. The covariance structure of the bivariate Poisson model was formally derived, and the identifiability of the latent shared defect component was analytically demonstrated. Furthermore, the monotonic likelihood property of the EM algorithm, the consistency of the repeatability variance estimator, the nonnegativity of the reproducibility variance estimator, and the asymptotic consistency of the proposed Percent R&R estimator were established through mathematical propositions and proofs. These theoretical results provide a rigorous statistical foundation for the proposed methodology.

The numerical analysis demonstrated that the proposed framework successfully reproduced several important statistical characteristics commonly observed in semiconductor wafer AOI inspection environments. The simulated defect-count data exhibited strong zero inflation,

 doi:10.20944/preprints202605.0948.v1

overdispersion, correlated defect behavior, and spatial wafer-zone variation. The results further showed that wafer-edge regions consistently exhibited higher defect counts and larger measurement variability than center and middle wafer regions, which is highly consistent with practical semiconductor manufacturing behavior.

The measurement system analysis results indicated that repeatability variation constituted the dominant source of total measurement-system variation, whereas between-system variation remained relatively small. Moreover, highly sensitive AOI inspection systems exhibited both stronger defect-detection capability and larger measurement variability. These findings suggest that semiconductor manufacturers may face an important tradeoff between inspection sensitivity and measurement-system stability. Consequently, additional inspection calibration, recipe optimization, threshold adjustment, or sensitivity normalization may be required to improve measurement-system reliability in high-sensitivity AOI environments.

Overall, the proposed framework extends conventional Gauge R&R analysis from continuous Gaussian measurements to zero-inflated correlated defect-count inspection data. The methodology developed in this study therefore provides a more realistic and statistically appropriate measurement-system evaluation framework for advanced semiconductor manufacturing environments.

Although the proposed framework demonstrated promising analytical and numerical performance, several limitations remain. First, the numerical analysis was based primarily on simulated semiconductor wafer inspection data rather than large-scale industrial AOI datasets. Second, the current framework considered only two correlated defect categories. In practical semiconductor manufacturing environments, multiple defect categories may coexist simultaneously and exhibit more complex dependence structures. Third, the current study assumed static inspection-system behavior across the observation horizon and did not explicitly consider temporal process drift, inspection-tool degradation, or dynamic calibration effects.

Future research may therefore extend the proposed framework in several directions. One important extension involves applying the proposed methodology to real semiconductor wafer AOI datasets obtained from advanced manufacturing facilities. Another promising direction involves extending the model toward multivariate zero-inflated defect-count structures capable of simultaneously modeling multiple correlated defect categories. Future studies may also integrate temporal process dynamics, Bayesian hierarchical modeling, deep-learning-based defect classification systems, or digital-twin-based semiconductor manufacturing environments into the proposed measurement-system framework. In addition, integrating the proposed R&R methodology with inline process control and predictive yield-management systems may further enhance its practical value in intelligent semiconductor manufacturing applications.

In conclusion, this study provides a statistically rigorous and practically relevant framework for evaluating measurement-system variation in semiconductor wafer AOI inspection environments characterized by correlated defect counts and excessive zero-defect observations. The proposed methodology contributes both theoretically and practically to the fields of measurement system analysis, semiconductor manufacturing quality control, and intelligent precision inspection systems.

Author Contributions: Conceptualization, M.N.C. and C.-C.F.; Data Curation, M.N.C. and C.-C.F.; Formal Analysis, M.N.C. and C.-C.F.; Funding Acquisition, M.N.C.; Investigation, C.-C.F.; Methodology, M.N.C. and C.-C.F.; Project Administration, M.N.C. and C.-C.F.; Resources, M.N.C. and C.-C.F.; Supervision, M.N.C.; Writing—Review and Editing, M.N.C. and C.-C.F. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Guangdong Basic and Applied Basic Research Foundation and the Guangdong Soft Science Foundation, China [grant number 2024A0505050043].

Data Availability Statement: The original contributions presented in this study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author(s).

