electronics

MDPI

Review

Review of Applications of Regression and Predictive Modeling in Wafer Manufacturing

Hsuan-Yu Chen [1] and Chiachung Chen [2,*]

[1] Africa Industrial Research Center, National Chung Hsing University, Taichung 40227, Taiwan; wakaharu37@gmail.com
[2] Department of Bio-Industrial Mechatronics Engineering, National Chung Hsing University, Taichung 40227, Taiwan
* Correspondence: ccchen@dragon.nchu.edu.tw; Tel.: +886-4-22857562

Abstract

Semiconductor wafer manufacturing is one of the most complex and data-intensive industrial processes, comprising 500–1000 tightly interdependent steps, each requiring nanometer-level precision. As device nodes approach 3 nm and beyond, even minor deviations in parameters such as oxide thickness or critical dimensions can lead to catastrophic yield loss, challenging traditional physics-based control methods. In response, the industry has increasingly adopted regression analysis and predictive modeling as essential analytical frameworks. Classical regression, long used to support design of experiments (DOE), process optimization, and yield analysis, has evolved to enable multivariate modeling, virtual metrology, and fault detection. Predictive modeling extends these capabilities through machine learning and AI, leveraging massive sensor and metrology data streams for real-time process monitoring, yield forecasting, and predictive maintenance. These data-driven tools are now tightly integrated into advanced process control (APC), digital twins, and automated decision-making systems, transforming fabs into agile, intelligent manufacturing environments. This review synthesizes foundational and emerging methods, industry applications, and case studies, emphasizing their role in advancing Industry 4.0 initiatives. Future directions include hybrid physics–ML models, explainable AI, and autonomous manufacturing. Together, regression and predictive modeling provide semiconductor fabs with a robust ecosystem for optimizing performance, minimizing costs, and accelerating innovation in an increasingly competitive, high-stakes industry.

Keywords: regression; predictive modeling; AI; machine learning; wafer manufacturing

Academic Editors: Jie Zhou, Shuoyang Qiu and Jiarui Gong

Received: 22 September 2025
Revised: 16 October 2025
Accepted: 16 October 2025
Published: 17 October 2025

Citation: Chen, H.-Y.; Chen, C. Review of Applications of Regression and Predictive Modeling in Wafer Manufacturing. *Electronics* **2025**, *14*, 4083. https://doi.org/10.3390/electronics14204083

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Semiconductor wafer manufacturing is widely recognized as one of the most intricate, multivariate, and data-intensive industrial processes in existence [1]. The fabrication of a single wafer involves 500 to 1000 sequential process steps, each one interdependent and highly sensitive to variation [2,3]. Semiconductor fabrication involves photolithography, thin-film deposition, ion implantation, etching, cleaning, CMP, metrology, inspection, and electrical testing [4]. Engineers analyze vast data—such as film thickness, line widths, critical dimensions (CD), overlay accuracy, defect counts, and electrical parameters—to ensure wafer quality [2]. Each layer requires sub-nanometer precision; even minor deviations can compound across hundreds of steps, resulting in significant yield loss [2,4].

The overall classification of the wafer manufacturing processes is shown in Figure 1.

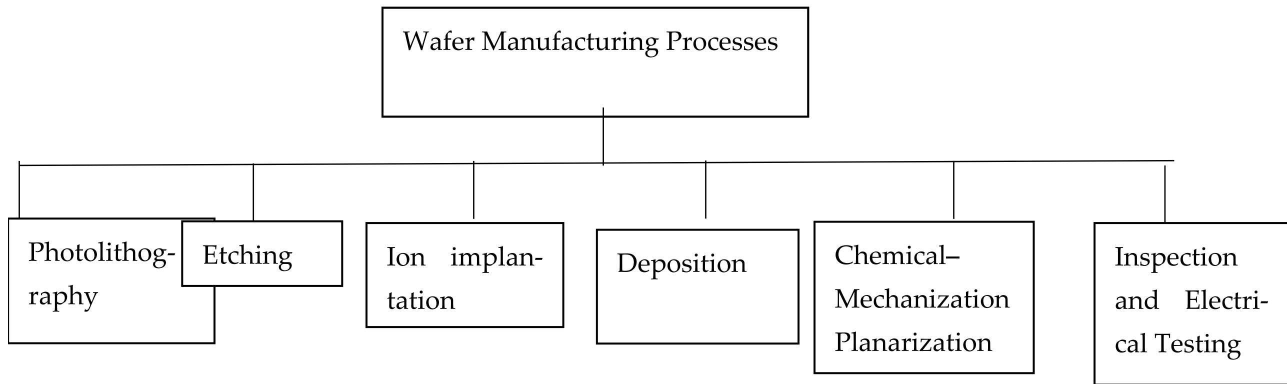

Figure 1. The overall classification of the wafer manufacturing processes.

As technology nodes shrink toward 3 nm and beyond, manufacturing tolerances have become incredibly stringent [5]. Variations as minor as a few angstroms in oxide thickness or a few nanometers in CD can have profound impacts on device performance and overall yield [6]. These tight margins are further challenged by factors such as tool-to-tool variability, recipe drift, environmental fluctuations, consumable wear, and raw material lot differences [5]. In earlier technology nodes, engineers often relied on physics-based heuristics and empirical control strategies. Still, these traditional approaches struggle to provide sufficient precision and adaptability in the face of modern manufacturing complexity [7,8]. The semiconductor industry has therefore embraced advanced statistical modeling and machine learning techniques to manage this growing variability and maintain competitiveness [7,9].

Wafer fabrication aims to maximize yield, minimize variability, detect process excursions early, and sustain high tool uptime and throughput [10]. Achieving these goals requires proactive, data-driven control rather than reactive troubleshooting [5,6]. Regression analysis and predictive modeling now underpin semiconductor manufacturing, empowering engineers to monitor, forecast, and optimize processes with exceptional precision and efficiency [11].

Regression, a long-standing staple of semiconductor data analytics, is used to quantify the relationships between process inputs and outputs [11–13]. Inputs may include recipe parameters, tool settings, consumable lifetimes, or incoming material characteristics, while outputs encompass key metrics like wafer yield, defect density, and device electrical performance [12]. Regression models, including linear, polynomial, and multivariate regression [11,13], have historically underpinned the design of experiments, helping process engineers identify critical factors, model process interactions, and fine-tune recipes [4,14]. However, regression has evolved well beyond its traditional role, now serving as a foundation for high-dimensional analysis, fault detection models, and virtual metrology systems [4,7,13,14]. Wan et al. [14] evaluate the performance of four typical regression methods for VM: multiple linear regression (MLR), least absolute shrinkage and selection operator (LASSO), neural networks (NN), and Gaussian process regression (GPR) on the virtual metrology.

Predictive modeling represents a natural extension of regression-based analytics [13], incorporating machine learning (ML) and artificial intelligence (AI) to handle large-scale, nonlinear, and high-dimensional data streams [15]. For example, virtual metrology models predict measurements, such as film thickness or CD, using upstream sensor data, allowing fabs to reduce reliance on time-consuming or destructive physical metrology steps [1,16,17]. Dreyfus et al. [17] reported a detailed review of the development of virtual metrology (VM). Similarly, fault detection and classification (FDC) systems integrate predictive algorithms

to detect subtle deviations in process signals before they escalate into yield-impacting issues [6,8]. By embedding predictive models into run-to-run (R2R) control loops, fabs can achieve real-time adjustments and minimize process drift [6,14].

Another critical application of predictive modeling is yield forecasting, which enables proactive decision-making by predicting wafer-level or lot-level yield distributions [2,18]. These forecasts inform engineering teams about potential risks and guide preventive measures, such as equipment calibration or recipe modification, before failures occur [19]. Machine learning models trained on sensor and maintenance logs can anticipate tool failures, optimize maintenance schedules, and minimize costly downtime [15]. This proactive approach represents a paradigm shift from reactive troubleshooting to autonomous optimization, aligning with the semiconductor industry's vision of "lights-out" manufacturing. In this future, fabs run continuously with minimal human intervention [17,19].

Device complexity, market pressures, and rapid innovation cycles require fabs to maintain agility while ensuring uncompromising quality [8]. Regression and predictive modeling not only enhance process visibility but also provide a framework for continuous improvement, where insights from historical and real-time data feed into iterative optimization cycles [2,8]. By integrating these models with advanced process control, fabs achieve tighter control limits, faster ramp-to-yield for new technology nodes, and significant cost savings [16].

This article reviews regression and predictive modeling in semiconductor wafer manufacturing, covering classical regression, multivariate statistics, machine learning, AI, and predictive maintenance frameworks. It examines how these methods enhance understanding of process interactions, failure prediction, and efficiency optimization. Emphasizing integration with SPC, DOE, virtual metrology, digital twins, and analytics platforms, the review positions these tools as pillars of smart fabs—where data-driven intelligence enables highly efficient, predictive, and scalable wafer production.

In this review of applications of regression and predictive modeling in wafer manufacturing, the background is introduced in the first section. The application of regression and predictive modeling in wafer manufacturing is described in Sections 2 and 3. The fourth and fifth sections describe the regression and predictive modeling methods. The sixth section is the integration with machine learning. The seventh section describes the challenges of regression and predictive modeling. The future trends are discussed in Section 8. Finally, there is Section 9, the conclusion.

1.1. Role of Regression and Predictive Modeling

Wafer manufacturing is a highly complex, multi-step process (including lithography, etching, deposition, CMP, implantation, etc.) with thousands of controllable parameters [2,10]. The variability in these parameters strongly affects critical dimension uniformity, overlay accuracy, film thickness, electrical test performance (parametric yield), and overall yield and reliability [2].

Wafer fabrication is highly multivariate, with each step involving numerous controllable factors such as etch time, gas flow, and pressure [3]. Massive, nonlinear, and correlated datasets emerge from metrology, inspection, testing, and in-tool sensors capturing temperature, pressure, and plasma signals [14]. Statistical modeling and machine learning—particularly regression and predictive modeling—enable fabs to convert these complex, high-dimensional data streams into actionable insights, driving process optimization and accelerated technology scaling [20,21].

Regression models serve as a bridge to quantify relationships between process inputs and outputs. They help engineers answer the following questions:

1. How do process conditions (e.g., etch time, temperature, pressure) influence wafer characteristics (e.g., film thickness, CD uniformity)? That is, how do we quantify complex relationships between inputs (recipes, tool settings, material properties) and outputs (film thickness, line width, yield) [4,14]?
2. What is the impact of multiple interacting factors on yield and defect density [4,8]?

Predictive modeling aims to transform these large, noisy datasets into actionable predictions for wafer yield, defect density, process excursions, and equipment health and maintenance needs [22]. Therefore, the intermediate in-line data can be used to predict wafer and die-level outcomes without destructive measurement (virtual metrology) [20]. This helps fabs reduce cycle time, improve yield, and detect deviations and anomalies early in real time (fault detection and classification), without the need for exhaustive measurement [23]. Therefore, it could forecast the health of the forecast tool and wafer yield distributions well before the final test [15,18].

The key differences and complementarity of regression and predictive modeling in semiconductor wafer manufacturing are listed in Table 1.

Table 1. Key differences and complementarity of regression and predictive modeling in semiconductor wafer manufacturing.

Aspect	Regression	Predictive Modeling
Focus	Explains and quantifies relationships between process variables and outputs [2,4,14]	Anticipates outcomes and future states such as yield, defect density, and equipment health [15,18,22]
Methods	Linear, multiple, logistic, partial least squares (PLS) regression [3,4,8]	Regression extended with machine learning: random forests (RF), support vector machines (SVM), deep learning (DL), Bayesian models [20–22]
Applications	Identifying parameter correlations, yield drivers, process window optimization, and virtual metrology (VM) [2,5,14]	Yield forecasting, fault detection and classification (FDC), tool health prediction, digital twins, real-time anomaly detection [15,18,20,23]
Strength	High interpretability, intense statistical rigor, practical for hypothesis testing [4,8,14]	High predictive accuracy, adaptability to nonlinear, high-dimensional, and noisy data streams [20–22]
Limitation	Less effective with highly nonlinear interactions, limited scalability to big data [3,14]	May lose interpretability, requires large datasets and high computing power [20–22]

1.2. The Applications in the Process Steps of Wafer Manufacturing

Some typical examples of the applications in wafer manufacturing are introduced as follows:

a. Lithography: This includes regression models for predicting CD uniformity, focus-exposure process windows, and overlay error [24] and regression models for overlay error prediction based on stage temperature, wafer bow, and lens aberration [24].

b. Etching and Deposition: This predicts etch depth and film thickness uniformity from tool sensor signals [25], optimizing plasma etching by integrating precise three-dimensional etching simulation and machine learning [26].

c. Chemical–Mechanical Polishing (CMP): Predicting the wafer material removal rate for semiconductor chemical–mechanical polishing is performed by a regression model [27]. Five regression models were proposed to predict the material removal of a wafer [28]. Yu et al. [29] used models to predict the material removal rate (MRR) in the chemical–mechanical planarization process, with the asperity radius and asperity density of the polishing pad as the responses.

d. Metrology Data Correlation: Regression links in-line inspection data with final wafer yield [8]. Rana et al. [22] showed that predictive metrology and analytics, which rely on multivariate, nonlinear, and correlated data, are difficult to model mathemati-

cally but can be effectively learned by computing machines for process prediction and control.

e. Electrical Test Data: Regression can be used for predicting die pass/fail before final test [30]. Skinner et al. [11] selected two traditional multivariate statistical methods—principal component methods and regression-based methods—and a classification and regression tree (CART) method to study the effects of the wafer fabrication process on the quality or yield of the wafers.

1.3. Benefits of Regression and Predictive Modeling

The benefits of regression and predictive modeling are listed below.

1.3.1. Early Yield Estimation: Predicts Outcomes Before Wafers Reach the Testing Stage

Wang et al. [31] developed a discrete spatial Bayesian model using wafer map defect data. They employed a hierarchical generalized linear mixed framework to capture global trends and spatial correlations, enabling improved analysis and prediction of wafer yield across spatially clustered die locations.

Ahmadi et al. [32] analyzed production data from two 65 nm RF transceivers to validate a yield estimation method based on wafer measurements, suitable for fab or design migrations. Kumar et al. [33] reviewed probabilistic yield models addressing spatial defects, radial yield loss, and complex process variations, highlighting techniques for accurately modeling interconnected wafer manufacturing processes.

1.3.2. Process Optimization: Identifies Critical Parameters with the Most Significant Impact

Kim et al. [34] introduced an ROI-based productivity model that uses deep neural networks to predict yields from wafer map configurations and employs differential evolution to optimize wafer dimensions. Comparative results demonstrated significant improvements in fab productivity and wafer production capacity [2,34].

1.3.3. Cost Reduction: Reduces Metrology Steps and Wafer Scrap and Enables Predictive Maintenance and Fewer Reworks

Pfitzner et al. [35] proposed two cost-reduction strategies for wafer fabrication: integrating metrology directly into processing tools and expanding wafer recycling. Standardized hardware and software interfaces enable seamless metrology integration, reducing testing costs, transport, and handling while improving tool utilization. Application-specific recycling models with graded wafer quality further enhance material efficiency and overall manufacturing sustainability [35].

1.3.4. Improved Reliability: Detects Subtle Drifts Before They Cause Widespread Failures

Kuo and Kim [36] introduce the realities of yield, reliability, aging, cost factors, and fault coverage in semiconductor manufacturing. They discuss the advantages and disadvantages of various yield models. Establishing a model that encompasses both reliability and yield can serve as a foundation for developing effective stress management strategies to ensure high-quality semiconductor products [36].

Chou et al. [37] combined artificial intelligence (AI) and simulation techniques to evaluate the long-term reliability of wafer-level packages (WLPs). This study applied artificial neural network (ANN) theory to analyze an extensive database of WLP structural reliability data to develop a regression model that correlates WLP structure with reliability. Once the regression model was validated, reliability results could be instantly obtained by simply inputting the geometric parameters of the WLP.

1.3.5. Shorter Ramp-Up: Accelerates Development of New Technology Nodes and Accelerates Learning Cycles for New Nodes

Xu et al. [38] establish a fast-ramp-up framework for the yield improvement model, which enables rapid yield improvements for new products and provides design companies with prompt feedback on failure analysis results. Tirkel [39] applied machine learning (ML) and data mining (DM) methods to develop a cycle time prediction model. Historical production line data from the fab's manufacturing execution system (MES) contained detailed wafer batch processing for various operations.

1.3.6. Higher Yield: Early Detection of Yield Excursions

Dong et al. [40] developed a wafer yield prediction model that integrates spatial defect clustering into functional testing. Using a fusion LASSO algorithm, spatial covariates were derived and combined with test data via logistic regression. Incorporating defect cluster characteristics improved prediction accuracy and enabled reliable yield forecasting for new wafers using historical production data [41,42].

Kumar et al. [43] reviewed the literature on key yield modeling issues, including simple probabilistic yield models and models that incorporate key characteristics such as spatial defects and radial yield loss. The authors emphasized the importance of systematically integrating various factors in yield modeling.

