Data-centric Approach to Decision Making in Semiconductor Manufacturing: Best Practices and Future Directions

Tarun Parmar
(Independent Researcher)
Austin, TX
ptarun@ieee.org

***Abstract*—Data-centric approach to decision making has become increasingly crucial in semiconductor manufacturing, revolutionizing the industry's approach to efficiency, quality control, and cost reduction. The integration of advanced analytics, machine learning, and artificial intelligence enables real-time monitoring, predictive maintenance, and adaptive control systems, thereby minimizing downtime, reducing waste, and improving the overall equipment effectiveness. This study explores various types of data collected in semiconductor manufacturing, such as process parameters, equipment sensors, yield data, and quality metrics, and examines the role of advanced analytics techniques in extracting insights from these data. The importance of real-time data processing and analysis for rapid decision-making in semiconductor fabs is highlighted, along with the challenges of data quality, integration, and governance. The study also addresses the use of data visualization tools and techniques to present complex manufacturing data in an easily understandable format for decision-makers. Case studies of successful data-centric approaches in semiconductor manufacturing are examined, showing the benefits and lessons learned. The role of Industry 4.0 and the Internet of Things in enabling more comprehensive data collection and analysis is discussed, as well as the potential of edge computing and fog computing in processing data closer to the source. The integration of supply chain data with manufacturing data for more holistic decision making is explored, and the human factors in data-driven decision making, including the need for training and upskilling of the workforce, are addressed. Finally, the paper concludes with a discussion of future directions, including emerging technologies and trends that may shape data-centric decision-making in semiconductor manufacturing, such as advanced artificial intelligence, cellular networks, quantum computing, digital twins, and focus on sustainability and energy efficiency.**

Keywords*— data-driven decision making, semiconductor manufacturing, advanced analytics, machine learning, artificial intelligence, predictive maintenance, adaptive control systems, overall equipment effectiveness (OEE), equipment sensors, yield data, quality metrics, real-time data processing, data quality and governance, data visualization

I. INTRODUCTION

Data-centric decision making has become increasingly crucial in semiconductor manufacturing, revolutionizing the industry's approach to efficiency, quality control, and cost reduction [1]. As the demand for smaller, faster, and more complex semiconductor devices continues to grow, manufacturers are facing mounting pressure to optimize their processes and maintain a competitive edge. By leveraging the vast amounts of data generated throughout the production lifecycle, semiconductor companies can gain valuable insights into their operations, identify areas for improvement, and make informed decisions that drive innovation and productivity.

The integration of advanced analytics, machine learning, and artificial intelligence into semiconductor manufacturing processes enables real-time monitoring, predictive maintenance, and adaptive-control systems. These technologies allow manufacturers to detect anomalies, predict equipment failures, and optimize process parameters with unprecedented accuracy [2]. By harnessing the power of data, semiconductor companies can minimize downtime, reduce waste, and improve the overall equipment effectiveness (OEE) [3]. Data-driven approaches facilitate the development of more robust quality control measures, ensuring that products meet increasingly stringent specifications and reliability standards.

Cost reduction is another significant benefit of data-driven decision making in semiconductor manufacturing. By analyzing production data, companies can identify inefficiencies, streamline workflows, and optimize resource allocation [4]. This leads to reduced material waste, improved energy efficiency, and enhanced labor productivity. Additionally, data-driven insights enable manufacturers to make more informed decisions regarding capacity planning, inventory management, and supply chain optimization, further contributing to cost savings and improved profitability.

II. OVERVIEW

The semiconductor manufacturing process is a complex, multistep procedure that involves the creation of integrated circuits on silicon wafers. It begins with the production of ultrapure silicon ingots, which are sliced into thin wafers. These wafers then underwent a series of precise fabrication steps, including photolithography, etching, doping, and deposition. Photolithography uses light to transfer circuit patterns onto the wafer, whereas etching removes specific areas to create the desired structures. Doping introduces impurities that alter the electrical properties of the wafer, and deposition adds layers of various materials. Throughout the process, stringent cleanroom conditions were maintained to prevent contamination. After fabrication, the wafers were cut into individual chips, which were then packaged and tested before integration into electronic devices. This highly sophisticated process requires state-of-the-art equipment, specialized knowledge, and significant capital investment, making semiconductor manufacturing a technologically advanced and economically important industry.