 doi:10.20944/preprints202605.0948.v1

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Abd Al Rahman, M.; Mousavi, A. A review and analysis of automatic optical inspection and quality monitoring methods in electronics industry. *IEEE Access* **2020**, *8*, 183192–183271.
2. Wang, C.C.; Yang, Y.Y. A machine learning approach for improving wafer acceptance testing based on an analysis of station and equipment combinations. *Mathematics* **2023**, *11*, 1569.
3. Wu, P.H.; Hoang, T.P.; Chou, Y.T.; Mayol, A.P.; Lai, Y.W.; Kang, C.H.; Chen, S.H. Elevating wafer defect inspection with denoising diffusion probabilistic model. *Mathematics* **2024**, *12*, 3164.
4. Fu, H.; Lai, Y.; Pan, C.; Zhang, S.; Bai, L.; Li, J. A central array method to locate chips in AOI systems in semiconductor manufacturing. *Electronics* **2024**, *13*, 1070.
5. Fu, C.J.; Chen, H.L.; Tseng, H.Y. Application of artificial intelligence to improve chip defect detection using semiconductor equipment. *Engineering Proceedings* **2025**, *98*, 26.
6. Tabatabaeemoshiri, P.; Kumar, N.; Khairuddin, A.S.M.; Ting, D.; Regeev, V. A novel approach to test-induced defect detection in semiconductor wafers using graph-based semi-supervised learning. *IEEE Access* **2025**, *13*, 21678–21694.
7. De Mast, J.; Trip, A. Gauge R&R studies for destructive measurements. *J. Qual. Technol.* **2005**, *37*, 40–49.
8. Hoffa, D.W.; Laux, C. Gauge R&R: An effective methodology for determining the adequacy of a new measurement system for micron-level metrology. *J. Ind. Technol.* **2007**, *23*, 1–8.
9. Kazerouni, A.M. Design and analysis of gauge R&R studies: Making decisions based on ANOVA method. *World Acad. Sci. Eng. Technol.* **2009**, *52*, 31–35.
10. Weaver, B.P.; Hamada, M.S.; Vardeman, S.B.; Wilson, A.G. A Bayesian approach to the analysis of gauge R&R data. *Qual. Eng.* **2012**, *24*, 486–500.
11. Sharma, M.; Sahni, S.P.; Sharma, S. Validating a destructive measurement system using Gauge R&R—A case study. *Eng. Manag. Prod. Serv.* **2019**, *11*, 34–42.
12. Mircea, B.R.; Mirescu, R.; Tucu, D. The use of gauge R&R method in manufacturing of agricultural machinery. **2023**.
13. Petitjean, A.; Musset, O. Application for determining repeatability and reproducibility in an honest Gauge R&R study. *Qual. Reliab. Eng. Int.* **2025**, *41*, 1953–1970.
14. Hamada, M.S. Operator linear analysis for Gauge R&R studies. *Qual. Eng.* **2025**, *37*, 419–429.
15. Malindzakova, M.; Ambrisko, L.; Simonova, T. Quality control of measurements using the repeatability and reproducibility (R&R) method. *Manag. Prod. Eng. Rev.* **2025**.
16. Hamada, M.S.; Ryan, K.J. On nominal data Gauge R&R studies. *Qual. Eng.* **2025**, *37*, 560–570.
17. Petitjean, A. Use A-, E- and D-optimal designs to reduce the number of measurements required for a Gauge R&R study. *Measurement* **2026**, 120834.
18. Wagh, Y.S.; Kamalja, K.K. Zero-inflated models and estimation in zero-inflated Poisson distribution. *Commun. Stat. Simul. Comput.* **2018**, *47*, 2248–2265.
19. Zhou, Y.; Wan, X.; Zhang, B.; Tong, T. Classifying next-generation sequencing data using a zero-inflated Poisson model. *Bioinformatics* **2018**, *34*, 1329–1335.
20. Kamalja, K.K.; Wagh, Y.S. Estimation in zero-inflated generalized Poisson distribution. *J. Data Sci.* **2018**, *16*, 183–206.
21. Truong, B.C.; Pho, K.H.; Dinh, C.C.; McAleer, M. Zero-inflated Poisson regression models: Applications in the sciences and social sciences. *Ann. Financ. Econ.* **2021**, *16*, 2150006.
22. Choi, J.; Chapkin, R.; Ni, Y. Bayesian causal structural learning with zero-inflated Poisson Bayesian networks. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 5887–5897.
23. Zhang, Q.; Yi, G.Y. Zero-inflated Poisson models with measurement error in the response. *Biometrics* **2023**, *79*, 1089–1102.
24. Fang, C.C.; Chen, M.N. Integrated quality inspection and production run optimization for imperfect production systems with zero-inflated non-homogeneous Poisson deterioration. *Mathematics* **2025**, *13*, 3901.
25. Lu, C.; Rastelli, R.; Friel, N. A zero-inflated Poisson latent position cluster model. *Netw. Sci.* **2026**, *14*, e2.

 doi:10.20944/preprints202605.0948.v1

26. Chien, C.F.; Hung, W.T.; Pan, C.W.; Van Nguyen, T.H. Decision-based virtual metrology for advanced process control to empower smart production and an empirical study for semiconductor manufacturing. *Comput. Ind. Eng.* **2022**, *169*, 108245.
27. Senoner, J.; Netland, T.; Feuerriegel, S. Using explainable artificial intelligence to improve process quality: Evidence from semiconductor manufacturing. *Manag. Sci.* **2022**, *68*, 5704–5723.
28. Maitra, V.; Su, Y.; Shi, J. Virtual metrology in semiconductor manufacturing: Current status and future prospects. *Expert Syst. Appl.* **2024**, *249*, 123559.
29. Kim, D.; Kim, H.; Kim, Y.; Chae, M.; Ko, Y.M.; Bae, Y.M.; Noh, K.H. Using statistical models for optimal packaging in semiconductor manufacturing processes. *J. Korean Stat. Soc.* **2025**, *54*, 1–19.
30. Suwattananuruk, B.; Chien, C.F. Denoising variational autoencoders for smart inspection of wafer probe card PCB channels for advancing quality control for semiconductor manufacturing. *Comput. Ind. Eng.* **2025**, 111453.
31. Fiorino, F.M.; Ruffino, F.; Catena, A. Investigation of defect propagation in 4H-SiC: From substrate to epitaxial layers. *Appl. Sci.* **2026**, *16*, 2727.
32. Jiang, Y.; Yang, F.; Li, C. Q-learning enhanced multi-objective particle swarm optimization for photolithography-etching scheduling in semiconductor manufacturing. *Expert Syst. Appl.* **2026**, *301*, 130506.
33. Nurmi, T. Determination of resolution, repeatability, and detection capability improvements of automatic optical inspection in micro-electro-mechanical systems manufacturing. **2022**.
34. Li, J.; Hu, E.; Wei, W.; Shi, F. From wafers to electrodes: Transferring automatic optical inspection (AOI) for multiscale characterization of smart battery manufacturing. *Adv. Funct. Mater.* **2026**, e28142.