1.3.7. Reduced Scrap: Predict Failing Wafers Before Packaging

Dimaculangan et al. [30] identified the final test as a frequent source of low yield. Using three years of production data, they applied regression analysis with continuous and categorical front- and back-end variables (excluding WAT) to train, validate, and evaluate 11 regressors for predicting final test yield in semiconductor manufacturing.

1.4. Data Landscape in Wafer Manufacturing

Before addressing regression and predictive modeling, it is crucial to understand the nature of data in wafer fabs [41]:

a. Process parameters: The tool settings data are temperature, pressure, RF power, gas flow, and rotation speed. The recipe variables are bake times, etch durations, and polish rates [42–44].
b. In situ sensor data: These time-series signals from plasma optical emission, chamber pressure traces, and motor current profiles often include thousands of correlated variables per wafer run [45,46].
c. Metrology data: This includes critical dimensions (CD-SEM, scatterometry) and film thickness (ellipsometry, X-ray Reflectivity (XRR)) [44,45].
d. Overlay, flatness, defects: They include defect and inspection data and defect counts, wafer maps, defect types (particles, pattern defects, scratches) [42,44].
e. Electrical test and yield data: They include wafer acceptance test (WAT), parametric tests (threshold voltage, leakage currents), final test yield at the die and wafer level, etc. [42,44].

Such data is high-volume, high-dimensional, and hierarchical (wafer-within-lot, die-within-wafer). Regression and predictive modeling are essential for extracting insights and deploying run-to-run control, yield learning, and predictive maintenance [41,45].

2. The Application of Regression in Wafer Manufacturing

Regression models quantify the mathematical relationships between process variables (inputs) and wafer outcomes (outputs) [4,8,14]. They are often the foundation for predictive modeling.

2.1. Process–Parameter Relationship Modeling

It is necessary to establish the relationship between the process variables (etch time, plasma power, deposition rate, temperature, etc.) and wafer-level metrics (oxide thickness, line width, dopant concentration, etc.) [47]. Based on the process–parameter relationships, these critical process parameters are identified [11]. A typical example is the use of multiple linear regression to predict film thickness variation from tool settings [13].

2.2. Metrology Correlation

Metrology correlation in wafer manufacturing leverages regression and predictive modeling to relate inline and end-of-line measurements, minimizing costly inspections [11]. Regression models quantify relationships between process parameters, metrology readings, and product quality, while predictive analytics forecast unseen measurements [48]. This enables virtual metrology, rapid feedback, and proactive control, improving yield, reducing cycle time, and enhancing tool performance monitoring in advanced semiconductor fabrication [14].

Zeng [6] used regression to correlate in-line metrology (critical dimension, overlay, thickness) with end-of-line electrical test results. For example, partial least squares (PLS) regression is used to relate spectra from scatterometry with actual CD measurements [6].

Krueger et al. [49] proposed a method to identify which data should be collected, integrated, and summarized, and to predict yield using a generalized linear model based on defect metrology data. This technique identifies important factors driving process improvement. It allows the model to account for the nested structure of the process, thereby improving predictive power and enabling its application to situations where assumptions are violated.

2.3. Fault Detection and Excursion Analysis

Regression models can be used to reveal abnormal deviations between expected and actual performance, aiding in the identification of root causes. A typical example is to use regression residuals to highlight when a chamber drifts out of calibration [50].

Lee and Kim [51] defined fault detection (FD) in semiconductor manufacturing processes as a feature extraction, feature selection, and classification algorithm. They tested FD models on six datasets using 117 possible algorithm combinations. Statistical analysis was conducted from both an algorithmic perspective and an FD model perspective. Multiple comparisons of algorithms were conducted to evaluate the performance of competing algorithms in the three modeling steps, and the strengths and weaknesses of each algorithm were discussed [51]. Chen et al. [52] used statistical metrics to convert physical data from state variable identification (SVID) into fault detection and classification parameters (FDC parameters). The authors developed a multivariate analysis model to identify key factors that may cause process deviations from a large amount of FDC data. The proposed two-stage multivariate analysis framework included applying the least absolute shrinkage and selection operator (LASSO) for key operation screening and using random forest (RF) to rank the FDC parameters based on the key operations [50,52].

A summary of the applications of regression in wafer manufacturing is listed in Table 2.

Table 2. Applications of regression in wafer manufacturing.

Application Area	Description and Role of Regression	Examples/Methods	References
2.1. Process–parameter relationship modeling	Quantifies mathematical relationships between process variables (etch time, plasma power, deposition rate, temperature) and wafer-level metrics (oxide thickness, line width, dopant concentration). Identifies critical parameters driving process performance.	Multiple linear regression to predict film thickness variation from tool settings.	[11,13,47]
2.2. Metrology correlation	Relates inline metrology (critical dimension, overlay, film thickness) with end-of-line electrical test results. This system enables virtual metrology, rapid feedback, and proactive process control. This process improves yield, reduces cycle time, and enhances tool monitoring.	PLS regression to link scatterometry spectra with CD; regression to connect process parameters, metrology readings, and product quality; GLM on defect metrology for yield prediction and data integration strategies.	[6,11,14,48,49]
2.3. Fault detection and excursion analysis	Uses regression residuals to reveal deviations between expected vs. actual performance. This system supports excursion analysis, fault detection, and identification of root causes of low yield and process drifts.	Regression residuals for chamber drift detection; data-driven diagnostic approaches for low yield; FD models with feature extraction, selection, and classification; two-stage frameworks with LASSO and random forest for FDC parameter ranking.	[50–52]

3. The Application of Predictive Modeling in Wafer Manufacturing

Predictive modeling expands beyond regression to include advanced machine learning and statistical forecasting. The focus is on anticipating wafer quality, yield, or tool health before issues arise.

3.1. Yield Prediction

Regression helps identify which variables most significantly affect wafer yield and supports real-time yield monitoring. It involves predicting wafer/die yield using defect density, metrology data, and test parameters [33,53].

Predicting the final wafer yield could optimize yield and ensure optimal results early.

Logistic regression is an example to classify wafers as pass/fail based on defect density and parametric test data [54].

Because early detection of low yields related to wafer materials is crucial for effectively controlling cost and quality, Jiang et al. [53] proposed using a Gaussian mixture model (GMM) clustering approach to develop a weighted ensemble regressor to predict back-end final test (FT) yield during the wafer fabrication (WF) stage. This approach was validated on actual production data from a new wafer product line, demonstrating significant improvement in prediction performance [53].

3.2. Equipment Health Monitoring and Predictive Maintenance

Predictive models are used to anticipate tool failures before they occur using sensor logs and historical maintenance data. Some examples are the prediction of chamber clean frequency based on RF power stability and endpoint signal drift; predictive models forecast tool drift and failures and are used in predictive maintenance (PdM) to avoid costly downtime [55,56].

Equipment health monitoring and predictive maintenance in wafer manufacturing use predictive modeling to track tool performance, detect anomalies, and prevent failures. By analyzing sensor data, vibration signals, and historical maintenance records, machine learning models identify degradation patterns and predict component lifetimes [57]. This approach enables condition-based maintenance, reduces unplanned downtime, extends

equipment life, and ensures process stability, supporting higher yield, cost efficiency, and reliability in advanced semiconductor manufacturing environments [56,58].

3.3. Defect Density and Pattern Prediction

Defect density and pattern prediction in wafer manufacturing leverage predictive modeling to identify defect sources, spatial patterns, and yield risks early [23]. Using inspection data, wafer maps, and process parameters, machine learning detects correlations and recurring defect signatures [59]. This enables rapid root cause analysis, targeted corrective actions, and reduced variability, improving yield, quality, and manufacturing efficiency [60].

It is possible to use predictive models to generate wafer defect maps based on lithography, CMP, or etch parameters, predict wafer-level defect hotspots using tool/process data, and correlate equipment signals with their impact on wafer-level yield [23]. The primary techniques are machine learning classifiers and spatial predictive models [59,60].

3.4. Process Control and Run-to-Run Optimization

Process control and run-to-run (R2R) optimization in wafer manufacturing leverage predictive modeling to maintain stability, reduce variability, and improve yield across complex process steps [61,62]. Predictive models analyze metrology data, tool parameters, and historical trends to adjust recipes dynamically between wafer runs. Machine learning enhances fault detection, drift compensation, and excursion prevention, enabling fabs to respond proactively to process changes [62]. Integrated with advanced process control systems, these models optimize throughput, minimize scrap, and ensure consistent device performance, supporting high-volume manufacturing at advanced technology nodes [62].

Wan et al. [63] proposed that integrating virtual metrology (VM) with batch-to-batch (R2R) control can maintain the advantages of R2R control while avoiding the negative impact of physical metrology on cost and cycle time. They proposed using a Gaussian process regression (GPR) model in VM-enabled R2R control to provide this information in an integrated manner. They also demonstrated the effectiveness of the GPR-enabled VM R2R control approach using a case study of a chemical–mechanical polishing process.

3.5. Reliability and Lifetime Forecasting

Reliability and lifetime forecasting in wafer manufacturing applies predictive modeling to anticipate equipment wear, process drift, and product degradation over time [64]. By analyzing sensor data, maintenance logs, and historical failure patterns, models such as survival analysis, Weibull modeling, and machine learning predict tool replacement intervals and product reliability [65,66]. These insights enable proactive maintenance, minimize downtime, and extend the lifespan of equipment [67]. Integrated forecasting improves process stability, reduces costs, and ensures consistent wafer quality, supporting advanced-node manufacturing with higher efficiency, yield, and operational resilience [64,66]. The typical application is to predict transistor/device degradation based on process variations and test data and improve design-for-reliability (DfR) strategies [68].

Lenhard et al. [69] introduce a screening method for wafer-level inline defect detection. This method utilizes an advanced prediction engine to generate wafer-level failure probabilities, screening out wafers with higher reliability risks when the predicted probability exceeds a predefined limit. A saliency graph clustering algorithm is employed to increase the granularity of latent defect detection beyond supervised defect classification.

3.6. Virtual Fabrication and Digital Twin

Digital twin models replicate semiconductor processes by integrating physical models, sensor data, and AI-based analytics for real-time monitoring and optimization. Model

construction involves synchronizing virtual and physical systems through continuous data streams [20–23]. Validation is achieved via historical data matching, real-time feedback, and performance benchmarking. However, limitations arise from data latency, model drift, computational demands, and incomplete physics-to-data coupling, which can reduce accuracy and scalability in complex, multistage wafer fabrication environments [20–23].

Virtual fabrication and digital twin technology in wafer manufacturing leverages predictive modeling to create high-fidelity simulations of processes, tools, and wafer behavior [2]. These digital replicas integrate physics-based models with machine learning to optimize recipes, predict yield outcomes, and evaluate process changes without physical trials [70]. Engineers can run "what-if" scenarios, accelerating development and reducing costly experiments [71]. Real-time data feeds continuously update the twin, enabling adaptive control, early fault detection, and process optimization, ultimately improving efficiency, product quality, and fab decision-making [71,72].

Nguyen et al. [73] proposed a physical model at a characteristic scale that uses the roughness and elastic deformation of the polishing pad to calculate the pressure distribution. This topography simulation module is coupled with various tools, including process simulators and electrical characterization modules. This allows for integrated 3D simulation of barrier and copper deposition, making it suitable for dam scenes. Graves et al. [74] identified plasma–tool digital twins to track reactor state, guide control, and predict maintenance, which is crucial for equipment-level twins in fabs.

3.7. Design–Technology Co-Optimization (DTCO)

Design–technology co-optimization (DTCO) in wafer manufacturing uses predictive modeling to align chip design and process technology for optimal performance, power, and cost [75]. By integrating layout-aware simulations, process variability models, and manufacturing constraints early in the design process, DTCO identifies trade-offs between design rules, materials, and device architectures [76]. Machine learning enhances the prediction of yield, performance, and manufacturability, reducing design iterations [77]. This synergy accelerates technology node development, improves yield ramp, and ensures robust designs that fully leverage advanced lithography and process innovations, driving efficient semiconductor scaling [75,77].

Chen and Chen [78] explore the application of generative adversarial networks (GANs) in the chip manufacturing process to generate wafer-level wafer acceptance test (WAT) and wafer probe (CP) test data, emphasizing design–technology co-optimization (DTCO). The generated virtual silicon wafer data contains essential performance characteristics, physical electrical properties, wafer-level process parameter distributions, and implicit information about wafer-level uniformity and defects. This enables the collaborative optimization of the manufacturing and chip design processes [78].

A summary of the applications of predictive modeling in wafer manufacturing is listed in Table 3.

Table 3. Applications of predictive modeling in wafer manufacturing.

Application Area	Description and Role	Methods/Examples	References
3.1. Yield prediction	Identifies variables most affecting yield and supports real-time monitoring. Predicts wafer/die yield from defect density, metrology, and test parameters. This process enables early detection of low-yield wafers to optimize cost and quality.	Logistic regression to classify wafers as pass/fail; GMM clustering + weighted ensemble regressor for FT yield prediction using WF data.	[33,53,54]

Table 3. *Cont.*

Application Area	Description and Role	Methods/Examples	References
3.2. Equipment health monitoring and predictive maintenance	The system anticipates tool failures using sensor logs, RF power, endpoint drift, vibration signals, and maintenance records. Enables condition-based maintenance, prevents downtime, and ensures process stability.	Predict chamber clean frequency; ML models (ANN, SVR, MLP, RF, ARIMA) tested on fab data; degradation pattern recognition; PdM frameworks.	[55–58]
3.3. Defect density and pattern prediction	Predicts defect sources, spatial patterns, and recurring defect signatures using inspection and wafer map data. Supports rapid root cause analysis and implementation of corrective actions.	Machine learning classifiers; spatial predictive models; wafer defect maps from lithography, CMP, etch; correlation of tool/process data with yield impact.	[23,59,60]
3.4. Process control and run-to-run (R2R) optimization	Maintains process stability and reduces variability via predictive modeling integrated into APC systems. Dynamically adjusts recipes using metrology and tool data.	VM-enabled R2R control with GPR models; metrology delay analysis; drift compensation; batch-to-batch optimization.	[61–63]
3.5. Reliability and lifetime forecasting	Forecasts equipment wear, process drift, and product degradation. Supports DfR strategies and predictive maintenance to minimize downtime.	Survival/Weibull models; ML for TDDB, BTI, EM; wafer inline defect screening (prediction engine + clustering); RF + FEM for WLP reliability lifecycle.	[64–69]
3.6. Virtual fabrication and digital twin	Creates high-fidelity digital replicas of processes, tools, and wafers. Integrates physics + ML to simulate recipes, yield outcomes, and "what-if" scenarios.	FEOL/BEOL transistor prediction; 3D CMP simulation (pad deformation, pressure); plasma–tool digital twins for reactor control/maintenance.	[2,70–74]
3.7. Design–Technology Co-Optimization (DTCO)	Aligns chip design with process constraints to optimize performance, power, cost, and yield. ML accelerates prediction of manufacturability and design-yield trade-offs.	Layout-aware variability models; GANs for WAT/CP data generation; virtual wafer data for co-optimization of process + design.	[75–78]

4. Regression Methods in Wafer Manufacturing

4.1. Regression Models

4.1.1. Linear Regression

Ordinary Least Squares (OLS) regression is the simplest and historically most widely used tool. It models a dependent variable (e.g., oxide thickness) as a linear combination of predictors (e.g., deposition time, temperature, gas flows) [5]. While easy to interpret, linear regression assumes independence and homoscedasticity, which are often violated in real fab data (due to wafer-to-wafer correlation, tool drifts, and hierarchical nesting) [12]. It is frequently used for DOE and initial process modeling [11].

The deposition rate is modeled as a function of chamber pressure and power, allowing for control of film thickness [79]. Critical dimensions (CDs) in lithography are modeled as a function of exposure dose and focus, which helps define the process window using a crucial dimension linear model [80]. Defect counts or parametric yield can be regressed against process settings across lots with a yield loss correlation equation [12,49].

Krueger et al. [49] proposed a modeling strategy for yield prediction using generalized linear models based on defect metrology data. This technique can identify important factors that drive process improvements. Purwins et al. [79] evaluated the use of high-dimensional multivariate input data against a small parameter set, comparing simple linear regression, multiple linear regression, partial least squares regression, and ridge linear regression. Regression parameter optimization and model selection were performed and evaluated using root mean square error (RMSE) through cross-validation and grid search. The results suggest that virtual metrology can benefit from regression methods that exploit collinearity and incorporate process expertise [79].

4.1.2. Multiple Regression

As wafers involve multi-factor interactions, multiple regression is widely adopted. The functions of multiple regression include incorporating multiple input variables simultaneously [13,81], helping capture multi-factor interactions in processes such as lithography focus/exposure or etch gas composition, capturing interactions between multiple parameters, and being used for process window analysis (e.g., lithography focus × dose) [13].

