Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 3, No. 1, 2026 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

Data-Driven Bottleneck Detection with Minimal Information in Manufacturing Systems

FRANK, Leede [1*]

[1] University of California-Berkeley, USA

***** FRANK, Leede is the corresponding author, E-mail: LedeeFrank123@gmail.com***

Abstract: Bottleneck detection is crucial for optimizing production systems, but many current methods in manufacturing environments rely on large amounts of process data that are difficult to obtain in real time. This paper explores a novel data-driven approach that uses minimal information to identify and analyze production bottlenecks. By combining minimal information with various activity cycle and queuing cycle methods, it maintains bottleneck detection accuracy while reducing data collection requirements.

Keywords: Bottleneck Detection, Minimal Information, Manufacturing Systems, Queue Directed Graph.

Disciplines: Applied Mathematics. **Subjects:** Mathematical Modeling.

DOI: https://doi.org/10.70393/6a6374616d.333634 **ARK:** https://n2t.net/ark:/40704/JCTAM.v3n1a09

1 INTRODUCTION

In modern manufacturing systems, optimizing production efficiency remains a critical objective for organizations striving to maintain competitive advantage. Despite the advancements in automation and data collection technologies, many manufacturing processes continue to face significant challenges due to production bottlenecks. These bottlenecks, often the result of inefficiencies or delays in certain parts of the production line, can severely impact overall system performance. The traditional methods for bottleneck detection typically require large volumes of detailed data, which, while highly informative, can be difficult to collect or process in real-time, especially in complex manufacturing environments. This presents a fundamental limitation for companies that either lack the infrastructure to gather such data or require timely insights to prevent delays.[1]

A shift in focus toward more accessible, data-driven approaches is gaining traction. Minimal Information (MI) methods, which utilize only essential data points—such as product IDs, workstation IDs, and station departure timestamps—offer a promising alternative. The primary advantage of this approach lies in its ability to simplify the data requirements without sacrificing the accuracy needed to effectively detect bottlenecks. Although MI-based models have seen some application in various contexts, their use for bottleneck detection in manufacturing systems remains an under-explored area. Therefore, this study aims to fill this gap by investigating the potential of MI to offer robust bottleneck detection capabilities, even in scenarios where data access is limited. This exploration is particularly relevant to industries where real-time data collection is constrained or not feasible, and the need for scalable solutions is urgent.[2]

The central research problem addressed in this study concerns the feasibility and reliability of using minimal data to identify production bottlenecks. Specifically, it questions whether the limited set of data, such as product IDs, workstation IDs, and timestamps, is sufficient for constructing an accurate model capable of pinpointing bottlenecks in a manufacturing environment. While existing research has explored the use of minimal data in other contexts, its application for bottleneck detection in manufacturing systems remains largely unexplored.[3]

To guide the investigation, the following research questions are proposed:

RQ1: To what extent can Minimal Information (MI) effectively reconstruct the production layout and identify bottleneck points in a manufacturing system?

RQ2: How do different bottleneck detection methods, such as the Active Period Method (APM) and Queue Period Method (QPM), perform when applied to MI-based models, and what are the comparative benefits and limitations of these methods?

RQ3: What are the key challenges in applying MI for real-time bottleneck detection, and how can these challenges be addressed in the modeling process?

These questions not only guide the development of a new methodology but also seek to assess its real-world applicability and limitations, particularly in the context of sparse data environments.

The primary objective of this research is to introduce a novel, data-driven approach for bottleneck detection in manufacturing systems that relies on minimal information. By focusing on a limited set of key data points—such as product IDs, workstation IDs, and station departure timestamps—this model aims to minimize the burden of data collection while ensuring that bottleneck identification remains accurate and efficient.[4] This shift toward using minimal data is particularly significant for environments where comprehensive data collection is not feasible or is constrained by existing technological infrastructure. In this sense, the approach not only promises to optimize operational efficiency but also offers scalability, making it adaptable to diverse industrial settings.