III. EVOLUTION OF DATA ANLYTICS FROM TRADITIONAL METHODS TO MODERN APPROACHES

The evolution of data collection and analysis in semiconductor manufacturing has undergone significant transformations over the years, driven by technological advancements and the increasing complexity of the manufacturing processes.

Traditional methods of data collection in semiconductor manufacturing rely heavily on manual processes and a limited number of automated systems. Engineers and operators would manually record key parameters such as temperature, pressure, and processing times at various stages of the manufacturing process. This approach is time consuming, prone to human error, and provides only a snapshot of the production environment. Analysis of this data is often retrospective, with limited real-time insights into process variations or potential issues.

As semiconductor manufacturing became more sophisticated, the industry began to adopt more automated data-collection systems. Sensors and monitoring equipment are integrated into production lines, allowing continuous data capture of critical process parameters. This shift has enabled manufacturers to collect larger volumes of data with greater accuracy and consistency. However, the analysis of these data is still largely confined to statistical process control (SPC) methods, which focus on identifying deviations from predetermined specifications.

The advent of advanced process control (APC) has marked a significant leap forward in data utilization. APC systems incorporate feedback loops and predictive modeling, allowing real-time adjustments to manufacturing processes based on the collected data. This approach improves yield rates and reduces variability in semiconductor production. However, the full potential of the collected data is still not fully realized, owing to limitations in data storage and processing capabilities.

The emergence of big data technologies has revolutionized data collection and analysis in semiconductor manufacturing. Modern fabrication facilities are equipped with thousands of sensors, which generate massive amounts of data across multiple process steps. These data include not only traditional process parameters but also equipment sensor data, metrology results, and unstructured data from maintenance logs and operator notes.

Big-data approaches in semiconductor manufacturing leverage advanced analytics, machine learning, and artificial intelligence to extract meaningful insights from this vast pool of information. These techniques enable manufacturers to:

1. Predictive maintenance is performed, equipment failures are anticipated before they occur, and downtime is minimized.

2. Optimize process parameters in real time and continuously adjust manufacturing conditions to maximize the yield and quality.

3. Conduct root cause analysis with unprecedented speed and accuracy, identifying the source of defects or process variations across complex production chains.

4. Virtual metrology is implemented using historical data and machine-learning models to predict wafer characteristics without the need for physical measurements, thereby reducing cycle times and increasing throughput.

5. Develop digital twins of manufacturing processes, allowing for the simulation and optimization of production scenarios without disrupting actual operations.

The integration of Internet of Things (IoT) technologies has further enhanced data collection capabilities, enabling seamless communication between various manufacturing equipment and centralized data repositories. This interconnectedness facilitates a holistic view of the entire manufacturing ecosystem from raw material intake to final product testing.

Cloud and edge computing technologies have also played a crucial role in the evolution of data analysis in semiconductor manufacturing. These platforms provide the necessary computational power and storage capacity to handle enormous volumes of data generated in modern fabs. Edge computing, in particular, allows for real-time processing of critical data directly at the source, enabling immediate decision-making and reducing latency.

As the semiconductor industry continues to push the boundaries between miniaturization and complexity, the role of advanced data collection and analysis has become increasingly critical. The future of semiconductor manufacturing is likely to see an even greater integration of artificial intelligence and machine learning techniques, potentially leading to fully autonomous fabrication facilities capable of self-optimization and adaptive manufacturing.

IV. ROLE AND IMPORTANCE OF DATA AND RELATED FACTORS

The evolution of data collection and analysis in semiconductor

A. Type of Data collected in semiconductor manufacturing:

Semiconductor manufacturing involves collecting and analyzing vast amounts of data to ensure product quality, optimize processes, and improve the overall efficiency. The collected data can be broadly categorized into the following groups:

Process parameters: These include the temperature, pressure, gas flow rates, and other variables that are critical to the various steps in semiconductor fabrication. The process parameters were monitored and controlled throughout the manufacturing process to maintain consistency and quality.

Equipment sensors: Modern semiconductor manufacturing equipment is equipped with numerous sensors that continuously monitor the state of machinery. These sensors collect data on factors such as vibration, power consumption, and component wear, which can be used for predictive maintenance and to detect potential issues before they impact the production.