Mahandran et al. [82] investigated factors contributing to oxide thickness instability. The coefficient of determination, R^2, was used to indicate a good model prediction. Significant influencing factors were gas flow rate, oxidation time, and temperature. Multiple regression was used to determine the optimal operating factor. Shumate et al. [83] investigated the characteristics of SF_6/sub$_6$He plasmas for TiW etching using fractional factors and a mixed response surface model. Factors influencing the process included RF power, reactor pressure, and the SF_6/He gas ratio. A quadratic/cubic model captured the interaction between the gas ratio and pressure. The model employed both quadratic and special cubic response surface models. Contour lines of this empirical model were used for joint analysis to optimize the process. Liu et al. [84] developed a chemical–mechanical polishing process at low downforce. They employed response surface methodology (RSM) to optimize the slurry composition, which included silica sol, H_2O_2, and an FA/O chelating agent. A central composite design (the standard design for RSM) was used to evaluate the influence and interaction of the three factors. Ultimately, the optimal conditions for silica sol concentration, H_2O_2 concentration, and FA/O chelating agent concentration were determined [84].

4.1.3. Polynomial Regression

Because semiconductor processes are rarely purely linear, polynomial regression is used to approximate nonlinear relationships. For example, the photoresist CD shrinkage during baking is nonlinear with respect to time and temperature, and the ion implantation dose–response curves follow quadratic-like patterns [85,86]. For example, plasma etch rate is modeled as a function of RF power, gas mixture ratios, and pressure, and interaction terms reveal synergies (e.g., pressure × gas ratio) [87].

Polynomial regression is suitable for nonlinear processes such as chemical vapor deposition (CVD) or CMP dishing [88]. Polynomial regression improves accuracy over simple linear models but risks overfitting without careful validation [84]. Liu et al. [84] employ RSM with second-order polynomial models to capture nonlinear CMP behavior and optimize slurry composition for a route commonly used to control planarization metrics like dishing/erosion.

4.1.4. Logistic Regression

Logistic regression in wafer manufacturing predicts binary or categorical outcomes, such as the presence of defects, pass/fail classification, or yield risk. It identifies key process variables that influence defects, supports early decision-making, and optimizes quality control by modeling the relationships between process parameters and product quality, thereby enhancing predictive maintenance and excursion detection [40].

Bae et al. [89] used models based on Poisson regression, negative binomial regression, and zero-inflated Poisson (ZIP) regression, with the spatial location of each IC die and the corresponding number of defects in the wafer map as dependent variables. Their model significantly improved yield prediction accuracy by capturing the spatial distribution of defects on the wafer map. Jizat et al. [90] established and evaluated the best machine learning classifier for wafer defect inspection applications. Logistic regression was one of the evaluated models, achieving high accuracy in classifying pass/fail defects.

4.1.5. Nonlinear Regression

Nonlinear regression in wafer manufacturing models involves complex, nonlinear relationships between process variables and outputs such as critical dimensions, film thickness, and yield. It enables accurate process optimization, virtual metrology, and root cause analysis, capturing intricate tool–material interactions, improving parameter tuning, and supporting predictive control for advanced technology nodes [91]. It is suitable for highly nonlinear processes (e.g., plasma etch rate vs. gas flow). Typical examples include diffusion processes, where dopant concentration profiles follow error function shapes [92].

Jia et al. [93] employed a nonlinear polynomial neural network VM model for CMP removal rate prediction, which supports adaptive feature and model selection. They found that the highly nonlinear CMP dynamics were well-suited for the virtual metrology in semiconductor manufacturing. Nami et al. [94] measured the deposition rate under a range of deposition conditions using statistically designed experiments. A semi-empirical metal–organic chemical vapor deposition (MOCVD) model based on a "hybrid" neural network is proposed to characterize the MOCVD properties of titanium dioxide (TiO_2) thin films. They demonstrate a nonlinear neural network regression for CVD film growth, capturing intricate process–material interactions beyond linear models [94].

Huang et al. [95] derived a relationship between the polar component of film stress and the polar component of system curvature in the presence of arbitrary radial inhomogeneities, inferring such stresses from full-field curvature measurements. The relations linking film stress to temperature fields and resulting curvature (bow) were modeled in a nonlinear equation [95].

4.1.6. Principal Component Regression (PCR)

Principal component regression (PCR) in wafer manufacturing reduces multicollinearity among highly correlated process variables by transforming them into orthogonal principal components [25,96]. It enables robust modeling of yield, defect density, and critical dimensions, enhances virtual metrology, improves fault detection, and supports process optimization in high-dimensional, data-rich semiconductor fabrication environments [25,96].

Phatak et al. [97] showed that PCR models relating sensor features to CMP material removal rate outperform alternatives and demonstrated robust modeling and process optimization with many correlated inputs. Jeong et al. [98] built plasma etch prediction models using PCR on high-dimensional data, compared them with neural networks, and illustrated VM and multicollinearity reduction in practice.

4.1.7. Partial Least Squares Regression (PLSR)

PLSR is particularly important in wafer manufacturing because it handles collinear and high-dimensional input spaces and extracts latent variables linking tool signatures to outputs [99,100]. Some applications include spectroscopic ellipsometry, which models film refractive indices and thickness from multichannel spectra, and plasma emission monitoring, which is used to predict etch profiles from thousands of spectral intensities [99]. Pan et al. [100] proposed a robust linear regression method based on variable selection for predicting final quality variables required in complex industrial processes. Initial model coefficients were generated using partial least squares regression, and then the non-negative Gallot method was used to shrink the original coefficients, implicitly eliminating irrelevant variables. This method significantly improved both the accuracy and robustness of variable selection. Khan et al. [101] applied partial least squares (PLS) modeling techniques to develop a linear regression model for the underlying process of VM. They also created a recursive moving window method to update the VM model as metrology data became

available. Then they employed batch-by-batch control techniques to develop W2W process control capabilities.

4.1.8. Regularized Regression (Ridge, Lasso, Elastic Net)

Regularized regression methods, such as ridge, Lasso, and Elastic Net, are vital in wafer manufacturing for handling high-dimensional, collinear process data. Ridge stabilizes models by shrinking coefficients, Lasso performs feature selection to identify critical parameters, and Elastic Net combines the benefits of both [102].

These techniques enhance defect prediction, virtual metrology, yield modeling, and fault detection, enabling the development of robust and interpretable models that generalize well, support advanced process control, and optimize performance in data-intensive semiconductor fabrication environments [103].

Ridge regression stabilizes coefficients by shrinking them toward zero [102]. Lasso regression performs variable selection, identifying critical tool signals among thousands [39,104,105].

Chen and Leclair [106] proposed using kernel ridge regression (KRR) to tune multi-input and multi-output dry etching recipes, thereby generating optimal recipes for multiple-input, multiple-output (MIMO) systems. They also developed a KRR parameter optimization method. Lee et al. [107] proposed an Elastic Net regularization that enforces sparsity in latent factors to reduce multicollinearity and improve variable selection.

4.1.9. Hierarchical and Mixed-Effect Regression

Hierarchical and mixed-effect regression models variability in wafer manufacturing across nested structures, such as wafers, lots, and tools, capturing both fixed process effects and random variations [108]. This approach improves yield prediction, process optimization, and equipment health monitoring by accounting for batch-to-batch and tool-to-tool variability. It enhances root cause analysis, enables more accurate virtual metrology, and supports advanced control strategies in high-volume manufacturing environments where multi-level data structures and correlated measurements are common [109]. Yuan et al. [110] proposed hierarchical Bayesian regression models to predict yield for wafer map defects within and across wafers. Wang et al. [31] propose a hierarchical generalized linear mixed model to incorporate global trends and spatially correlated random effects on wafer maps. They found that the proposed model provides a better fit for spatially correlated wafer map data.

Liu et al. [111] developed a profile-based statistical process control scheme to monitor wafer thickness profiles with abnormal distributions during industrial wafer slicing. The authors propose a mixed-effect profile monitoring (MEPM) scheme that adaptively groups profile data into clusters and models inter-cluster variation, resulting in a robust statistical process control scheme for detecting abnormal profile data. The comparison of these regression models is listed in Table 4.

Table 4. Comparative overview of regression methods in semiconductor wafer manufacturing.

Regression Method	Purpose/Description	Advantages in Wafer Manufacturing	Limitations/Challenges	Literature
Linear Regression (OLS)	Models a dependent variable as a linear combination of predictors (e.g., oxide thickness vs. process parameters).	Simple, interpretable; helpful for DOE and first-order process understanding.	Sensitive to multicollinearity and heteroscedasticity; poor for nonlinear effects.	[49,79]
Multiple Regression	This method extends OLS by modeling multiple inputs simultaneously to capture multi-factor interactions.	Captures interdependencies between process variables; supports DOE and RSM optimization.	Requires careful factor selection; collinearity can inflate variance.	[82,84]

Table 4. *Cont.*

Regression Method	Purpose/Description	Advantages in Wafer Manufacturing	Limitations/Challenges	Literature
Polynomial Regression	Fits nonlinear relationships using polynomial terms of predictors (e.g., quadratic, cubic).	This approach captures nonlinear behavior in etch, CMP, and CVD processes, improving accuracy over linear models.	Risk of overfitting and instability; requires cross-validation.	[84,88]
Logistic Regression	Predicts categorical or binary outcomes (e.g., pass/fail, defect/no defect).	This tool enables yield risk estimation, defect prediction, and quality classification.	Limited to binary/categorical outcomes; nonlinear decision boundaries.	[89,90]
Nonlinear Regression	Models nonlinear relationships between variables and outputs (e.g., etch rate, stress, yield).	Captures complex process behaviors (e.g., diffusion, plasma chemistry); supports virtual metrology.	Computationally intensive; may converge to local minima.	[93–95]
Principal Component Regression (PCR)	This method reduces multicollinearity by transforming correlated predictors into orthogonal principal components.	Effective for high-dimensional tool data (sensor signals, spectra); enhances model stability.	Principal components may lack direct physical meaning; may exclude minor but important factors.	[97,98]
Partial Least Squares Regression (PLSR)	Extracts latent variables that maximize covariance between predictors and outputs.	Robust for spectral and plasma data; enables real-time VM and R2R control.	Model interpretation is complex; it requires latent variable tuning.	[100,101]
Regularized Regression (Ridge, Lasso, Elastic Net)	Adds penalty terms to reduce overfitting and stabilize high-dimensional models.	This approach handles multicollinearity, identifies key process variables, and enhances generalization.	Requires hyperparameter tuning; interpretability may be reduced.	[106,107]
Hierarchical/Mixed-Effect Regression	Models multi-level data structures (e.g., wafer → lot → tool) with fixed and random effects.	Accounts for lot-to-lot and tool-to-tool differences; improves yield and defect prediction.	Complex parameter estimation; computational cost for large datasets.	[31,110,111]

4.2. Generalization of the Regression Model

Regression models—whether linear, polynomial, or regularized—generalize only moderately well across heterogeneous fab datasets without deliberate harmonization and adaptation. The challenge arises because wafer manufacturing data are not identically distributed: they vary by tool, time, and context. These datasets are characterized by data heterogeneity, non-stationarity, multicollinearity, and confounding due to measurement system differences, sampling bias, and domain mismatch. The most effective treatments combine hierarchical regression, domain adaptation, and physics-constrained learning, augmented by digital twin calibration and drift-aware retraining. This layered approach maintains interpretability while achieving robustness—turning regression from a static estimator into a dynamic, transferable model aligned with real fab variability.

Some regression models that transfer robustly are introduced as follows:

Lynn et al. [91] investigated the application of global and local virtual metrology from an industrial plasma etch chamber and found that the Gaussian process regression models provided accurate results, capable of producing valid estimates of plasma etch rates over multiple chamber maintenance events and many thousands of wafers. Some regression cases achieved robust transferability. Melhem et al. [102] presented a survey of regularized linear regression methods that utilize feature reduction and variable selection to address multicollinearity in high-dimensional input data. These methods could predict wafer quality based on production equipment data, demonstrating model robustness and prediction accuracy. Yu et al. [29] introduced a novel physics-informed machine learning approach, developing physics-based and data-driven regression models to predict the material removal rate (MRR) in the CMP process. Experimental results revealed that the physics-informed model could predict MRR with high accuracy. Miyaguchi et al. [112] presented a novel framework, partial trajectory regression (PTR, to address the limitations of conventional vector-based regression models for wafer defect root cause analysis. This model was validated for its effectiveness using real wafer history data from the NY CREATES fab in Albany.

Some regression cases that failed to generalize are introduced as follows:

Orji et al. [113] proposed that barriers to virtual metrology (VM) adoption include low confidence in regression models, poor data quality, limited process understanding, weak metrology correlation, high maintenance costs, and IP concerns. Key solutions emphasize developing standardized data and model quality metrics, improving data completeness and communication, and establishing model maintenance frameworks to ensure robustness under process drift. Regression models remain central to VM by linking process parameters to wafer outcomes; however, their reliability depends on transparent quality evaluation, validated metrology correlation, and adaptive retraining. Kim and Kang [114] investigated the effect of irrelevant variables on three learning algorithms for regression tasks—artificial neural network, decision tree (DT), and k-nearest neighbors (k-NN)—and collected datasets from a semiconductor manufacturer to examine how the number of irrelevant variables affects the performance of prediction models. The results indicate that higher model complexity of learning algorithms leads to a higher sensitivity to irrelevant variables. However, low sensitivities are preferred as a first trial for building prediction models.

Maitra et al. [16] surveyed the state-of-the-art VM works for prominent processes involved in semiconductor manufacturing, and the unique contributions of those VM predictive works are summarized. They found that the VM models were significant hurdles that prevent VM implementation and acceptance in the semiconductor industry. These VM models need to be trained faster and be more compatible with the existing VM models in the factory.

4.3. Demonstration of the Regression Methods

4.3.1. Virtual Metrology

Wan and McLoone [63] introduced a Gaussian process regression (GPR) technique that consistently surpasses classic linear/NN baselines on benchmark VM datasets and provides uncertainty bands useful for R2R. Comparative studies and CMP-focused work show that this GPR tracked drifts well. Ishikawa et al. [115] applied recursive PLS to a plasma etcher with time-varying endpoint characteristics, showing improved prediction of etch rate and uniformity under continuous process drift. The model is updated via exponentially weighted recursive least squares, emphasizing recent runs while gradually discounting outdated samples. Similarly, Lee et al. [116] employed LW-PLS in oxide and poly-etch VM, using local similarity weighting to adapt to process condition changes across different chambers, achieving more stable residuals and reduced RMSE over extended production periods.

4.3.2. Yield Prediction

Lee and Roh [117] establish an expandable machine learning framework for semiconductor yield prediction that integrates heterogeneous process data. After data preprocessing, multiple regression models were optimized and compared to identify the most effective predictor, with random forest (RF) regression achieving the best performance. The model captures complex, nonlinear dependencies among fabrication variables and demonstrates strong generalization across diverse process stages. It provides actionable insights for adaptive process optimization and data-driven production management across scalable semiconductor fabrication environments.

Amato et al. [43] established a data-driven framework to optimize semiconductor testing by predicting final yield from Scanning Electron Microscope (SEM)-detected defects. Two modeling approaches were developed: an Odds Ratio-based model to rank wafer layers most correlated with electrical failures, guiding targeted SEM inspections, and a gradient boosting regression/classification model to predict device yield from defect distributions. The gradient boosting model validated the Odds Ratio findings and demon-

strated strong predictive capability despite sparse data. Both models were validated on two STMicroelectronics datasets, enhancing inspection efficiency and yield prediction accuracy.

4.3.3. Excursion Detection

Puggini and McLoone [118] proposed a variable selection and Isolation Forest–based anomaly detection model for analyzing high-dimensional optical emission spectroscopy (OES) data in semiconductor manufacturing. The framework could perform dimensionality reduction while retaining critical isolated variables often missed by conventional PCA methods. By validating through simulated and industrial plasma etch case studies, the approach demonstrates strong capability for efficient, interpretable, and data-driven anomaly diagnosis in production environments.

Kazemi et al. [119] established an adaptive neural-based Principal Component Analysis (PCA) framework for fault detection and diagnosis in time-varying industrial processes. This model recursively updates the correlation matrix, the number of principal components, and confidence limits for the monitoring indices, enabling dynamic adaptation to process variations, and is validated through three studies. The result indicated that the model exhibits superior detection accuracy and diagnostic reliability compared to conventional PCA, enhancing real-time monitoring and fault isolation in complex dynamic systems.

5. Predictive Modeling in Wafer Manufacturing

5.1. Virtual Metrology (VM) Models

Virtual metrology models predict wafer properties without physical measurement, using tool sensor data and regression models. VM reduces costly metrology steps, enables 100% wafer monitoring, and improves run-to-run control [7]. It predicts metrology outcomes (e.g., CD, thickness, overlay) using statistical or machine learning regression models trained on tool/sensor data, thereby reducing reliance on time-consuming and expensive inline metrology [9]. The application examples include predicting post-etch CD from chamber pressure, RF power, and in situ spectral data [120], and predicting CMP film thickness from platen pressure and slurry parameters [121,122].