This study contributes to the field in two significant ways. First, it proposes a methodology that reduces the need for extensive data while still providing reliable insights into bottleneck identification.[5] By constructing a Queue Directed Graph (QDG) based on minimal information, the methodology provides a foundation for analyzing production bottlenecks without overwhelming the system with excessive data. Second, this paper presents a case study at Bosch Thermotechnology, applying the proposed methodology to a complex manufacturing environment. This real-world example not only illustrates the practical applicability of the model but also highlights the challenges encountered during implementation. By doing so, the research emphasizes how minimal data, if strategically utilized, can reveal underlying inefficiencies within the production process.

The significance of this research lies in its potential to revolutionize bottleneck detection in manufacturing systems by providing a simple yet effective approach that relies on minimal information. The ability to detect bottlenecks using only key data points—such as product and workstation IDs and departure timestamps—presents a scalable solution for manufacturers with limited access to comprehensive data. This research opens the door to more flexible, cost-effective methods of production monitoring, which could lead to significant improvements in operational efficiency. Furthermore, the findings contribute to the growing body of knowledge on how data-driven approaches can be applied in contexts where full-scale data acquisition is challenging or impractical.

2 LITERATURE REVIEW

Bottleneck detection has been a fundamental area of research in manufacturing systems optimization. Identifying bottlenecks is crucial as they determine the maximum throughput and efficiency of production processes. Historically, bottleneck detection relied on comprehensive and often complex datasets that captured various aspects of the manufacturing process, such as machine performance, throughput rates, and worker activities. These methods, although effective in certain contexts, often demand a large amount of real-time data, which may not always be readily available or feasible, especially in smaller-scale or less automated settings. Moreover, such methods may not be adaptable to dynamic production environments, where conditions can fluctuate frequently.

Recent advancements in bottleneck detection have focused on leveraging more accessible and minimal data, which can reduce the burden of data collection while maintaining the accuracy of bottleneck identification. This shift towards minimal information (MI) approaches, which utilize basic data points like product IDs, workstation IDs, and timestamps, offers the potential for broader applicability, especially in industries with limited access to detailed data. However, despite the promising nature of these techniques, the full potential of MI-based bottleneck detection remains underexplored in the literature, and further studies are required to refine and validate these approaches in real-world manufacturing contexts.

Several models have emerged to address these challenges. One such approach is the use of Queue Directed Graphs (QDGs) to model production workflows. These graphs represent the flow of materials through a manufacturing system, capturing the relationships between workstations, products, and times, which can then be analyzed to identify potential bottlenecks. The ability of QDGs to provide a visual and analytical representation of production flows makes them particularly useful for understanding complex manufacturing systems, where multiple factors contribute to delays.

In addition, models such as the Active Period Method (APM) and the Queue Period Method (QPM) have been applied in various studies to quantify bottleneck impacts by focusing on different operational periods within a production process. These models offer flexibility in terms of identifying which stages of the production cycle contribute most significantly to bottlenecks. However, while APM and QPM have been applied successfully in specific contexts, the need for robust methods that can handle sparse and real-time data remains an unresolved issue in the field. In this regard, Yin (2025) addresses these challenges through a real-time bottleneck detection framework that incorporates MI, using APM and QPM as foundational methods for analysis.[6]

2.1 Active Period Methods and Queue Period Methods

Active Period Methods (APM) and Queue Period Methods (QPM) represent two of the most widely applied techniques for bottleneck detection in manufacturing systems. Both approaches focus on different phases of the production cycle to identify when and where congestion occurs. APM measures the active times when a workstation or process is actively engaged in production, while QPM focuses on the queue lengths at various workstations, tracking how long items remain in a waiting state.

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 3, No. 1, 2026 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

Despite their widespread use, these methods are not without limitations. One of the primary challenges is the assumption that data about queues or active periods is consistently available, which is often not the case in real-world manufacturing systems where data sparsity and irregularity can lead to incomplete or inaccurate analyses. Sun (2024) suggests that incorporating additional visualization techniques alongside these models could help mitigate this issue, providing more actionable insights for bottleneck detection and allowing for better interpretation of the data . However, further research is needed to refine these methods and explore their applicability in dynamic environments with incomplete data.[7]