Yield data: The yield refers to the percentage of functional chips produced from a wafer. Yield data are collected at various stages of the manufacturing process, including wafer-level and final package testing. This information is crucial for identifying process bottlenecks and improving overall production efficiency.

Quality metrics: These encompass a wide range of measurements and tests performed on semiconductor devices to ensure that they meet the specifications. Quality metrics may include electrical characteristics, physical dimensions, and reliability test results. The defective parts metric is a crucial quality metric from the customer perspective. This metric measures the number or percentage of defective parts that reach the customer and is often expressed as Defective Parts Per Million (DPPM) or Defective Parts Per Billion (DPPB). A lower DPPM or DPPB indicates higher quality and reliability of the semiconductor products.

The defective part metric is significant for several reasons:

1. Customer satisfaction: Lower defect rates lead to higher customer satisfaction and loyalty as customers receive products that consistently meet their expectations.

2. Cost reduction: By identifying and addressing defects early in the manufacturing process, companies can reduce costs associated with warranty claims, returns, and rework.

3. Competitive advantage: A low rate of defective parts can be a key differentiator in the highly competitive semiconductor industry, potentially leading to increased market share and customer trust.

4. Continuous improvement: Tracking and analyzing defective parts data helps manufacturers identify trends, root causes, and opportunities for process improvements, thereby driving ongoing quality enhancements.

To minimize defective parts and improve customer quality, semiconductor manufacturers employ various strategies, including statistical process control, advanced inspection techniques, and rigorous testing protocols, throughout the production process. Additionally, data analytics and machine learning algorithms are increasingly being used to predict and prevent defects before they occur, further enhancing product quality and customer satisfaction.

B. Advanced analytics to extract insights from data:

Advanced analytics techniques, such as machine learning, artificial intelligence, and predictive modeling, are essential for extracting valuable insights from manufacturing data. These technologies enable manufacturers to optimize processes, improve quality, and enhance decision making by analyzing vast amounts of operational data.

Machine learning algorithms identify patterns and correlations in historical production data, allowing manufacturers to predict issues, optimize schedules, and fine-tune processes [5]. Artificial intelligence, particularly deep learning and neural networks, enhances quality control through real-time image analysis, and enables adaptive automation systems.

Predictive modeling forecasts future outcomes based on historical data and current conditions, facilitating proactive maintenance scheduling, downtime reduction, and resource optimization. The integration of these advanced analytics techniques promotes data-driven decision making across organizations, leading to significant improvements in efficiency, quality, and productivity.

Furthermore, these technologies contribute to the development of smart factories and Industry 4.0 initiatives, creating highly responsive and adaptive production systems that can quickly adjust to changing market demands and operational conditions [6].

C. Real-time data process for rapid decision-making in fabs:

Real-time data processing and analysis are essential for rapid decision making in semiconductor fabrication facilities (fabs). This approach enables prompt issue detection and response, predictive maintenance, quality control, yield optimization, and rapid experimentation. By leveraging real-time data, fabs can improve the yield, reduce downtime, and enhance the overall performance of the competitive semiconductor industry. Continuous monitoring and analysis of equipment and process data allows for immediate intervention, preventing potential yield losses and minimizing production disruptions. Predictive maintenance strategies based on real-time data analysis help optimize maintenance schedules and reduce unplanned downtimes. Real-time quality control and yield optimization ensure consistent product quality and maximize financial gains. Additionally, real-time data analysis facilitates agile process optimization and experimentation, enabling semiconductor manufacturers to remain competitive and responsive to technological advancements.

D. Integration of data driven support systems with MES:

Data-driven decision support systems (DSS) have become increasingly important in modern manufacturing environments, offering valuable insights and recommendations for optimizing production processes. The implementation of these systems involves several key steps:

1. *Data collection and integration*: The first step is to gather relevant data from various sources within a manufacturing facility, including sensors, machines, and existing software systems. These data were then integrated into a centralized repository or data lake to ensure the data quality and consistency.

2. *Data preprocessing and analysis*: Raw data were cleaned, normalized, and transformed to make them suitable for analysis. Advanced analytics techniques, such as machine learning algorithms and statistical models, are applied to extract meaningful patterns and insights from data.

3. *Development of decision models*: Based on the analyzed data, decision models were created to address specific manufacturing challenges. These models include predictive maintenance algorithms, production scheduling optimizers, or quality control systems.