5.2. Yield Prediction Models

Yield is the ultimate KPI in wafer manufacturing [110]. Predictive models forecast yield before wafers finish processing and enable adaptive lot dispositioning (deciding early whether to scrap or continue wafers) [53,89].

Some examples include the wafer acceptance test (WAT) parametrics regressed against final test yield [123], defect density mapping to predict die yield distributions [124], and hierarchical regression modeling of wafer-within-lot correlations [49].

Spatial covariates are obtained from the fused LASSO on the wafer map, and logistic regression is used to combine these derived spatial covariates with functional test values to predict the final yield, thereby improving accuracy by exploiting the clustered defect structure [40].

5.3. Fault Detection and Classification (FDC) Models

Fault detection and classification (FDC) models in wafer manufacturing monitor tool sensors and process data to detect anomalies early, classify fault types, and prevent excursions [125]. They enable predictive maintenance, improve process stability, and reduce scrap by identifying subtle deviations in real time, supporting higher yields, equipment reliability, and robust manufacturing control in semiconductor fabs [126,127].

Some models of the FDC systems include the regression residuals to signal deviations (e.g., predicted CD differs significantly from the measured CD); machine learning classifiers

predict whether a tool run will be "normal" or "abnormal", and time-series predictive models anticipate drifts before out-of-control events [128].

Kim et al. [128] evaluated seven novelty-detection machine learning methods applied to real-world FDC data for wafer fabrication and tackled challenges such as high-dimensional input (150+ variables) through dimensionality reduction, achieving high true-positive rates in detecting faulty wafers. Fan et al. [129] used traditional and multivariate ML techniques to build predictive models for equipment fault detection in wafer fabrication. They emphasized real-time predictive maintenance to maintain high yields through rapid anomaly detection.

5.4. Time-Series Predictive Models

Time-series predictive models in wafer manufacturing analyze sequential process and sensor data to forecast tool performance, process drift, and yield trends [130]. They enable proactive fault detection, predictive maintenance, and run-to-run control by capturing temporal patterns and correlations, ensuring process stability, minimizing downtime, and optimizing throughput in highly automated, data-intensive semiconductor fabrication environments [131].

Chen and Blue [132] developed a sensor trajectory health metric based on time series and generalized moving variance to predict equipment performance degradation, enabling predictive maintenance. This was validated using case studies of plasma-enhanced chemical vapor deposition and physical vapor deposition tools. A hierarchical, recipe-independent monitoring system based on multi-stream time-series sensor data is developed for proactive fault detection and tool status prediction [133].

5.5. Hybrid Models

Hybrid models integrate physics-based equations with data-driven algorithms to enhance predictive accuracy and interpretability in semiconductor process modeling. Model construction involves coupling physical laws with machine learning structures, enabling generalization beyond the training domain [134–136]. Validation typically employs cross-validation, residual analysis, and comparison with physical benchmarks to ensure robustness. However, limitations include high computational cost, complexity in parameter tuning, and challenges in balancing physical fidelity with data flexibility, particularly when process mechanisms or sensor data are incomplete [134–136].

Hybrid models in wafer manufacturing combine physics-based simulations with data-driven machine learning to enhance process understanding and prediction [134]. They enable accurate virtual metrology, defect prediction, and yield modeling by leveraging domain knowledge and large datasets [134,135]. This integration improves root cause analysis, accelerates process optimization, and supports advanced control strategies in complex semiconductor fabrication workflows [135].

Deivendran et al. [122] proposed a novel hybrid modeling framework for chemical–mechanical polishing that uses machine learning models (random forest, Lasso, SVM, XGBoost, etc.) to accurately predict slurry temperature, mean abrasive particle size (MAPS), and material removal rate (MRR). Model-based optimization (via particle swarm optimization) is used to determine the optimal operating conditions [136]. Liu introduced a hierarchical modeling approach for virtual wafer metrology based on machine learning [136]. This approach leverages batch-to-batch (RTS) and post-processing quality characteristics to enhance wafer property prediction. The authors propose that this approach can capture the structure of physical information and enable data-driven adaptability, thereby improving the accuracy of virtual metrology [137].

6. Integration with Machine Learning

Machine learning models include decision trees/random forests to capture nonlinear interactions, support vector machines (SVM) to classify wafers/dies as good/bad, and neural networks (NNs) to predict yield from complex multivariate data [138].

Beyond regression, fabs now use machine learning-based predictive models.

Random forests and gradient boosting are used to handle nonlinearities and interactions. Neural networks are used to capture complex sensor–output mappings, while support vector regression (SVR) is effective in high-dimensional metrology data [139]. Deep learning, including CNNs, is applied to wafer map defect classification, and RNNs are used for time-series tool data [140].

Shih et al. [141] used decision trees and random forests to classify and predict white-pixel defects in Complementary Metal–Oxide–Semiconductor (CMOS) wafers. They identified critical process factors (e.g., grinding load/torque) and achieved high predictive accuracy (~97%) using interpretable tree-based models. Choi et al. [142] combined handcrafted features (e.g., radon, geometric, density) with an ensemble approach using a stacking model that included base-level classifiers and extreme gradient boosting (XGBoost) as the meta-level predictor. It is particularly effective in classifying wafer map defect patterns—even under severe class imbalance and limited data—and demonstrates how gradient boosting handles nonlinear interactions well. Jeong et al. [143] review a range of traditional ML methods, including SVM, applied to features such as radon transforms and geometry for wafer map defect classification. They explored how convolutional neural networks (CNNs) are increasingly used for deep learning-based defect pattern classification, showing performance improvements and automation of feature extraction.

The clear advantages of hybrid physics–ML models over purely data-driven or purely physics-based approaches in wafer manufacturing include predictive modeling of material removal rate in chemical–mechanical planarization [29,144], improved accuracy on complex 3D effects of lithography [145], prediction of fan-out wafer-level package [146], plasma etch rate optimization [147], and prediction of Si etch profile [148].

6.1. Support Vector Regression (SVR)

Support vector regression (SVR) is used in wafer manufacturing to predict nonlinear critical dimensions (CD) with high accuracy [3,20]. By mapping process variables into higher-dimensional spaces, SVR captures complex relationships between tool settings and CD variations. It supports virtual metrology, enhances process optimization, and reduces inspection costs, enabling tighter control in advanced technology nodes [25].

Dailey et al. [149] proposed a new approach for virtual metrology (VM) of plasma etch processes. They used full optical emission spectroscopy (OES) signatures and regression analysis, including support vector regression (SVR), to predict post-etch critical dimensions (CDs) for plasma etch VM. The proposed SVD-based OES signature consistently outperformed the baseline VM model signature.

Kang et al. [150] introduced a self-training-based semi-supervised support vector regression (SS-SVR) method. Two probabilistic local reconstruction (PLR) models were used to estimate the label distribution of unlabeled data. Training data was generated by oversampling from the unlabeled data and its estimated label distribution. Results showed that the proposed method improved accuracy by 8% compared to traditional supervised support vector regression (SVR) and reduced training time [150].

6.2. Random Forests/Gradient Boosting

Random forests and gradient boosting are powerful ensemble methods for yield prediction in wafer manufacturing, excelling with noisy, high-dimensional fab data [151].

They capture nonlinear relationships, rank critical features, and provide robust predictions despite measurement variability. These models support root cause analysis, defect classification, and process optimization, driving higher yield and reliability in advanced semiconductor production [152,153].

Lee et al. [117] developed a complete yield prediction system. They compared multiple machine learning models, including gradient boosting, using various types of manufacturing data from wafer fabrication processes to determine the model with the best predictive performance [117]. They also integrated SHAP (Shapley Additive exPlanations) to explain feature importance and illuminate the relationship between yield and features. This approach can analyze a scalable range of manufacturing conditions, understand the influencing factors, and optimize the process [117]. Mohammed's research describes an integrated model using random forests, XGBoost (a form of gradient boosting), LightGBM, and stacking to predict defect density (a key yield driver) [154]. Performance evaluation using metrics such as RMSE shows that the stacking model outperforms traditional single models. This integrated approach can be applied to complex, high-dimensional fab data, improving robustness and prediction accuracy [154]. Kalir et al. [155] applied long short-term memory (LSTM) networks and optimized random forests (RF) to address productivity and cost challenges for predictive maintenance, forecasting equipment behavior, and unplanned downtime. They demonstrated the robustness of RF on noisy fab data and how reliability indirectly supports yield by reducing downtime [155].

6.3. Neural Networks

Neural networks in wafer manufacturing model highly nonlinear relationships between process parameters and outputs, such as yield, defects, and critical dimensions [45]. They learn complex patterns from large fab datasets, enabling virtual metrology, predictive maintenance, and early detection of anomalies [103]. Their adaptability supports advanced nodes, optimizing process control and decision-making in data-rich semiconductor fabrication environments [156].

Choi et al. [157] proposed a multivariate virtual metrology (VM) model based on a convolutional neural network. This model uses process data from multiple sensors to predict various physical metrology results, capturing the complex nonlinear mapping between device signals and outputs within a single deep learning framework. The model was evaluated using VM modeling of the etch process in wafer fabrication as a real-world example. Bao et al. [158] combined autoencoder-based data augmentation techniques with a CNN classifier to process a noisy and unbalanced dataset of wafer defect images. This deep learning approach significantly improved defect pattern classification accuracy, outperforming traditional classifiers such as random forests and support vector machines (SVMs) [158].

Jeong et al. [143] proposed a rotation- and flip-invariant method. This method is based on labeling rules and is insensitive to the rotation and flipping of labels. Therefore, it could improve class discrimination performance even without data by utilizing a convolutional neural network backbone, combined with radon transforms and kernel flipping, to achieve geometric invariance. Qualitative applications suggest using multi-branch layer-by-layer correlation propagation to explain model decisions. Quantitative applications validate the superiority of this method through ablation studies [143].

6.4. Deep Learning (CNNs, LSTMs)

Deep learning models, including CNNs and LSTMs, enable the recognition of wafer map patterns and temporal prediction in semiconductor manufacturing [151]. CNNs excel at identifying spatial defect patterns, while LSTMs capture time-dependent process

trends [139]. Together, they support early excursion detection, root cause analysis, and predictive control, enhancing yield, equipment reliability, and process optimization in advanced wafer fabrication [159].

Hybrid approaches often combine statistical regression + machine learning to balance interpretability and predictive power [139].

Tsai et al. [160] designed a compact convolutional neural network for wafer image classification, using depth-wise separable convolutions. They also utilized encoder–decoder enhancement techniques to achieve fast and accurate spatial defect recognition. The results showed that lightweight deep convolutions can reduce model parameters and computational complexity, making them highly effective for high-volume test fabs [160]. Yoon et al. [161] proposed a semi-automated wafer image pattern classification method that selectively utilizes a convolutional neural network to classify new wafer images. The convolutional neural network classifier incorporates uncertainty perception and engineer-in-the-loop (EIL) decision logic, improving the practicality and interpretability of wafer fab deployments [161].

Kim et al. [162] employed a long short-term memory (LSTM) network in conjunction with a feedforward neural network (FFNN) for temporal modeling. The LSTM utilizes time-series data, whereas the FFNN employs non-time-series data. They predicted edge yield trends based on fab time series, enabling early deviation detection and predictive control [162].

A summary of the integration of machine learning in wafer manufacturing is listed in Table 5.

Table 5. Integration of machine learning in wafer manufacturing.

Model/ Approach	Role and Application	Methods/Examples	References
Support Vector Regression (SVR)	This method captures nonlinear CD variation and maps process variables into higher-dimensional spaces. This supports VM, process optimization, and reduced inspection costs.	SVR for CD prediction [3,20,25]; Dailey et al. [149] used SVD-based OES + SVR for plasma etch VM; Kang et al. [150] proposed SS-SVR with PLR models, improving accuracy (+8%) and reducing training time.	[3,20,25,149,150]
Random Forests/Gradient Boosting	Handle noisy, high-dimensional fab data. Capture nonlinearities, rank features, and support yield prediction, defect classification, and root cause analysis.	Shih et al. [141] classified CMOS wafer defects with RF (~97% accuracy); Choi et al. [142] stacked models with XGBoost for wafer defect patterns; Mohammed et al. [154] integrated RF, XGBoost, LightGBM, and stacking for defect density; Kalir et al. [155] applied RF + LSTM for predictive maintenance; Lee et al. [158] used gradient boosting + SHAP for feature interpretation.	[138,139,141,142,151–155]
Neural Networks (NNs)	Model nonlinear input–output relations for yield, defects, and CD. Enable VM, predictive maintenance, anomaly detection, and decision optimization.	Choi et al. [157]: CNN for multivariate VM; Bao et al. [158]: CNN + autoencoder for noisy defect data; Jeong et al. [143]: CNN with radon transform and kernel flipping for geometric invariance.	[45,103,143,156–158]
Deep Learning (CNNs, LSTMs)	CNNs classify wafer map defect patterns; LSTMs capture time-series fab data for tool drift, yield forecasting, and predictive control.	Tsai et al. [160]: compact CNN for defect recognition; Yoon et al. [161]: CNN + EIL for wafer images; Kim et al. [159]: CNN with inception + skip connection for fast defect classification; Kim et al. [162]: LSTM + FFNN for edge yield trends.	[139,140,143,151,159–162]
Hybrid/Ensemble ML	Combine regression + ML for interpretable yet robust predictions. Effective on limited, imbalanced, or noisy data.	Choi et al. [142] used a stacked ensemble with boosting; Mohammed et al. [154] integrated RF/XGBoost/LightGBM; hybrid regression + DL models balance interpretability and accuracy [139].	[139,142,154]

7. Challenges of Regression and Predictive Modeling

7.1. Regression

7.1.1. Multicollinearity

Tool parameters are often strongly correlated, inflating the variance of estimates in the regression application of wafer manufacturing [47].

Chang et al. [163] noted that predictive modeling in semiconductor processes often faces multicollinearity due to causal relationships among numerous variables. To address this, they proposed two algorithmic variable elimination methods and compared them with principal component and ridge regression. Their study, focused on high-dimensional EUV lithography models, demonstrated improved handling of multicollinearity [163]. Magklaras et al. [164] proposed a ridge regression method using a machine learning (ML) algorithm to adapt to datasets with high multicollinearity. They applied this method to perturbed data from a 300 mm wafer fab. Kim et al. [12] developed an outlier detection technique using multivariate partial least squares regression with dimensionality reduction and function-based independent component analysis. The method captures OES data continuity, extracts structural components, and enhances prediction accuracy for wafer performance during plasma etching, demonstrating strong applicability to optical emission spectroscopy-based process monitoring [12].

7.1.2. Nonlinear and Dynamic Processes

Many interactions are highly nonlinear and time-dependent in the regression application of wafer manufacturing [165]. Advanced models (deep learning, Gaussian processes) outperform linear regression [53].

Wu et al. [166] proposed a hybrid CNN–Gaussian process regression (GPR) model for virtual metrology, combining deep learning feature extraction with probabilistic GPR yield prediction, outperforming linear models in accuracy and uncertainty. Chan et al. [167] introduced a just-in-time GPR framework that adapts to time-varying nonlinear process data, enhancing predictive accuracy. Zhou et al. [168] developed a recursive GPR model for dynamic batch processes, enabling online updates and robust handling of process uncertainty and nonstationarity.

7.1.3. High-Dimensional Data

Modern sensors produce thousands of signals, requiring dimension reduction before regression [169]. Thousands of parameters per wafer make overfitting a risk. Hundreds of variables per wafer step need dimensionality reduction (PLS, PCA, autoencoders) [169,170].

Serani and Diez [171] compared linear (PCA) and nonlinear (autoencoder) methods for functional surface and shape optimization, showing how dimensionality reduction improves tractability and mitigates overfitting in high-dimensional engineering models. Ghojogh et al. [172] presented sufficient dimensionality reduction (SDR) for high-dimensional regression, detailing SIR, SAVE, and MAVE techniques to address overfitting and parameter redundancy. Ajana et al. [173] demonstrated that combining dimensionality reduction with penalized regressions, such as LASSO and ridge, enhances predictive performance and reduces overparameterization, improving model generalization in complex semiconductor process data.

7.1.4. Concept Drift

Regression coefficients change as processes age, requiring frequent recalibration.

Lima et al. [174] reviewed regression coefficient drift, proposing ensemble and neural drift detection methods to address nonstationarity, with a focus on the effects of batch size and drift location on model calibration. Igual et al. [175] explored proactive concept drift strategies, introducing adaptive linear filtering to maintain model accuracy during gradual or sudden process changes, ideal for real-time applications. Samuel and Ahmed [176] developed a continuous learning approach that achieves complete retraining accuracy while efficiently adapting models over time, effectively managing concept drift in dynamic engineering systems.