Additionally, QPM has been shown to work well in controlled settings where queue sizes can be easily monitored. However, its utility is diminished when there is a high degree of variability in production cycles or when some workstations do not report data consistently. As discussed by Yin (2025), this issue can be partially addressed by integrating machine learning models that can predict queue lengths based on minimal data inputs . Yet, the success of these integrated models hinges on their ability to work with sparse data, making it imperative to continue exploring alternative ways to enhance the predictive power of QPM.[8]

2.2 Minimal Information for Bottleneck Detection

Minimal Information (MI) has emerged as a promising approach for bottleneck detection, particularly in environments where extensive data collection is either impractical or impossible. The concept of MI involves using only the most essential data—such as product IDs, workstation IDs, and timestamps—to construct a framework for identifying and addressing bottlenecks. The theoretical underpinnings of MI-based bottleneck detection are grounded in the belief that a limited but strategically chosen set of data can offer meaningful insights into the production process, especially when combined with advanced data analysis techniques.

Yin (2025) emphasizes the potential of MI-based approaches in semiconductor manufacturing, where the volume of data available for analysis is often constrained.[9]By leveraging MI, his research has shown that it is possible to detect and optimize bottlenecks with a minimal data footprint, thus overcoming some of the limitations associated with traditional methods that require extensive data collection . However, the challenge remains in adapting these methods to more complex, real-world manufacturing environments, where data may be sparse or inconsistent.

One of the key contributions of MI approaches is their ability to reduce the data collection burden, which is particularly important in industries with high operational costs and where the infrastructure for detailed data gathering may be lacking. Further research in this area should explore how MI can be used in combination with real-time data streams to enhance bottleneck detection, especially in environments where production schedules and processes are subject to frequent changes. To some extent, the integration of real-time data could help address the issue of data sparsity, providing a more holistic view of production performance and facilitating quicker interventions to resolve bottlenecks.[10]

2.3 Gaps in the Literature and Future Research Directions

While the literature on bottleneck detection has grown significantly in recent years, several gaps remain that warrant further exploration. Most notably, there is a lack of comprehensive studies that combine MI with advanced machine learning models to predict and optimize bottlenecks in real-time manufacturing environments. [11]Furthermore, many existing methods, including APM and QPM, rely on consistent and extensive data inputs, which are not always available in practical scenarios. This highlights the need for more research into how MI-based approaches can be combined with predictive analytics to improve both the accuracy and reliability of bottleneck detection.

Additionally, while QDGs and queue-based models have been widely used in theoretical settings, their real-world applications, particularly in environments characterized by sparse data or high variability, remain underexplored. [12]The integration of these models with machine learning algorithms, as suggested by Huang (2025), holds significant promise for improving the adaptability and scalability of bottleneck detection systems . However, further studies are needed to refine these models, address their limitations, and explore their applicability in diverse industrial settings.[13]

3 METHODOLOGY

This paper outlines the methodology employed to investigate the feasibility and effectiveness of using Minimal Information for real-time bottleneck detection in manufacturing systems. The central premise of this study is that by leveraging only essential data points—product IDs, workstation IDs, and departure timestamps—it is possible to build a model that can reliably identify bottlenecks in the production process. The methodology is divided into several key stages, including data collection, construction of the Queue Directed Graph, and the application of bottleneck detection methods. These stages are designed to be iterative, with each stage building on the previous one to ensure that the final model is both robust and adaptable to real-world manufacturing environments.[14]

The decision to adopt a MI-based approach stems from the recognition that, in many manufacturing settings, collecting exhaustive data for each production process is neither feasible nor practical. The approach described in this chapter, therefore, seeks to minimize the data requirements

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 3, No. 1, 2026 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

while maintaining a high degree of accuracy in identifying production inefficiencies. Given the constraints of available data, it is crucial that the methodology is flexible and can adapt to different production environments where data may be sparse or irregular.[15]

3.1 Data Collection

Data collection is the first critical step in the methodology, as the accuracy of the model depends on the quality and type of data used. In this study, the data required for bottleneck detection is minimal, consisting primarily of product IDs, workstation IDs, and timestamps marking the departure of products from each workstation. These data points are often available in manufacturing systems, particularly in those with automated tracking capabilities. However, even in environments where data collection is not fully automated, these basic data points can often be manually extracted or inferred from other system logs.[16]