4. *User interface design*: A user-friendly interface was developed to present insights and recommendations generated by the decision support system. This interface should be intuitive and accessible to various stakeholders within the organization.

5. *Integration with Manufacturing Execution Systems (MES)*: To maximize the effectiveness of data-driven DSS, it is crucial to integrate them with the existing MES. This integration can be achieved using the following approach:

a. *API-based integration*: Developing application programming interfaces (APIs) that allow seamless communication between DSS and MES, enabling real-time data exchange and decision implementation.

b. *Middleware solutions*: Implementation of middleware software that acts as a bridge between the DSS and MES, facilitating data transfer and system interoperability.

c. *Cloud-based integration*: Utilizing cloud platforms to host both DSS and MES allows for easier integration and scalability.

6. *Training and Change Management*: Implementing a data-driven DSS requires training employees to effectively use the new system and adapt to new decision-making processes. Change management strategies should be employed to ensure smooth adoption and minimize resistance to change.

7. *Continuous improvement and maintenance*: Regular monitoring, evaluation, and updating of DSS are essential to maintain its effectiveness and relevance. This

includes refining decision models, incorporating new data sources, and addressing technical issues that may arise.

By successfully implementing and integrating data-driven decision support systems with existing MES, manufacturers can achieve significant improvements in operational efficiency, product quality, and overall productivity. These systems enable more informed decision making, faster response times to production issues, and better utilization of resources throughout the manufacturing process.

E. Data visualization tools to present data to decision makes:

Data visualization is crucial for presenting complex manufacturing data to decision makers. Interactive dashboards provide real-time updates on key performance indicators and production metrics, incorporating various chart types for trend analysis, production-volume comparisons, and bottleneck identification. 3D modeling and virtual reality enable virtual walkthroughs on production floors, helping optimize workflows. Sankey diagrams visualize material flows and energy consumption, whereas data clustering and dimensionality reduction techniques manage large datasets. Geospatial visualization aids in understanding multisite operations and supply chains, and time-series analysis visualizations reveal trends and patterns over time. These tools and techniques transform raw data into actionable insights, facilitating informed decision-making and process optimization in manufacturing.

V. DATA QUALITY, INTEGRATION AND GOVERNANCE CHALLENGES

Data quality, integration, and governance present significant challenges in semiconductor manufacturing environments, owing to the complexity and scale of operations.

Data quality issues can arise from various sources, including sensor malfunctions, human errors in data entry, and inconsistencies across different systems. Even minor inaccuracies in semiconductor manufacturing can lead to significant problems in product quality and yield. Ensuring data accuracy and completeness requires robust validation processes, regular calibration of measurement equipment, and the implementation of automated error detection mechanisms.

Data integration challenges stem from the diverse arrays of systems and equipment used in semiconductor fabrication. Each piece of equipment can generate data in different formats, making it difficult to consolidate and analyze information across the entire production line. Integrating data from legacy systems with new technologies further complicates this issue. Implementing standardized data formats and developing flexible integration platforms are crucial steps in addressing these challenges.

Data governance in semiconductor manufacturing involves the establishment of policies and procedures for data management throughout the manufacturing lifecycle. This includes defining the data ownership, access control, and retention policies. Given the sensitive nature of semiconductor manufacturing processes, maintaining data security and confidentiality is of paramount importance. The implementation of robust data governance frameworks ensures compliance with industry regulations and protects intellectual property.

Semiconductor manufacturers have adopted several strategies to address these challenges. Implementing advanced data management systems with built-in quality checks and integration capabilities can improve data quality and streamline integration processes. Establishing cross-functional teams to oversee data governance initiatives ensures that policies are comprehensive and aligned with the business objectives. In addition, leveraging artificial intelligence and machine learning technologies can enhance data quality through predictive maintenance and anomaly detection.

Continuous monitoring and improvement of data management practices are essential for the rapidly evolving semiconductor industry. Regular audits of data quality, integration processes, and governance policies can help identify areas for improvement, and ensure that data management strategies remain effective in supporting manufacturing excellence.