7.1.5. Interpretability vs. Accuracy

Engineers prefer interpretable regression models, but AI/ML often provides higher predictive accuracy [177]. Alangari et al. [178] compared seven generalized additive model (GAM) variants with seven machine learning models using 20 benchmark datasets. After extensive hyperparameter tuning and cross-validation, advanced GAMs—particularly interpretable boosting machines (EBMs)—achieved predictive accuracy comparable to black-box models while maintaining transparency [178]. Muckley et al. [179] contrasted simple linear regression with black-box methods, finding that despite higher interpolation error, the simple model performed within 5% of complex models during extrapolation and even outperformed them in 40% of cases [179]. Nazir et al. [180] proposed a framework to balance interpretability and accuracy using the Simulatable Operation Count (SOC) metric. They demonstrated that interpretability can improve by 97% with only a 2.5% loss in accuracy, offering a practical trade-off between explainability and performance [180].

A summary of the challenges associated with regression in wafer manufacturing is listed in Table 6.

Table 6. Challenges of regression in wafer manufacturing.

Category	Challenge	Description	Representative Studies
7.1 Regression	**7.1.1 Multicollinearity**	Strong correlations among tool/process parameters inflate the variance of estimates; wafer datasets are high-dimensional and interdependent. Methods include variable elimination, PCA, ridge regression, and PLS.	[12,163,164]
	7.1.2 Nonlinear and dynamic processes	Semiconductor processes are nonlinear, time-varying, and nonstationary. Linear regression struggles; advanced models (CNNs, Gaussian processes, recursive updates) improve predictions and uncertainty handling.	[166–168]
	7.1.3 High-dimensional data	Sensors produce thousands of signals per wafer, creating overfitting risks. Requires dimensionality reduction (PCA, PLS, autoencoders, SDR, penalized regression).	[171–173]
	7.1.4 Concept drift	Regression coefficients evolve as tools age, recipes shift, or fab conditions drift. Models must adapt via online learning, adaptive filtering, ensembling, or continuous learning strategies.	[174–176]
	7.1.5 Interpretability vs. accuracy	Engineers prefer transparent models (linear regression, GAMs), but AI/ML often yields higher accuracy—recent approaches balance interpretability (EBMs, SOC metrics) with performance.	[178–180]

7.2. Predictive

7.2.1. Data Quality and Integration

Sensor data is often noisy, missing, or unaligned across tools. Missing, loud, or inconsistent fab data complicates regression accuracy [181].

Teh et al. [182] discussed key sensor data challenges, including error types, detection, quantification, and correction, providing theoretical and practical strategies for engineers. Siang et al. [183] emphasized that reliable industrial modeling relies on substantial data collection and preprocessing rather than algorithmic complexity, offering guidelines for preparing noisy or incomplete time series for soft sensors. Halder et al. [184] combined physics-informed neural networks with energy-based modeling and Fisher fraction regularization to embed physical knowledge for real-time noise reduction in laser powder bed fusion (LPBF) systems [184].

7.2.2. Model Interpretability

Black-box AI models are more complex for engineers to trust in high-stakes manufacturing [185].

Sofianidis et al. [185] emphasized explainable AI (XAI) to enhance operator trust by increasing transparency in industrial AI systems, which are often viewed as opaque "black boxes." Lisboa et al. [186] argued that manufacturing decisions require ethically cautious use of interpretable, inherently transparent models over opaque ones. Kusiak [187] highlighted that while digital manufacturing leverages predictive models for optimization and simulation, challenges remain in balancing interpretability, complexity, and reusability. This underscores the persistent trade-off between model performance and operator trust in adopting advanced machine learning systems.

7.2.3. Tool-to-Tool Variation

Models may not generalize across different equipment sets [110]. Models must adapt to new tools, recipes, and products [188].

Shim et al. [189] observed that recipe changes in semiconductor manufacturing often shift sensor data distributions, degrading existing fault detection and classification (FDC) model performance. To address this, they developed an adaptive FDC framework employing unsupervised adaptation to limit degradation and semi-supervised adaptation using partial inspection results for new recipes. Chien et al. [190] proposed a CNN-based intelligent FDC strategy that analyzes SVID data to detect wafer status, enabling rapid redefinition of monitoring rules and self-learning, which significantly reduces cycle time. Li et al. [191] introduced the causal mechanism transfer network (CMTN), a domain-adaptive model for multivariate time series that captures invariant causal mechanisms across tools and timescales, allowing effective adaptation with limited new data and improving reliability in tool-to-tool model transfer within semiconductor and mechanical systems.

7.2.4. Computational Efficiency

Real-time predictions require fast inference at fab scale in wafer manufacturing [192].

Horr [193] explored reduced-order modeling (ROM) for wire arc additive manufacturing (WAAM), comparing hybrid ML-ROM approaches that enable fast, reliable, real-time predictions for dynamic manufacturing environments. Horr et al. [194] extended this within a digital twin framework, integrating data-driven, physical, and machine learning models for efficient real-time monitoring and control in additive and extrusion processes. Chen et al. [195] proposed a time-series dense encoder (TiDE) surrogate model to accelerate model predictive control in digital twin systems, enabling multi-step predictions per execution and significantly reducing computational latency for real-time fab-scale decision-making.

7.2.5. Imbalanced Data

Defective wafers/dies are rare and require advanced classification methods [54].

Kim et al. [196] developed an oversampling strategy to balance wafer image data by generating synthetic samples, improving rare defect detection [196]. Choi et al. [142] addressed severe class imbalance with a framework combining a convolutional autoencoder-based data augmentation and a stacked ensemble classifier, enhancing robustness when defect samples are limited [142]. Taha [139] reviewed machine learning methods for wafer defect classification, noting that ResNet achieves about 99% accuracy. At the same time, XGBoost balances strong accuracy (~94.8%) with high computational efficiency, making it suitable for practical industrial applications [139]. A summary of the challenges associated with predictive modeling in wafer manufacturing is listed in Table 7.

Table 7. Challenges of predictive modeling in wafer manufacturing.

Category	Challenge	Description	Representative Studies
7.2 Predictive Modeling	**7.2.1 Data quality and integration**	Sensor streams may be noisy, missing, or misaligned across fab tools. Requires preprocessing, error correction, and domain-knowledge integration (PINNs).	[182–184]
	7.2.2 Model interpretability	Black-box AI (DL, ensembles) undermines trust in high-stakes fabs. Explainable AI (XAI) and transparent models improve adoption and ethical decision-making.	[185–187]
	7.2.3 Tool-to-tool variation	Models trained on one tool often fail on others due to recipe/equipment differences. Solutions include adaptive learning, CNN-based FDC, and causal mechanism transfer networks.	[189–191]
	7.2.4 Computational efficiency	Real-time fab-scale inference demands lightweight models. Hybrid ROMs, digital twins, and surrogate models accelerate prediction and control.	[193–195]
	7.2.5 Imbalanced data	Defect samples are rare compared to regular wafers. Synthetic oversampling, CAE-based augmentation, and ensemble classifiers mitigate imbalance.	[139,142,196]

8. Future Trends

8.1. Hybrid Physics–ML Models

Process physics may be combined with regression/AI for more robust predictions in wafer manufacturing. Kasilingam et al. [134] provide an overview of various hybrid modeling strategies in manufacturing, combining physics-based models with machine learning methods. They describe the different components and architectures of hybrid models, providing a taxonomy and practical insights. This engineering-oriented framework is used to incorporate physics-based knowledge into predictive models, with potential applications in wafer processing [134]. Yu et al. [29] proposed a physics-based machine learning approach to predict material removal rate (MRR) during chemical–mechanical polishing (CMP). This approach combines a multiscale mechanical physics model, including contact mechanics and abrasive wear, with a data-driven model for predicting pad topology variables. Experimental validation demonstrates that the hybrid model achieves extremely high prediction accuracy. This result illustrates how hybrid models can improve regression accuracy and reliability by embedding physical process understanding [29]. Han et al. [197] reviewed the application of physical information neural networks (PINNs) in semiconductor thin-film deposition. They analyzed how to embed physical laws and governing equations into neural network architectures to improve interpretability, robustness, and data efficiency. They also proposed future research directions for integrating PINNs into semiconductor manufacturing [197].

8.2. Real-Time Predictive Analytics

Integration with fab control systems allows immediate corrective action. Predictive models must operate in a real-time environment [198]. Deenen et al. [21] proposed a network of aggregated models, each representing a work area within a wafer fab. These models, built based on real-time arrival and departure data, dynamically predict key operational performance indicators, such as cycle time and work-in-process (WIP) levels. Initialization and prediction of ongoing indicators using a real-time wafer fab snapshot support real-time proactive scheduling and control decisions [21]. Singgih et al. [199] used an AnyLogic-based digital twin model of Intel's small fab to identify key factors affecting

yield. They then trained multiple machine learning models, including adaptive boosting and random forest models, achieving real-time yield prediction accuracy exceeding 97.8%. This technology can quickly detect production disruptions and guide control measures to maintain a smooth wafer flow [199]. Chien et al. [190] investigated the integration of reduced-order models (ROMs), including data-driven and hybrid models, into digital twin systems. These models can rapidly process real-time sensor data and adjust process parameters, thereby facilitating real-time monitoring, predictive, and corrective control. This provides a scalable approach for real-time predictive control in complex manufacturing environments, including wafer fabs [190].

8.3. Digital Twins

Digital twins are virtual replicas of wafer processes driven by regression + AI predictive modeling, allowing the combination of physics-based wafer simulators with predictive ML regression for proactive process control [200]. Behrendt et al. [70] introduced a data-driven approach for automated simulation model generation (ASMG). This approach builds simulation models based on production data. By using machine learning-based device simulation, it can capture device behavior even when data is noisy or incomplete. This technology connects physical and virtual systems with predictive fidelity through automated models, enabling rapid deployment of fab digital twins [70]. Karkaria et al. [201] proposed a digital twin framework for real-time control. They combined machine learning (LSTM) for temperature prediction with Bayesian optimization (BOTSPO) for dynamic adjustment of process parameters. This approach demonstrates the integration of predictive regression models with physics-guided optimization to achieve proactive control [201]. Liu et al. [202] developed a high-fidelity digital twin model that integrates a physics-based LPBF process melt pool computational model, an efficient surrogate model based on Fourier Neural Operator (FNO) regression, and optimization-driven closed-loop control to reduce defects such as surface roughness and porosity, while also quantifying uncertainty. This concept is highly transferable to wafer manufacturing scenarios [202].

8.4. Explainable AI

Enhancing the interpretability of complex predictive models to gain engineers' trust is crucial. Engineers need to interpret ML-based regression results in wafer fabs [190].

Senoner et al. [203] developed a data-driven decision model that combines nonlinear machine learning with Shapley Additive Explanations (SHAP) to quantify essential features. This model, applied to semiconductor manufacturing, effectively identified process drivers of yield fluctuations and guided corrective actions [203]. Lee et al. [117] proposed a yield prediction framework tailored for semiconductor manufacturing, emphasizing model interpretability and scalability through explainable artificial intelligence (XAI). By transparently analyzing the input–output relationship in regression models, they provide a practical modeling framework centered on interpretability for yield prediction in semiconductor processes [117]. Mohammed [154] introduced Trace Shapley Attribution (TSA) to address root cause diagnosis in continuous fab processes. TSA identifies process measurements that are primarily responsible for defect occurrence, addresses inherent temporal dependencies, and avoids the pitfalls of arbitrary baselines, enabling an interpretable technique for regression-driven root cause analysis in the fab [154].

8.5. Edge AI Deployment

Running predictive models directly on equipment controllers enables fast feedback.

INFICON [204] introduced FabGuard®, (INFICON com., Bad Ragaz, Switzerland) an edge AI system deployed at the sensor layer of semiconductor devices. Using the Libtorch library (PyTorch C++ runtime, version 2.2) to run deep learning models in the

field enables real-time endpoint detection in complex etch processes and anomaly detection using multivariate sensor streams.

Chen et al. [205] introduced an architecture for deploying predictive models on edge devices to perform real-time product quality classification. This system, which uses Synthetic Minority Oversampling Technique (SMOTE) + extreme gradient boosting (XGBoost) (version 0.4-2) to handle imbalanced data, has been validated on a real-world industrial brake disc production line. This technology demonstrates the feasibility of edge-based machine learning inference in high-throughput manufacturing, a concept that can be applied to wafer fabs [205].

SemiEngineering [206] explores practical edge AI applications in equipment control systems used by companies like Applied Materials. For example, AI-driven image processing models are embedded in CMP equipment for endpoint and uniformity control, enabling proactive in-tool control rather than post-process analysis. This can effectively enable AI inference within controllers for semiconductor fabs [206].

A summary of the future trends in regression and predictive modeling for wafer manufacturing is listed in Table 8.

Table 8. Future trends in regression and predictive modeling for wafer manufacturing.

Category	Future Trend	Description	Representative Studies
8.1 Hybrid physics–ML models	Integration of physical models with AI/ML	Embeds domain physics into machine learning to enhance robustness, interpretability, and accuracy in wafer processes (e.g., CMP, deposition). Combines governing equations, contact mechanics, and data-driven regression.	[29,133,197]
8.2 Real-time predictive analytics	Models embedded in fab control systems	Predictive models must run in real time for corrective actions. Digital twins and reduced-order models enable cycle-time prediction, yield optimization, and proactive scheduling.	[21,190,199]
8.3 Digital twins	Virtual replicas of wafer fab processes	Combines physics-based simulators and ML regression models to replicate process behavior, optimize parameters, and reduce variability. Supports proactive process monitoring and control.	[70,201,202]
8.4 Explainable AI (XAI)	Improving interpretability of predictive models	Use of SHAP, Trace Shapley Attribution, and interpretable frameworks to balance accuracy and transparency. This tool helps engineers identify key process drivers, root causes, and yield variations.	[117,154,203]
8.5 Edge AI deployment	AI running on equipment controllers	Pushes predictive inference to the sensor/equipment level for low-latency feedback. This feature enables in-tool control, anomaly detection, and real-time quality monitoring.	[204–206]

9. Conclusions

Regression and predictive modeling have become indispensable pillars of modern semiconductor wafer manufacturing, evolving from basic statistical tools into advanced AI-driven systems that shape the industry's future. Historically, linear regression was used to support the design of experiments (DOEs) and yield correlations. Today, these approaches underpin a broad spectrum of applications, including virtual metrology, yield prediction, fault detection and classification (FDC), and predictive maintenance. Their ability to quantify complex relationships, predict outcomes, reduce metrology costs, and detect anomalies directly impacts yield, cost, cycle time, and equipment reliability.

What makes these methods transformative is not only their predictive power but also their integration into control loops and decision-making frameworks, enabling fabs to move beyond static dashboards toward real-time, adaptive manufacturing intelligence. As Industry 4.0 unfolds, regression and predictive modeling will increasingly converge with hybrid physics–data approaches, transfer learning, explainable AI, and digital twin technologies. The next decade will see fabs relying more on AI-augmented regression to manage the growing complexity of advanced nodes while maintaining interpretability, speed, and governance, ultimately enabling greater autonomy, robustness, and competitiveness across the semiconductor value chain.

The conclusion uniquely underscores the evolution of regression and predictive modeling from simple statistical tools to intelligent, adaptive systems that drive Industry 4.0 manufacturing. It highlights their integration into real-time control loops, enabling autonomous decision-making and process optimization. Furthermore, it envisions a future where hybrid physics–AI models, transfer learning, and digital twins enhance interpretability and robustness, ensuring semiconductor fabs achieve greater yield, reliability, and competitiveness amid growing process complexity and data intensity.