While this minimal data approach significantly reduces the complexity of data collection, it also introduces potential challenges related to data completeness and consistency. [17]One of the primary challenges faced in this study was ensuring that the timestamps were accurate and that products could be reliably tracked through the system. In some cases, missing or incorrect timestamps led to gaps in the data, which required careful handling. To address these issues, missing data points were estimated based on neighboring production cycles or, where necessary, excluded from the analysis to maintain data integrity.[18]

Further, the accuracy of workstation IDs was also a critical factor in ensuring that the system could correctly trace products through the production line. In cases where products were moved manually or bypassed certain workstations, the system needed to be able to handle these exceptions. The methodology proposed here involves a hybrid approach, using both the available data and expert knowledge to fill in gaps or adjust for anomalies that might occur due to production irregularities.[19]

3.2 Queue Directed Graph (QDG) Construction

Once the necessary data points are collected, the next step involves constructing the Queue Directed Graph (QDG), which serves as the foundational framework for bottleneck detection. A QDG is a directed graph where nodes represent workstations, and edges represent the flow of products between them. Each node stores information about the workstation, such as the time products spend at the station, while each edge stores information about the transitions between workstations.[20]

The construction of the QDG is a multi-step process. First, the product IDs and workstation IDs are mapped to the nodes in the graph. The timestamps serve as the edges, representing the flow of products from one workstation to the next. The weight of each edge is determined by the time difference between the departure timestamps of products at adjacent workstations. This process is iterative, as the QDG is refined over time based on the accumulation of new data.[21]

One of the main challenges in constructing the QDG is dealing with production processes that involve parallel workstations or rework. In cases where products are processed simultaneously at multiple workstations, the QDG must account for these parallel flows without double-counting the products or artificially inflating bottleneck measures. To address this, the methodology incorporates a flexible approach to parallelism, where multiple edges can be created for a single product flow, depending on the production configuration.[22]

Another challenge faced during QDG construction was handling cases where products bypass certain workstations due to production variances. In such cases, the QDG must be adapted to ensure that the product flow is accurately represented, even though it does not follow the standard sequence of workstations. This required a dynamic adjustment of the graph structure, allowing for missing or out-of-sequence workstations to be temporarily excluded or marked as exceptions.[23]

3.3 Bottleneck Detection Methods

With the QDG constructed, the next step is to apply bottleneck detection methods to identify production delays and inefficiencies. Several methods have been proposed for bottleneck detection in manufacturing systems, with the two most commonly used being the Active Period Method (APM) and the Queue Period Method (QPM). These methods focus on different aspects of the production process, offering complementary insights into where bottlenecks may be occurring.

Active Period Method (APM): This method identifies bottlenecks by analyzing the periods when workstations are actively processing products. Workstations with longer active periods, relative to their total operating time, are more likely to be bottlenecks. The method is particularly useful in detecting workstations that are frequently engaged in processing but experience delays due to limited capacity.[24]

Queue Period Method (QPM): In contrast, the QPM focuses on the length of time that products spend in queues at various workstations. The longer products spend in a queue, the more likely it is that the workstation is a bottleneck. The QPM method is particularly effective in identifying points in the production process where congestion is building up, as it tracks the accumulation of products over time.[25]

Both methods are implemented in the model, and the results from APM and QPM are compared to determine where bottlenecks occur. By combining these methods, it is possible to identify both stationary and dynamic bottlenecks, providing a comprehensive view of production inefficiencies.

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 3, No. 1, 2026 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

3.4 Challenges and Adjustments in the Research Process

While the methodology proposed here is based on a clear theoretical framework, several challenges emerged during the research process. One significant challenge was the integration of minimal data into the existing bottleneck detection models, which often assume access to more detailed information. To some extent, this limitation was addressed by adjusting the models to work with the available data, but there remains uncertainty regarding how well these models perform in highly dynamic or data-sparse environments.[26]

Another difficulty was related to the real-time application of the methodology. While the theoretical framework is robust, real-time data streams introduce additional complexities, such as delays in data transmission or incomplete datasets due to system failures. Addressing these issues required the development of a dynamic, adaptive approach that could accommodate real-time fluctuations and maintain the integrity of the model's outputs.[27]