VI. OTHER DECISION-MAKING FACTORS

A. Supply Chain data:

The integration of supply chains and manufacturing data is crucial for holistic decision-making in modern industries. This integration provides a comprehensive view of the entire value chain, enabling organizations to optimize processes, improve demand forecasting, enhance inventory management, and strengthen quality control. It also facilitates better logistics and transportation planning and supports more effective risk management. By leveraging this integrated approach, companies can improve efficiency, reduce costs, enhance quality, and increase customer satisfaction across their operations.

B. Human factor including need for training and up-skilling:

Data-driven decision making is increasingly important across industries, leveraging advanced analytics and AI. However, human factors are crucial for the interpretation, implementation, and ethical considerations. To effectively integrate data-driven approaches, organizations should focus on the following:

1. Comprehensive training and upskilling initiatives

2. Developing data literacy, analytical skills, and critical thinking

3. Enhancing competencies in data collection, statistical analysis, and visualization

4. Understanding limitations and potential biases in data-driven approaches

5. Fostering cross-functional collaboration and communication skills

6. Encouraging continuous learning and adaptation to evolving technologies

By prioritizing human factors, organizations can harness the full potential of data-driven decision making while maintaining the importance of human expertise and judgment in the process.

VII. CHALLENGES AND OPPORTUNITIES OF IMPLEMENTING DATA-CENTRIC APPROACHES IN LEGACY FABS

The implementation of data-centric approaches in legacy semiconductor fabs presents significant challenges and opportunities. Challenges include integrating modern data systems with outdated equipment, cultural resistance to change, and ensuring data quality and standardization. However, the potential benefits are substantial. These include improved operational efficiency through advanced analytics and machine learning, enhanced predictive maintenance capabilities, more agile and responsive operations, data-driven innovation in product development and quality control, and better positioning for future technological advancement. While the transition may be complex and require substantial investment, the long-term advantages in efficiency, maintenance, agility, and competitiveness make it a crucial endeavor for legacy fabs to remain viable in an evolving industry landscape.

VIII. CONCLUSION

The future of data-centric decision-making in semiconductor manufacturing is poised for significant advancements. Key trends include:

1. AI and machine learning for process optimization and predictive maintenance

2. Edge computing and advanced communication networks for real-time analysis

3. Quantum computing for solving complex optimization problems

4. Digital twin technology for accurate simulations and scenario planning

5. Enhanced sensor technology and IoT for granular data collection

6. Data-driven approaches to improve sustainability and energy efficiency

7. Novel data analysis methods for emerging materials and manufacturing techniques

These developments promise to enhance the efficiency, reduce costs, and drive innovation in the semiconductor industry.

REFERENCES

[1] K.-H. Chang *et al.*, "Big data analytics energy-saving strategies for air compressors in the semiconductor industry – an empirical study," *International Journal of Production Research*, vol. 60, no. 6, pp. 1782–1794, Jan. 2021, doi: 10.1080/00207543.2020.1870015.

[2] M. Azamfar, X. Li, and J. Lee, "Deep Learning-Based Domain Adaptation Method for Fault Diagnosis in Semiconductor Manufacturing," *IEEE Transactions on Semiconductor Manufacturing*, vol. 33, no. 3, pp. 445–453, May 2020, doi: 10.1109/tsm.2020.2995548.

[3] M. S. Abd Rahman, E. Mohamad, and A. A. Abdul Rahman, "Enhancement of overall equipment effectiveness (OEE) data by using simulation as decision making tools for line balancing," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 2, p. 1040, May 2020, doi: 10.11591/ijeecs.v18.i2.pp1040-1047.

[4] E. Flores-Garcia, M. Wiktorsson, M. Jackson, and J. Bruch, "Decision-making approaches in process innovations: an explorative case study," *Journal of Manufacturing Technology Management*, vol. 32, no. 9, pp. 1–25, Dec. 2020, doi: 10.1108/jmtm-03-2019-0087.

[5] Q. Ma, H. Li, and A. Thorstenson, "A big data-driven root cause analysis system: Application of Machine Learning in quality problem solving," *Computers & Industrial Engineering*, vol. 160, p. 107580, Aug. 2021, doi: 10.1016/j.cie.2021.107580.

[6] A. Bousdekis, K. Lepenioti, G. Mentzas, and D. Apostolou, "A Review of Data-Driven Decision-Making Methods for Industry 4.0 Maintenance Applications," *Electronics*, vol. 10, no. 7, p. 828, Mar. 2021, doi: 10.3390/electronics10070828.