Author Contributions: Conceptualization, H.-Y.C. and C.C.; methodology, H.-Y.C. and C.C.; software, C.C.; formal analysis, H.-Y.C.; investigation, H.-Y.C. and C.C.; data curation, H.-Y.C.; writing—original draft preparation, H.-Y.C. and C.C.; writing—review and editing, H.-Y.C. and C.C.; visualization, C.C.; supervision, C.C.; project administration, C.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data are unavailable because a statement is still pending.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Chen, Q.; Wei, Q.; Xu, H. A multimodal hierarchical learning approach for virtual metrology in semiconductor manufacturing. *J. Manuf. Syst.* **2025**, *80*, 194–205. [CrossRef]
2. Hargrove, M.; Wen, S.; Yim, D.; Ruegger, K.E.; Nanja, P.; Sarkar, S.; Lowe, B.; Vincent, B.; Ervin, J.; Fried, D. Review of virtual wafer process modeling and metrology for advanced technology development. *J. Micro/Nanopatterning Mater. Metrol.* **2023**, *22*, 31209. [CrossRef]
3. Kang, P.; Lee, H.J.; Cho, S.; Kim, D.; Park, J.; Park, C.K.; Doh, S. A virtual metrology system for semiconductor manufacturing. *Expert Syst. Appl.* **2009**, *36*, 12554–12561. [CrossRef]
4. Melhem, M.; Ananou, B.; Ouladsine, M.; Pinaton, J. Regression methods for predicting the product's quality in the semiconductor manufacturing process. *IFAC-PapersOnLine* **2016**, *49*, 83–88. [CrossRef]
5. Morsy, M.; Znid, F.; Farraj, A. A critical review on improving and moving beyond the 2 nm horizon: Future directions and impacts in next-generation integrated circuit technologies. *Mater. Sci. Semicond. Proc.* **2025**, *190*, 109376. [CrossRef]
6. Zeng, D. Statistical Methods for Enhanced Metrology in Semiconductor/Photovoltaic Manufacturing. Ph.D. Thesis, UC Berkeley, Berkeley, CA, USA, 2012.
7. Suthar, K.; Shah, D.; Wang, J.; He, Q.P. Next-generation virtual metrology for semiconductor manufacturing: A feature-based framework. *Comput. Chem. Eng.* **2019**, *127*, 140–149. [CrossRef]
8. Scheibelhofer, P.; Gleispach, D.; Hayderer, G.; Stadlober, E. A methodology for predictive maintenance in semiconductor manufacturing. *Austrian J. Stat.* **2012**, *41*, 161–173. [CrossRef]
9. Zhang, J.; Jiang, Y.; Luo, H.; Yin, S. Prediction of material removal rate in chemical mechanical polishing via residual convolutional neural network. *Comput. Chem. Eng.* **2021**, *107*, 104673. [CrossRef]
10. Flechsig, C.; Lohmer, J.; Lasch, R.; Zettler, B.; Schneider, G.; Eberts, D. Streamlining semiconductor manufacturing of 200 mm and 300 mm wafers: A longitudinal case study on the lot-to-order-matching process. *IEEE Trans. Semicond. Manuf.* **2022**, *35*, 397–404. [CrossRef]

11. Skinner, K.R.; Montgomery, D.C.; Runger, G.C.; Fowler, J.W.; McCarville, D.R.; Rhoads, T.R.; Stanley, J.D. Multivariate statistical methods for modeling and analysis of wafer probe test data. *IEEE Trans. Semicond. Manuf.* **2002**, *15*, 523–530. [CrossRef]
12. Kim, K.; Oh, S.; Bae, K.; Oh, H.S. Prediction of wafer performance: Use of functional outlier detection and regression. *IEEE Access* **2025**, *13*, 3544244. [CrossRef]
13. Purwins, H.; Barak, B.; Nagi, A.; Engel, R.; Höckele, U.; Kyek, A.; Cherla, S.; Lenz, B.; Pfeifer, G.; Weinzierl, K. Regression methods for virtual metrology of layer thickness in chemical vapor deposition. *IEEE/ASME Trans. Mechatron.* **2013**, *19*, 1–8. [CrossRef]
14. Wan, J.; Pampuri, S.; O'Hara, P.G.; Johnston, A.B.; McLoone, S. On regression methods for virtual metrology in semiconductor manufacturing. In Proceedings of the 25th IET Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communications Technologies (ISSC 2014/CIICT 2014), Limerick, Ireland, 26–27 June 2014; IET: Stevenage, UK, 2014; pp. 380–385.
15. Zhai, W.; Han, Q.; Chen, L.; Shi, X. Explainable AutoML (xAutoML) with adaptive modeling for yield enhancement in semiconductor smart manufacturing. In Proceedings of the 2024 2nd International Conference on Artificial Intelligence and Automation Control (AIAC), Guangzhou, China, 20–22 December 2024; pp. 162–171.
16. Maitra, V.; Su, Y.; Shi, J. Virtual metrology in semiconductor manufacturing: Current status and prospects. *Expert Syst. Appl.* **2024**, *249*, 123559. [CrossRef]
17. Dreyfus, P.A.; Psarommatis, F.; May, G.; Kiritsis, D. Virtual metrology as an approach for product quality estimation in Industry 4.0: A systematic review and integrative conceptual framework. *Int. J. Prod. Res.* **2022**, *60*, 742–765. [CrossRef]
18. Nam, W.S.; Kim, S.B. A prediction of wafer yield using product fabrication virtual metrology process parameters in semiconductor manufacturing. *J. Korean Inst. Ind. Eng.* **2015**, *41*, 572–578. [CrossRef]
19. Kumar, M.G.V.; Krishna, V.R.; Chinthamu, N.; Gupta, A.; Venkatakiran, S. IoT sensor data-driven programming approach for predictive maintenance of wafer stick machines. In *Technological Applications for Smart Sensors*; Apple Academic Press: Palm Bay, FL, USA, 2025; pp. 447–464.
20. Lenz, B.; Barak, B. Data mining and support vector regression machine learning in semiconductor manufacturing to improve virtual metrology. In Proceedings of the 2013 46th Hawaii International Conference on System Sciences, Wailea, HI, USA, 7–10 January 2013; pp. 3447–3456.
21. Deenen, P.C.; Middelhuis, J.; Akcay, A.; Adan, I.J. Data-driven aggregate modeling of a semiconductor wafer fab to predict WIP levels and cycle time distributions. *Flex. Serv. Manuf. J.* **2024**, *36*, 567–596. [CrossRef]
22. Rana, N.; Zhang, Y.; Wall, D.; Dirahoui, B. Predictive data analytics and machine learning enabling metrology and process control for advanced node IC fabrication. In Proceedings of the 2015 26th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), Saratoga Springs, NY, USA, 3–6 May 2015; pp. 313–319.
23. Kang, S. Joint modeling of classification and regression for improving faulty wafer detection in semiconductor manufacturing. *J. Intell. Manuf.* **2020**, *31*, 319–326. [CrossRef]
24. Ahmad, J. Using ML to Model and Optimize Chip Geometry for Improved Lithography. Master's Thesis, The University of South Bohemia in České Budějovice, Branišovská, České Budějovice, Czech Republic, 2023.
25. Chien, K.C.; Chang, C.H.; Djurdjanovic, D. Virtual metrology modeling of reactive ion etching based on statistics-based and dynamics-inspired spectral features. *J. Vac. Sci. Technol. B* **2021**, *39*, 064003. [CrossRef]
26. Guo, J.; Geng, M.; Ren, K.; Ni, D.; Gao, D. Optimizing plasma etching: Integrating precise three-dimensional etching simulation and machine learning for multi-objective optimization. *IEEE Access* **2024**, *12*, 127065–127073. [CrossRef]
27. Liu, C.L.; Tseng, C.J.; Hsaio, W.H.; Wu, S.H.; Lu, S.R. Predicting the wafer material removal rate for semiconductor chemical mechanical polishing using a fusion network. *Appl. Sci.* **2022**, *12*, 11478. [CrossRef]
28. Zhao, G.; Wei, Z.; Wang, W.; Feng, D.; Xu, A.; Liu, W.; Song, Z. Review on modeling and application of chemical mechanical polishing. *Nanotechnol. Rev.* **2020**, *9*, 182–189. [CrossRef]
29. Yu, T.; Li, Z.; Wu, D. Predictive modeling of material removal rate in chemical mechanical planarization with physics-informed machine learning. *Wear* **2019**, *426*, 1430–1438. [CrossRef]
30. Dimaculangan, R.V.; de Luna, R.G.; Rosales, M.A.; Magsumbol, J.A.V.; Tubola, O.D. Semiconductor manufacturing final test yield prediction using regression with genetic algorithm-based feature selection. In Proceedings of the International Conference on Intelligent Computing & Optimization, Phnom Penh, Cambodia, 27–28 October 2023; Springer Nature: Cham, Switzerland, 2023; pp. 108–117.
31. Wang, H.; Li, B.; Tong, S.H.; Chang, I.K.; Wang, K. A discrete spatial model for wafer yield prediction. *Qual. Eng.* **2018**, *30*, 169–182. [CrossRef]
32. Ahmadi, A.; Stratigopoulos, H.G.; Huang, K.; Nahar, A.; Orr, B.; Pas, M.; Carulli, J.M., Jr.; Makris, Y. Yield forecasting across semiconductor fabrication plants and design generations. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **2017**, *36*, 2120–2133. [CrossRef]
33. Kumar, N.; Kennedy, K.; Gildersleeve, K.; Abelson, R.; Mastrangelo, C.M.; Montgomery, D.C. A review of yield modelling techniques for semiconductor manufacturing. *Int. J. Prod. Res.* **2006**, *44*, 5019–5036. [CrossRef]

34. Kim, J.S.; Jang, S.J.; Kim, T.W.; Lee, H.J.; Lee, J.B. A productivity-oriented wafer map optimization using yield model based on machine learning. *IEEE Trans. Semicond. Manuf.* **2018**, *32*, 39–47. [CrossRef]
35. Pfitzner, L.; Benesch, N.; Öchsner, R.; Schmidt, C.; Schneider, C.; Tschaftary, T.; Trunk, R.; Dudenhausen, H.M. Cost reduction strategies for wafer expenditure. *Microelectron. Eng.* **2001**, *56*, 61–71. [CrossRef]
36. Kuo, W.; Kim, T. An overview of manufacturing yield and reliability modeling for semiconductor products. *Proc. IEEE* **1999**, *87*, 1329–1344. [CrossRef]
37. Chou, P.H.; Chiang, K.N.; Liang, S.Y. Reliability assessment of wafer level package using artificial neural network regression model. *J. Mech.* **2019**, *35*, 829–837. [CrossRef]
38. Xu, H.W.; Zhang, Q.H.; Sun, Y.N.; Chen, Q.L.; Qin, W.; Lv, Y.L.; Zhang, J. A fast ramp-up framework for wafer yield improvement in semiconductor manufacturing systems. *J. Manuf. Syst.* **2024**, *76*, 222–233. [CrossRef]
39. Tirkel, I. Cycle time prediction in wafer fabrication line by applying data mining methods. In Proceedings of the 2011 IEEE/SEMI Advanced Semiconductor Manufacturing Conference, San Francisco, CA, USA, 16–18 May 2011; pp. 1–5.
40. Dong, H.; Chen, N.; Wang, K. Wafer yield prediction using derived spatial variables. *Qual. Reliab. Eng. Int.* **2017**, *33*, 2327–2342. [CrossRef]
41. PDF Solutions Semiconductor Manufacturing Data 101—An Introduction. Available online: https://www.pdf.com/semiconductor-manufacturing-data-101-an-introduction/?utm_source=chatgpt.com/ (accessed on 31 August 2025).
42. Ma, J.; Zhang, T.; Yang, C.; Cao, Y.; Xie, L.; Tian, H.; Li, X. Review of wafer surface defect detection methods. *Electronics* **2023**, *12*, 1787. [CrossRef]
43. Amato, U.; Antoniadis, A.; De Feis, I.; Doinychko, A.; Gijbels, I.; La Magna, A.; Pagano, D.; Piccinini, F.; Suviseshamuthu, E.S.; Severgnini, C.; et al. Detecting important features and predicting yield from defects detected by SEM in semiconductor production. *Sensors* **2025**, *25*, 4218. [CrossRef]
44. Chen, C.H.; Zhao, W.D.; Pang, T.; Lin, Y.Z. Virtual metrology of semiconductor PVD process based on combination of tree-based ensemble model. *ISA Trans.* **2020**, *103*, 192–202. [CrossRef]
45. Kenney, J.; Valcore, J.; Riggs, S.; Rietman, E. Deep learning regression of VLSI plasma etch metrology. *arXiv* **2019**, arXiv:1910.10067. [CrossRef]
46. Qin, F.; Zhang, L.; Chen, P.; An, T.; Dai, Y.; Gong, Y.; Yi, Z.; Wang, H. In situ wireless measurement of grinding force in silicon wafer self-rotating grinding process. *Mech. Syst. Signal Process.* **2021**, *154*, 107550. [CrossRef]
47. Chien, C.F.; Chen, C.C. Adaptive parametric yield enhancement via collinear multivariate analytics for semiconductor intelligent manufacturing. *Appl. Soft Comput.* **2021**, *108*, 107385. [CrossRef]
48. Pan, J.C.H.; Tai, D.H. Implementing virtual metrology for in-line quality control in semiconductor manufacturing. *Int. J. Syst. Sci.* **2009**, *40*, 461–470. [CrossRef]
49. Krueger, D.C.; Montgomery, D.C.; Mastrangelo, C.M. Application of generalized linear models to predict semiconductor yield using defect metrology data. *IEEE Trans. Semicond. Manuf.* **2010**, *24*, 44–58. [CrossRef]
50. Caglayan, C.; Sertel, K. Experimental analysis of repeatability and calibration residuals in on-wafer non-contact probing. *IEEE Trans. Microw. Theory Tech.* **2017**, *65*, 2185–2191. [CrossRef]
51. Lee, T.; Kim, C.O. Statistical comparison of fault detection models for semiconductor manufacturing processes. *IEEE Trans. Semicond. Manuf.* **2014**, *28*, 80–91. [CrossRef]
52. Chen, Y.J.; Wang, B.C.; Wu, J.Z.; Wu, Y.C.; Chien, C.F. Big data analytic for multivariate fault detection and classification in semiconductor manufacturing. In Proceedings of the 2017 13th IEEE Conference on Automation Science and Engineering (CASE), Xi'an, China, 20–23 August 2017; pp. 731–736.
53. Jiang, D.; Lin, W.; Raghavan, N. A Gaussian mixture model clustering ensemble regressor for semiconductor manufacturing final test yield prediction. *IEEE Access* **2021**, *9*, 22253–22263. [CrossRef]
54. Arnold, N.N.A. Wafer Defect Prediction with Statistical Machine Learning. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2016.
55. Liu, Y. Predictive Modeling for Intelligent Maintenance in Complex Semiconductor Manufacturing Processes. Ph.D. Thesis, University of Michigan, Ann Arbor, MI, USA, 2008.
56. Iskandar, J.; Moyne, J.; Subrahmanyam, K.; Hawkins, P.; Armacost, M. Predictive maintenance in semiconductor manufacturing. In Proceedings of the 2015 26th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), Saratoga Springs, NY, USA, 3–6 May 2015; pp. 384–389.
57. Zhu, L.; Chen, J.; Chen, C.I. Prognostics of tool failing behavior based on autoassociative Gaussian process regression for semiconductor manufacturing. In Proceedings of the 2020 IEEE International Conference on Industrial Technology (ICIT), Buenos Aires, Argentina, 26–28 February 2020; pp. 316–321.
58. Zhu, L.; Chen, J.; Chen, C.I. Prognostics for semiconductor sustainability: Tool failure behavior prediction in fabrication processes. *IEEE Trans. Syst. Man Cybern. Syst.* **2024**, *54*, 3450–3460. [CrossRef]