Furthermore, the reliance on timestamps for flow analysis introduced challenges in terms of data accuracy. In many manufacturing environments, timestamps are not always perfectly aligned, and discrepancies between workstations can lead to errors in the flow calculations. These discrepancies were mitigated by using estimation techniques and expert knowledge to align timestamps where necessary, but further research is needed to refine these methods for better accuracy and reliability in real-time applications.[28]

4 CASE STUDY

The applicability of the proposed methodology for bottleneck detection using Minimal Information (MI) is tested through a case study at Bosch Thermotechnology, a prominent player in the global manufacturing sector. Bosch operates in a complex environment where production workflows are intricate, and the need for effective bottleneck identification is particularly critical to maintain operational efficiency. This case study aims to assess the feasibility of applying the MI-based methodology to a real-world production system, focusing on its ability to detect bottlenecks with minimal data. The study was conducted on a segment of the production line at Bosch Thermotechnology, involving the manufacturing of heat pumps, an area that often experiences variability due to high customization in product configurations and varying production cycles.

This section outlines the specifics of the case study, including data collection, the application of the Queue Directed Graph (QDG), and the detection of bottlenecks using the Active Period Method (APM) and Queue Period Method (QPM). It also presents the challenges encountered in implementing the methodology in a real-world manufacturing setting and the adjustments made to address these challenges.

4.1 Data Collection and Preparation

The data collected for this case study consisted of three key variables: product IDs, workstation IDs, and departure timestamps. These variables were chosen because they represent the minimal data points required by the proposed MI-based methodology. Bosch Thermotechnology's manufacturing system generates this data automatically at various points in the production process, with each product being assigned a unique identifier and tracked through different workstations.

Data preparation involved cleaning and preprocessing the raw data to ensure consistency and accuracy. One significant challenge during this stage was ensuring the accuracy and synchronization of timestamps across different workstations. In some instances, timestamps recorded at different stations were slightly out of sync due to network delays or machine malfunctions, leading to discrepancies in the product flow. To address this issue, we implemented a smoothing algorithm that adjusted the timestamps by using neighboring workstation data. This step was crucial for maintaining the integrity of the QDG construction, as misaligned timestamps could result in incorrect flow calculations, potentially skewing bottleneck detection results.

Another challenge was handling missing data, as occasional errors in the production system led to incomplete entries for some products. In such cases, missing data were imputed based on the average time spent at similar workstations or, when feasible, by consulting operational logs for that period. While this approach was not without limitations, it allowed for the continuation of the analysis and ensured that the dataset remained sufficiently complete for meaningful results.

4.2 Construction of the Queue Directed Graph (QDG)

With the data prepared, the next step was to construct the Queue Directed Graph, which serves as the foundation for identifying bottlenecks. The QDG was constructed by mapping workstation IDs to graph nodes and using departure timestamps to define the edges representing product flows between stations. In this case study, the product flow was largely sequential, with only a few parallel workstations, allowing the QDG to follow a relatively straightforward structure. However, there were some instances where products bypassed certain workstations due to the customized nature of the production process, requiring dynamic adjustments to the graph to account for missing nodes.

One of the significant challenges encountered during the construction of the QDG was the handling of parallel workstations. Some products were processed simultaneously at multiple stations, which required careful modeling to avoid inflating the flow or incorrectly identifying bottlenecks. To address this, the methodology incorporated a virtual queue system, where parallel workstations were represented as

Journal of Computer Technology and Applied Mathematics
Journal Home: http://jctam.suaspress.org | CODEN: JCTAAD
Vol. 3, No. 1, 2026 | ISSN 3007-4126 (Print) | ISSN 3007-4134 (Online) | ISSN 3007-4142 (Digital)

separate nodes connected to a shared virtual edge. This approach allowed for an accurate representation of product flows without distorting the analysis.

The QDG also needed to accommodate the fact that some workstations involved rework or retries. In such cases, products returned to previous stations, creating feedback loops in the system. This presented a challenge for maintaining the accuracy of the QDG, as cycles and feedback loops could complicate the flow analysis. To resolve this, we introduced a dynamic re-entry mechanism that allowed products to be temporarily reassigned to their previous nodes without disrupting the overall structure of the QDG.