59. Kim, T.; Behdinan, K. Advances in machine learning and deep learning applications towards wafer map defect recognition and classification: A review. *J. Intell. Manuf.* **2023**, *34*, 3215–3247. [CrossRef]
60. Munirathinam, S.; Ramadoss, B. Predictive models for equipment fault detection in the semiconductor manufacturing process. *IACSIT Int. J. Eng. Technol. Innov.* **2016**, *8*, 273–285.
61. Del Castillo, E.; Hurwitz, A.M. Run-to-run process control: Literature review and extensions. *J. Qual. Technol.* **1997**, *29*, 184–196. [CrossRef]
62. Chemali, C.E.; Freudenberg, J.; Hankinson, M.; Bendik, J.J. Run-to-run critical dimension and sidewall angle lithography control using the PROLITH simulator. *IEEE Trans. Semicond. Manuf.* **2004**, *17*, 388–401. [CrossRef]
63. Wan, J.; McLoone, S. Gaussian process regression for virtual metrology-enabled run-to-run control in semiconductor manufacturing. *IEEE Trans. Semicond. Manuf.* **2017**, *31*, 12–21. [CrossRef]
64. Martin, A.; Vollertsen, R.P. An introduction to fast wafer level reliability monitoring for integrated circuit mass production. *Microelectron. Reliab.* **2007**, *44*, 1209–1231. [CrossRef]
65. Bensoussan, A. Microelectronic reliability models for more than Moore nanotechnology products. *Facta Univ. Ser. Electron. Energ.* **2016**, *30*, 1–25. [CrossRef]
66. Su, Q.H.; Chiang, K.N. Predicting wafer-level package reliability life using mixed supervised and unsupervised machine learning algorithms. *Materials* **2022**, *15*, 3897. [CrossRef] [PubMed]
67. Kim, K.O.; Oh, H.S. Reliability functions estimated from commonly used yield models. *Microelectron. Reliab.* **2008**, *48*, 481–489. [CrossRef]
68. Amrouch, H. Techniques for Aging, Soft Errors and Temperature to Increase the Reliability of Embedded On-Chip Systems. Ph.D. Thesis, Karlsruher Institut für Technologie (KIT), Karlsruhe, Germany, 2015.
69. Lenhard, P.; Kovalenko, A.; Lenhard, R. Die level predictive modeling to reduce latent reliability defect escapes. *Microelectron. Reliab.* **2023**, *148*, 115139. [CrossRef]
70. Behrendt, S.; Altenmüller, T.; May, M.C.; Kuhnle, A.; Lanza, G. Real-to-sim: Automatic simulation model generation for a digital twin in semiconductor manufacturing. *J. Intell. Manuf.* **2025**, 1–20. [CrossRef]
71. Cale, T.S.; Merchant, T.P.; Borucki, L.J.; Labun, A.H. Topography simulation for the virtual wafer fab. *Thin Solid Film.* **2000**, *365*, 152–175. [CrossRef]
72. Sheng, C.; Dong, X.; Zhu, Y.; Wang, X.; Chen, X.; Xia, Y.; Xu, Z.; Zhou, P.; Wang, J.; Bao, W. Two-dimensional semiconductors: From device processing to circuit integration. *Adv. Funct. Mater.* **2023**, *33*, 2304778. [CrossRef]
73. Nguyen, P.H.; Bär, E.; Lorenz, J.; Ryssel, H. Modeling of chemical–mechanical polishing on patterned wafers as part of integrated topography process simulation. *Microelectron. Eng.* **2004**, *76*, 89–94. [CrossRef]
74. Graves, D.B.; Labelle, C.B.; Kushner, M.J.; Aydil, E.S.; Donnelly, V.M.; Chang, J.P.; Mayer, P.; Overzet, L.; Shannon, S.; Rauf, S.; et al. Science challenges and research opportunities for plasma applications in microelectronics. *J. Vac. Sci. Technol. B* **2024**, *42*, 042202. [CrossRef]
75. Yao, A. Design-Technology Co-Optimization for Sub-2 nm Technology Node Based on 2D Materials. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2025.
76. Sturtevant, J.; Capodieci, L. Design for manufacturing and design process technology co-optimization. In *Microlithography*; CRC Press: Boca Raton, FL, USA, 2020; pp. 293–326.
77. Zhu, Y.; Tang, H.; Pan, Y.; Ouyang, P.; Ma, Y.; Xu, K. Meta-optimized LDMOS process–device co-optimization: Extrapolation-enhanced modeling with wafer-level validation. *IEEE Trans. Electron Devices* **2025**, *72*, 5089–5096. [CrossRef]
78. Chen, S.N.; Chen, S.H. Application of generative adversarial networks for virtual silicon data generation and design-technology co-optimization: A study on WAT and CP. *IEEE Access* **2024**, *12*, 6532–6545. [CrossRef]
79. Purwins, H.; Nagi, A.; Barak, B.; Höckele, U.; Kyek, A.; Lenz, B.; Pfeifer, G.; Weinzierl, K. Regression methods for prediction of PECVD Silicon Nitride layer thickness. In Proceedings of the 2011 IEEE International Conference on Automation Science and Engineering, Trieste, Italy, 24–27 August 2011; pp. 387–392.
80. Zhang, Q.; Poolla, K.; Spanos, C.J. Across wafer critical dimension uniformity enhancement through lithography and etch process sequence: Concept, approach, modeling, and experiment. *IEEE Trans. Semicond. Manuf.* **2007**, *20*, 488–505. [CrossRef]
81. Forsberg, M. Effect of process parameters on material removal rate in chemical mechanical polishing of Si (100). *Microelectron. Eng.* **2005**, *77*, 319–326. [CrossRef]
82. Mahandran, C.J.; Fatah, A.Y.A.; Bani, N.A.; Kaidi, H.M.; Muhtazaruddin, M.B.; Amran, M.E. Thermal oxidation improvement in semiconductor wafer fabrication. *Int. J. Power Electron. Drive Syst.* **2019**, *10*, 1141–1147. [CrossRef]
83. Shumate, D.A.; Montgomery, D.C. Development of a TiW plasma etch process using a mixture experiment and response surface optimization. *IEEE Trans. Semicond. Manuf.* **1996**, *9*, 335–343. [CrossRef]
84. Liu, X.; Liu, Y.; Liang, Y.; Liu, H.; Hu, Y.; Gao, B. Optimization of slurry components for a copper chemical mechanical polishing at low down pressure using response surface methodology. *Microelectron. Eng.* **2011**, *88*, 99–104. [CrossRef]

85. Shi, H.; Fang, Z.; Yang, D. Mechanism study of ion implantation on photoresist shrinkage. In *Advances in Patterning Materials and Processes XXXV*; SPIE: Bellingham, WA, USA, 2018; pp. 189–197.
86. Jankovič, D.; Pipan, M.; Šimic, M.; Herakovič, N. Polynomial regression-based predictive expert system for enhancing hydraulic press performance over a 5G network. *Appl. Sci.* **2024**, *14*, 12016. [CrossRef]
87. Mack, C.A.; Byers, J.D. Improved model for focus-exposure data analysis. In *Metrology, Inspection, and Process Control for Microlithography XVII*; SPIE: Bellingham, WA, USA, 2003; pp. 396–405.
88. George, P.; Meng, J.; Jaluria, Y. Optimization of gallium nitride metalorganic chemical vapor deposition process. *J. Heat Transf.* **2015**, *137*, 061007. [CrossRef]
89. Bae, S.J.; Hwang, J.Y.; Kuo, W. Yield prediction via spatial modeling of clustered defect counts across a wafer map. *IIE Trans.* **2007**, *39*, 1073–1083. [CrossRef]
90. Jizat, J.A.M.; Majeed, A.P.A.; Nasir, A.F.A.; Taha, Z.; Yuen, E. Evaluation of the machine learning classifier in wafer defects classification. *ICT Express* **2021**, *7*, 535–539. [CrossRef]
91. Lynn, S. Virtual Metrology for Plasma Etch Processes. Ph.D. Thesis, National University of Ireland, Maynooth, Ireland, 2011.
92. Gonzalez Debs, M.; Kuech, T.F. Phenomenological modeling of diffusion profiles: Sn in GaAs. *J. Appl. Phys.* **2006**, *99*, 123710. [CrossRef]
93. Jia, X.; Di, Y.; Feng, J.; Yang, Q.; Dai, H.; Lee, J. Adaptive virtual metrology for semiconductor chemical mechanical planarization process using GMDH-type polynomial neural networks. *J. Process Control* **2018**, *62*, 44–54. [CrossRef]
94. Nami, Z.; Misman, O.; Erbil, A.; May, G.S. Semi-empirical neural network modeling of metal-organic chemical vapor deposition. *IEEE Trans. Semicond. Manuf.* **1997**, *10*, 288–294. [CrossRef]
95. Huang, Y.; Rosakis, A.J. Extension of Stoney's formula to non-uniform temperature distributions in thin film/substrate systems. The case of radial symmetry. *J. Mech. Phys. Solids* **2005**, *53*, 2483–2500. [CrossRef]
96. Iwakoshi, T.; Hirota, K.; Mori, M.; Tanaka, J.I.; Itabashi, N. Prediction of etching results and etching stabilization by applying principal component regression to emission spectra during in-situ cleaning. *Thin Solid Film.* **2008**, *516*, 3464–3468. [CrossRef]
97. Phatak, U.; Bukkapatnam, S.; Kong, Z.; Komanduri, R. Sensor-based modeling of slurry chemistry effects on the material removal rate (MRR) in copper-CMP process. *Int. J. Mach. Tools Manuf.* **2009**, *49*, 171–181. [CrossRef]
98. Jeong, Y.S.; Hwang, S.; Ko, Y.D. Quantitative analysis for plasma etch modeling using optical emission spectroscopy: Prediction of plasma etch responses. *Ind. Eng. Manag. Syst.* **2015**, *14*, 392–400. [CrossRef]
99. Hong, S.J.; May, G.S.; Park, D.C. Neural network modeling of reactive ion etching using optical emission spectroscopy data. *IEEE Trans. Semicond. Manuf.* **2003**, *16*, 598–608. [CrossRef]
100. Pan, C.C.; Bai, J.; Yang, G.K.; Wong, D.S.H.; Jang, S.S. An inferential modeling method using enumerative PLS based nonnegative garrote regression. *J. Process Control* **2012**, *22*, 1637–1646. [CrossRef]
101. Khan, A.A.; Moyne, J.R.; Tilbury, D.M. Virtual metrology and feedback control for semiconductor manufacturing processes using recursive partial least squares. *J. Process Control* **2008**, *18*, 961–974. [CrossRef]
102. Melhem, M.; Ananou, B.; Ouladsine, M.; Pinaton, J. Regularized regression models to predict the product quality in multistep manufacturing. In Proceedings of the 2016 5th International Conference on Systems and Control (ICSC), Marrakesh, Morocco, 25–27 May 2016; pp. 31–36.
103. Durowoju, E.S.; Olowonigba, J.K. Machine learning-driven process optimization in semiconductor manufacturing: A new framework for yield enhancement and defect reduction. *Int. J. Adv. Res. Publ. Rev.* **2024**, *1*, 110–130. [CrossRef]
104. Park, C.; Kim, S.B. Virtual metrology modeling of time-dependent spectroscopic signals by a fused lasso algorithm. *J. Process Control* **2016**, *42*, 51–58. [CrossRef]
105. Choi, J.; Son, Y.; Kang, J. Group-exclusive feature group Lasso and applications to automatic sensor selection for virtual metrology in semiconductor manufacturing. *IEEE Trans. Semicond. Manuf.* **2024**, *37*, 505–517. [CrossRef]
106. Chen, H.; Leclair, J. Optimizing etching process recipe based on kernel ridge regression. *J. Manuf. Process* **2021**, *61*, 454–460. [CrossRef]
107. Lee, G.; Wang, T.; Kim, D.; Jeong, M.K. Sparse group principal component analysis using elastic-net regularisation and its application to virtual metrology in semiconductor manufacturing. *Int. J. Prod. Res.* **2025**, *63*, 865–881. [CrossRef]
108. Bao, L.; Wang, K.; Jin, R. A hierarchical model for characterising spatial wafer variations. *Int. J. Prod. Res.* **2014**, *52*, 1827–1842. [CrossRef]
109. Krueger, D.C.; Montgomery, D.C. Modeling and analyzing semiconductor yield with generalized linear mixed models. *Appl. Stoch. Models Bus. Ind.* **2014**, *30*, 691–707. [CrossRef]
110. Yuan, T.; Ramadan, S.Z.; Bae, S.J. Yield prediction for integrated circuits manufacturing through hierarchical Bayesian modeling of spatial defects. *IEEE Trans. Reliab.* **2011**, *60*, 729–741. [CrossRef]
111. Liu, J.P.; Jin, R.; Kong, Z.J. Wafer quality monitoring using spatial Dirichlet process based mixed-effect profile modeling scheme. *J. Manuf. Syst.* **2018**, *48*, 21–32. [CrossRef]

112. Miyaguchi, K.; Joko, M.; Sheraw, R.; Idé, T. Wafer defect root cause analysis with partial trajectory regression DM: Big data management and machine learning. In Proceedings of the 36th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), Saratoga Springs, NY, USA, 5–8 May 2025; pp. 1–6.
113. Orji, N.G.; Obeng, Y.S.; Beitia, C.; Mashiro, S.; Moyne, J. Virtual metrology white paper—International roadmap for devices and systems (IRDS). In *International Roadmap for Devices and Systems (IRDS)*; IEEE: Piscataway, NJ, USA, 2018.
114. Kim, D.; Kang, S. Effect of irrelevant variables on faulty wafer detection in semiconductor manufacturing. *Energies* **2019**, *12*, 2530. [CrossRef]
115. Ishikawa, Y.; Otsuka, M.; Nagai, T.; Utsumi, T.; Takeda, M. Recursive partial least squares for process drift com pensation in plasma etch virtual metrology. *IEEE Trans. Semicond. Manuf.* **2010**, *23*, 310–319.
116. Lee, H.; Kim, T.; Choi, B. Locally weighted PLS modeling for adaptive virtual metrology in plasma etching pro cesses. *J. Process Control* **2017**, *58*, 49–60.
117. Lee, Y.; Roh, Y. An expandable yield prediction framework using explainable artificial intelligence for semiconductor manufacturing. *Appl. Sci.* **2023**, *13*, 2660. [CrossRef]
118. Puggini, L.; McLoone, S. An enhanced variable selection and isolation forest-based methodology for anomaly detection with OES data. *Eng. Appl. Artif. Intell.* **2018**, *67*, 126–135. [CrossRef]
119. Kazemi, P.; Masoumian, A.; Martin, P. Fault detection and isolation for time-varying processes using neural-based principal component analysis. *Processes* **2024**, *12*, 1218. [CrossRef]
120. Lynn, S.A.; Ringwood, J.; MacGearailt, N. Global and local virtual metrology models for a plasma etch process. *IEEE Trans. Semicond. Manuf.* **2011**, *25*, 94–103. [CrossRef]
121. Pate, K.; Safier, P. Chemical metrology methods for CMP quality. In *Advances in Chemical Mechanical Planarization (CMP)*; Woodhead Publishing: Cambridge, UK, 2022; pp. 355–383.
122. Deivendran, B.; Masampally, V.S.; Nadimpalli, N.R.V.; Runkana, V. Virtual metrology for chemical mechanical planarization of semiconductor wafers. *J. Intell. Manuf.* **2025**, *36*, 1923–1942. [CrossRef]
123. Jiang, D.; Lin, W.; Raghavan, N. Semiconductor manufacturing final test yield optimization and wafer acceptance test parameter inverse design using multi-objective optimization algorithms. *IEEE Access* **2021**, *9*, 137655–137666. [CrossRef]
124. Miller, R.B.; Riordan, W.C. Unit level predicted yield: A method of identifying high defect density die at wafer sort. In Proceedings of the International Test Conference 2001 (Cat. No. 01CH37260), Baltimore, MD, USA, 1 November 2001; pp. 1118–1127.
125. Leite, D.; Andrade, E.; Rativa, D.; Maciel, A.M. Fault detection and diagnosis in industry 4.0: A review on challenges and opportunities. *Sensors* **2024**, *25*, 60. [CrossRef]
126. Park, H.; Choi, J.E.; Kim, D.; Hong, S.J. Artificial immune system for fault detection and classification of semiconductor equipment. *Electronics* **2021**, *10*, 944. [CrossRef]
127. Goodlin, B.E.; Boning, D.S.; Sawin, H.H.; Wise, B.M. Simultaneous fault detection and classification for semiconductor manufacturing tools. *J. Electrochem. Soc.* **2003**, *150*, G778–G784. [CrossRef]
128. Kim, D.; Kang, P.; Cho, S.; Lee, H.J.; Doh, S. Machine learning-based novelty detection for faulty wafer detection in semiconductor manufacturing. *Expert Syst. Appl.* **2012**, *39*, 4075–4083. [CrossRef]
129. Fan, S.K.S.; Hsu, C.Y.; Tsai, D.M.; He, F.; Cheng, C.C. Data-driven approach for fault detection and diagnostic in semiconductor manufacturing. *IEEE Trans. Autom. Sci. Eng.* **2020**, *17*, 1925–1936. [CrossRef]
130. Ren, J.; Ni, D. A real-time monitoring framework for wafer fabrication processes with run-to-run variations. *IEEE Trans. Semicond. Manuf.* **2021**, *34*, 483–492. [CrossRef]
131. Chen, J.H.; Kuo, T.W.; Lee, A.C. Run-by-run process control of metal sputter deposition: Combining time series and extended Kalman filter. *IEEE Trans. Semicond. Manuf.* **2007**, *20*, 278–285. [CrossRef]
132. Chen, A.; Blue, J. Recipe-independent indicator for tool health diagnosis and predictive maintenance. *IEEE Trans. Semicond. Manuf.* **2009**, *22*, 522–535. [CrossRef]
133. Blue, J.; Gleispach, D.; Roussy, A.; Scheibelhofer, P. Tool condition diagnosis with a recipe-independent hierarchical monitoring scheme. *IEEE Trans. Semicond. Manuf.* **2012**, *26*, 82–91. [CrossRef]
134. Kasilingam, S.; Yang, R.; Singh, S.K.; Farahani, M.A.; Rai, R.; Wuest, T. Physics-based and data-driven hybrid modeling in manufacturing: A review. *Prod. Manuf. Res.* **2024**, *12*, 2305358. [CrossRef]
135. Wang, J.; Li, Y.; Gao, R.X.; Zhang, F. Hybrid physics-based and data-driven models for smart manufacturing: Modelling, simulation, and explainability. *J. Manuf. Syst.* **2022**, *63*, 381–391. [CrossRef]
136. Liu, Y.J.; Ni, D.; Shao, X.; Gong, D.L.; Li, J.J. A hierarchical model-based method for wafer level virtual metrology under process information deficiency. *Qual. Eng.* **2024**, *36*, 534–547. [CrossRef]
137. Jeong, M.K.; Choi, J.; Son, Y.; Kang, J. Deep learning based virtual metrology and yield prediction in semiconductor manufacturing processes. In Proceedings of the PHM Society Asia-Pacific Conference, Jeju, Republic of Korea, 14 July 2017; Volume 1.
138. Baly, R.; Hajj, H. Wafer classification using support vector machines. *IEEE Trans. Semicond. Manuf.* **2012**, *25*, 373–383. [CrossRef]

139. Taha, K. Observational and experimental insights into machine learning-based defect classification in wafers. *J. Intell. Manuf.* **2025**, 1–51. [CrossRef]
140. Huang, A.C.; Meng, S.H.; Huang, T.J. A survey on machine and deep learning in semiconductor industry: Methods, opportunities, and challenges. *Clust. Comput.* **2023**, *26*, 3437–3472. [CrossRef]
141. Shih, D.H.; Yang, C.Y.; Wu, T.W.; Shih, M.H. Investigating a machine learning approach to predicting white pixel defects in wafers—A case study of wafer fabrication plant F. *Sensors* **2024**, *24*, 3144. [CrossRef]
142. Choi, J.; Suh, D.; Otto, M.O. Boosted stacking ensemble machine learning method for wafer map pattern classification. *Comput. Mater. Contin.* **2023**, *74*, 2945–2966. [CrossRef]
143. Jeong, I.; Lee, S.Y.; Park, K.; Kim, I.; Huh, H.; Lee, S. Wafer map failure pattern classification using geometric transformation-invariant convolutional neural network. *Sci. Rep.* **2023**, *13*, 8127. [CrossRef]
144. Rahman, M.W.; Vogl, G.W.; Jia, X.; Qu, Y. Physics-informed multi-task learning for material removal rate prediction in semiconductor chemical mechanical planarization. In Proceedings of the 2024 IEEE International Conference on Prognostics and Health Management (ICPHM), Denver, CO, USA, 10–13 June 2024; pp. 385–392.
145. Yang, Y.; Liu, K.; Gao, Y.; Wang, C.; Cao, L. Advancements and challenges in inverse lithography technology: A review of artificial intelligence-based approaches. *Light Sci. Appl.* **2025**, *14*, 250. [CrossRef]
146. Zhao, X.; Zheng, H.; Zhao, Z.; Cheng, M.; Li, W.; Wan, G.; Jia, Y. Warpage prediction of fan-out wafer-level package based on coupled deep learning and finite element simulation. *Microelectron. Reliab.* **2025**, *170*, 115759. [CrossRef]
147. Kota, P.R. Real-Time Plasma Etch Rate Optimization Using Machine Learning-Based Neural Network Surrogate Models. 2025, *Authorea Preprints*.
148. Kwon, J.W.; Ryu, S.; Park, J.; Lee, H.; Jang, Y.; Park, S.; Kim, G.H. Development of virtual metrology using plasma information variables to predict Si etch profile processed by $SF_6/O_2/Ar$ capacitively coupled plasma. *Materials* **2021**, *14*, 3005. [CrossRef]
149. Dailey, R.; Bertelson, S.; Kim, J.; Djurdjanovic, D. Virtual metrology of critical dimensions in plasma etch processes using entire optical emission spectrum. *IEEE Trans. Semicond. Manuf.* **2024**, *37*, 363–372. [CrossRef]
150. Kang, P.; Kim, D.; Cho, S. Semi-supervised support vector regression based on self-training with label uncertainty: An application to virtual metrology in semiconductor manufacturing. *Expert Syst. Appl.* **2016**, *51*, 85–106. [CrossRef]
151. Mamun, K.M.; Pala, N.; Shawkat, M.S.A. A comprehensive review of machine learning approaches for semiconductor device modeling and simulation. *IEEE Access* **2025**, *13*, 3605856. [CrossRef]
152. Katari, M.; Shanmugam, L.; Malaiyappan, J.N.A. Integration of AI and machine learning in semiconductor manufacturing for defect detection and yield improvement. *Int. J. Artif. Intell. Gen. Sci.* **2024**, *3*, 418–431. [CrossRef]
153. Soorya, K.; Michael, G. Prediction of yields in semiconductor using XGBoost classifier in comparison with random forest classifier. In *Information Visualization for Intelligent Systems*; Wiley Online Library: New York, NY, USA, 2025; Chapter 9; pp. 95–104.
154. Mohammed, H.U.H. A data-driven framework for predicting defect density in semiconductor wafer fabrication using ensemble learning. *Int. J. Intell. Syst. Appl. Eng.* **2021**, *9*, 429–437.
155. Kalir, A.A.; Lo, S.K.; Goldberg, G.; Zingerman-Koladko, I.; Ohana, A.; Revah, Y.; Chimol, T.B.; Honig, G. Leveraging machine learning for capacity and cost on a complex toolset: A case study. *IEEE Trans. Semicond. Manuf.* **2023**, *36*, 611–618. [CrossRef]
156. Puthanveettil Madathil, A.; Luo, X.; Liu, Q.; Walker, C.; Madarkar, R.; Qin, Y. A review of explainable artificial intelligence in smart manufacturing. *Int. J. Prod. Res.* **2025**, 1–44. [CrossRef]
157. Choi, J.; Zhu, M.; Kang, J.; Jeong, M.K. Convolutional neural network based multi-input multi-output model for multi-sensor multivariate virtual metrology in semiconductor manufacturing. *Ann. Oper. Res.* **2024**, *339*, 185–201. [CrossRef]
158. Bao, Y.Y.; Li, E.C.; Yang, H.Q.; Jia, B.B. Wafer map defect classification using autoencoder-based data augmentation and convolutional neural network. *arXiv* **2024**, arXiv:2411.11029. [CrossRef]
159. Kim, M.; Tak, J.; Shin, J. A deep learning model for wafer defect map classification: Perspective on classification performance and computational volume. *Phys. Stat. Sol.* **2024**, *261*, 2300113. [CrossRef]
160. Tsai, T.H.; Lee, Y.C. A light-weight neural network for wafer map classification based on data augmentation. *IEEE Trans. Semicond. Manuf.* **2020**, *33*, 663–672. [CrossRef]
161. Yoon, S.; Kang, S. Semi-automatic wafer map pattern classification with convolutional neural networks. *Comput. Ind. Eng.* **2022**, *166*, 107977. [CrossRef]
162. Kim, D.; Kim, M.; Kim, W. Wafer edge yield prediction using a combined long short-term memory and feed-forward neural network model for semiconductor manufacturing. *IEEE Access* **2020**, *8*, 215125–215132. [CrossRef]
163. Chang, Y.C.; Mastrangelo, C. Addressing multicollinearity in semiconductor manufacturing. *Qual. Reliab. Eng. Int.* **2011**, *27*, 843–854. [CrossRef]
164. Magklaras, A.; Gogos, C.; Alefragis, P.; Birbas, A. Enhancing parameters tuning of overlay models with ridge regression: Addressing multicollinearity in high-dimensional data. *Mathematics* **2024**, *12*, 3179. [CrossRef]
165. Li, Y.; Du, J.; Jiang, W. Reinforcement learning for process control with application in semiconductor manufacturing. *IISE Trans.* **2024**, *56*, 585–599. [CrossRef]

166. Wu, X.; Chen, J.; Xie, L.; Chan, L.L.T.; Chen, C.I. Development of convolutional neural network based Gaussian process regression to construct a novel probabilistic virtual metrology in multi-stage semiconductor processes. *Control Eng. Pract.* **2020**, *96*, 104262. [CrossRef]
167. Chan, L.L.T.; Wu, X.; Chen, J.; Xie, L.; Chen, C.I. Just-in-time modeling with variable shrinkage based on Gaussian processes for semiconductor manufacturing. *IEEE Trans. Semicond. Manuf.* **2018**, *31*, 335–342. [CrossRef]
168. Zhou, L.; Chen, J.; Song, Z. Recursive Gaussian process regression model for adaptive quality monitoring in batch processes. *Math. Probl. Eng.* **2015**, 761280. [CrossRef]
169. Hosseinpour, F.; Ahmed, I.; Baraldi, P.; Zio, E.; Behzad, M.; Lewitschnig, H. A novel methodology based on long short-term memory stacked autoencoders for unsupervised detection of abnormal working conditions in semiconductor manufacturing systems. *Proc. Inst. Mech. Eng. Part O J. Risk Reliab.* **2024**, *239*, 1115–1133. [CrossRef]
170. Xie, Y. Enhancing Efficiency and Accuracy in Virtual Metrology for Semiconductors with Innovative Sparse Dimensionality Reduction and Advanced Ensemble Learning for Multi-Source Data Analysis. Ph.D. Thesis, The State University of New Jersey, Rutgers, New Brunswick, NJ, USA, 2024.
171. Serani, A.; Diez, M. A survey on design-space dimensionality reduction methods for shape optimization. *arXiv* **2024**, arXiv:2405.13944. [CrossRef]
172. Ghojogh, B.; Ghodsi, A.; Karray, F.; Crowley, M. Sufficient dimension reduction for high-dimensional regression and low-dimensional embedding: Tutorial and survey. *arXiv* **2021**, arXiv:2110.09620.
173. Ajana, S.; Acar, N.; Bretillon, L.; Hejblum, B.P.; Jacqmin-Gadda, H.; Delcourt, C. Benefits of dimension reduction in penalized regression methods for high-dimensional grouped data: A case study in low sample size. *Bioinformatics* **2019**, *35*, 3628–3634. [CrossRef]
174. Lima, M.; Neto, M.; Silva Filho, T.; Fagundes, R.A.D.A. Learning under concept drift for regression—A systematic literature review. *IEEE Access* **2022**, *10*, 45410–45429. [CrossRef]
175. Igual, J.; Gomes, H.M.; Pfahringer, B.; Bifet, A. Linear adaptive filtering for regression in data streams. *Int. J. Data Sci. Anal.* **2025**, *50*, 5017–5032. [CrossRef]
176. Samuel, K.M.; Ahmed, F. Continual learning strategies for 3D engineering regression problems: A benchmarking study. *J. Comput. Inf. Sci. Eng.* **2025**, *25*, 101003. [CrossRef]
177. Ou, F.; Suherman, J.; Zhang, C.; Wang, H.; Bom, S.; Davis, J.F.; Christofides, P.D. Industrial multi-machine data aggregation, AI-ready data preparation, and machine learning for virtual metrology in semiconductor wafer and slider production. *Digit. Chem. Eng.* **2025**, *15*, 100242. [CrossRef]
178. Alangari, N.; El Bachir Menai, M.; Mathkour, H.; Almosallam, I. Exploring evaluation methods for interpretable machine learning: A survey. *Information* **2023**, *14*, 469. [CrossRef]
179. Muckley, E.S.; Saal, J.E.; Meredig, B.; Roper, C.S.; Martin, J.H. Interpretable models for extrapolation in scientific machine learning. *Digit. Discov.* **2023**, *2*, 1425–1435. [CrossRef]
180. Nazir, Z.; Kaldykhanov, D.; Tolep, K.K.; Park, J.G. A machine learning model selection considering tradeoffs between accuracy and interpretability. In Proceedings of the 2021 13th International Conference on Information Technology and Electrical Engineering (ICITEE), Chiang Mai, Thailand, 14–15 October 2021; pp. 63–68.
181. Lang, C.I. Applications of Probabilistic Machine Learning Models to Semiconductor Fabrication. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2022.
182. Teh, H.Y.; Kempa-Liehr, A.W.; Wang, K.I.K. Sensor data quality: A systematic review. *J. Big Data* **2020**, *7*, 11. [CrossRef]
183. Siang, L.C.; Elnawawi, S.; Rippon, L.D.; O'Connor, D.L.; Gopaluni, R.B. Data quality over quantity: Pitfalls and guidelines for process analytics. *IFAC-PapersOnLine* **2023**, *56*, 7992–7999. [CrossRef]
184. Halder, P.; Mojumder, S. Physics-guided denoiser network for enhanced additive manufacturing data quality. *arXiv* **2025**, arXiv:2508.02712.
185. Sofianidis, G.; Rožanec, J.M.; Mladenic, D.; Kyriazis, D. A review of explainable artificial intelligence in manufacturing. In *Trusted Artificial Intelligence in Manufacturing: A Review of the Emerging Wave of Ethical and Human Centric AI Technologies for Smart Production*; Now Publishers: Norwell, MA, USA, 2021; pp. 93–113.
186. Lisboa, P.J.; Saralajew, S.; Vellido, A.; Fernández-Domenech, R.; Villmann, T. The coming of age of interpretable and explainable machine learning models. *Neurocomputing* **2023**, *535*, 25–39. [CrossRef]
187. Kusiak, A. Predictive models in digital manufacturing: Research, applications, and future outlook. *Int. J. Prod. Res.* **2023**, *61*, 6052–6062. [CrossRef]
188. Davis, J.C.; Mozumder, P.K.; Burch, R.; Fernando, C.; Apte, P.P.; Saxena, S.; Rao, S.; Vasanth, K. Automatic synthesis of equipment recipes from specified wafer-state transitions. *IEEE Trans. Semicond. Manuf.* **2002**, *11*, 527–536. [CrossRef]
189. Shim, J.; Cho, S.; Kum, E.; Jeong, S. Adaptive fault detection framework for recipe transition in semiconductor manufacturing. *Comput. Ind. Eng.* **2021**, *161*, 107632. [CrossRef]

190. Chien, C.F.; Hung, W.T.; Liao, E.T.Y. Redefining monitoring rules for intelligent fault detection and classification via CNN transfer learning for smart manufacturing. *IEEE Trans. Semicond. Manuf.* **2022**, *35*, 158–165. [CrossRef]
191. Li, Z.; Cai, R.; Ng, H.W.; Winslett, M.; Fu, T.Z.; Xu, B.; Yang, X.; Zhang, Z. Causal mechanism transfer network for time series domain adaptation in mechanical systems. *ACM Trans. Intell. Syst. Technol.* **2021**, *12*, 1–21. [CrossRef]
192. Thavarajah, M. Real time inferencing of semiconductor wafer probing process using Machine Learning. In Proceedings of the 2024 IEEE 40th International Electronics Manufacturing Technology (IEMT), Penang, Malaysia, 16–18 October 2024; pp. 1–6.
193. Horr, A.M. Real-time modeling for design and control of material additive manufacturing processes. *Metals* **2024**, *14*, 1273. [CrossRef]
194. Horr, A.M.; Drexler, H. Real-time models for manufacturing processes: How to build predictive reduced models. *Processes* **2025**, *13*, 252. [CrossRef]
195. Chen, Y.P.; Karkaria, V.; Tsai, Y.K.; Rolark, F.; Quispe, D.; Gao, R.X.; Cao, J.; Chen, W. Real-time decision-making for Digital Twin in additive manufacturing with Model Predictive Control using time-series deep neural networks. *J. Manuf. Syst.* **2025**, *80*, 412–424. [CrossRef]
196. Kim, E.S.; Choi, S.H.; Lee, D.H.; Kim, K.J.; Bae, Y.M.; Oh, Y.C. An oversampling method for wafer map defect pattern classification considering small and imbalanced data. *Comput. Ind. Eng.* **2021**, *162*, 107767. [CrossRef]
197. Han, T.; Taheri, Z.; Ko, H. Physics-informed neural networks For semiconductor film deposition: A Review. *arXiv* **2025**, arXiv:2507.10983. [CrossRef]
198. Di Cairano, S.; Kolmanovsky, I.V. Real-time optimization and model predictive control for aerospace and automotive applications. In Proceedings of the 2018 annual American control conference (ACC), Milwaukee, WI, USA, 27–29 June 2018; pp. 2392–2409.
199. Singgih, I.K. Production flow analysis in a semiconductor fab using machine learning techniques. *Processes* **2021**, *9*, 407. [CrossRef]
200. Leray, P.; De Gendt, S. Exploring machine learning for semiconductor process optimization: A systematic review. *IEEE Trans. Artif. Intell* **2024**, *5*, 5949–5989.
201. Karkaria, V.; Goeckner, A.; Zha, R.; Chen, J.; Zhang, J.; Zhu, Q.; Cao, J.; Gao, R.X.; Chen, W. Towards a digital twin framework in additive manufacturing: Machine learning and bayesian optimization for time series process optimization. *J. Manuf. Syst.* **2024**, *75*, 322–332. [CrossRef]
202. Liu, N.; Li, X.; Rajanna, M.R.; Reutzel, E.W.; Sawyer, B.; Rao, P.; Lua, J.; Phan, N.; Yu, Y. Deep neural operator enabled digital twin modeling for additive manufacturing. *arXiv* **2024**, arXiv:2405.09572. [CrossRef]
203. Senoner, J.; Netland, T.; Feuerriegel, S. Using explainable artificial intelligence to improve process quality: Evidence from semiconductor manufacturing. *Manag. Sci.* **2022**, *68*, 5704–5723. [CrossRef]
204. INFICON. Edge AI: A Semiconductor Process Control Revolution (INFICON, 2025). Available online: https://www.inficon.com/en/news/edge-ai-a-semiconductor-process-control-revolution (accessed on 26 August 2025).
205. Chen, M.; Wei, Z.; Li, L.; Zhang, K. Edge computing-based proactive control method for industrial product manufacturing quality prediction. *Sci. Rep.* **2024**, *14*, 1288. [CrossRef]
206. SemiEngineering. When and Where to Implement AI/ML in Fabs. (SemiEngineering, 2023). Available online: https://semiengineering.com/when-and-where-to-implement-ai-ml-in-fabs/?utm_source=chatgpt.com (accessed on 28 August 2025).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.