The QDG is represented mathematically by the graph G = (N, E), where N is the set of nodes (workstations) and E is the set of edges (transitions between workstations). The weight of each edge w_{ij} is given by:

$$w_{ij}=T_{ij}-T_{i,start}$$

where:

T_{ij} is the departure timestamp of product p from workstation i to workstation j,

$T_{i,start}$ is the start timestamp at workstation *i*.

4.3 Bottleneck Detection

Once the QDG was constructed, the next step was to apply the bottleneck detection methods: the Active Period Method (APM) and the Queue Period Method (QPM). These methods were selected due to their ability to identify bottlenecks from different perspectives—APM by focusing on active processing periods and QPM by analyzing queue lengths. Both methods were applied in tandem to provide a comprehensive analysis of potential bottlenecks across the production line.

Active Period Method (APM): In this case study, the APM was used to identify workstations with prolonged active processing periods, which could indicate that the workstation was struggling to handle the workload efficiently. By analyzing the frequency and duration of workstation engagement, we identified several workstations where the time spent actively processing products exceeded the threshold typically associated with bottleneck stations. This initial identification helped pinpoint key areas that warranted further investigation.

The active processing time A_i for workstation iii over a time window is $[t_{start}, t_{end}]$ calculated as: $A_i=\sum_{p\in P}\left(\sum_{t\in[t_{start},t_{end}]} T_{i,p}\,(t)\right)$

Where $T_i(P_t)$ is the active processing time for product *p* at workstation *i* during time *t*.

The Active Ratio at workstation *i* is:

$$Active\ Ratio_i=\frac{A_i}{Total\ Time_i}$$

Workstations with higher active ratios are considered potential bottlenecks.

Queue Period Method (QPM): The QPM was applied to assess the amount of time products spent in queues at various workstations. The longer products spent in the queue, the more likely it was that the workstation was a bottleneck. Through QPM analysis, several workstations were flagged as potential bottlenecks, particularly where large queues built up and caused delays in the flow of products. These findings were consistent with observations made using APM, indicating that the identified workstations were indeed experiencing production delays.

The queue time Q_i at each workstation i is given by

$$Q_i=\sum_{p\in P}\left(\sum_{q\in Q_i} t_q\right)$$

Where t_q is the time product *p* spends in queue at workstation *i*, Q_i is the queue at workstation *i*.

The queue length QL_i at each workstation iii can be calculated using a queueing theory model:

$$QL_i=\frac{\lambda_i\cdot W_q}{1-\rho_i}$$

λ_i is the arrival rate of products to workstation *i*

W_q is the average waiting time in the queue,

P_i is the utilization factor for workstation *i*

The application of the MI-based bottleneck detection methodology to the case study at Bosch Thermotechnology yielded valuable insights into the production process. The results confirmed that, even with minimal data, it was possible to reliably identify bottlenecks using the QDG and the combination of APM and QPM. In particular, the QDG provided a clear visualization of the product flow, which helped pinpoint the workstations causing delays. The integration of APM and QPM further enhanced the analysis, providing a comprehensive understanding of both the active processing times and the queuing behavior at each workstation.

5 CONCLUSION

This study demonstrates the feasibility and effectiveness of using a Minimal Information (MI) approach for real-time bottleneck detection in manufacturing systems, with a case study at Bosch Thermotechnology showing promising results. By relying on minimal data—product IDs, workstation IDs, and departure timestamps—this methodology successfully identifies bottlenecks without the need for extensive data. However, challenges such as missing data and the inability to directly analyze underlying causes of delays were encountered. Future research could improve real-time adaptability, integrate predictive analytics, and expand the methodology to more complex manufacturing environments. Additionally, incorporating feedback loops and optimization techniques could enhance its utility, making it a powerful tool for real-time production flow optimization.

ACKNOWLEDGMENTS

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions.

FUNDING

Not applicable.

INSTITUTIONAL REVIEW BOARD STATEMENT

Not applicable.

INFORMED CONSENT STATEMENT

Not applicable.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

AUTHOR CONTRIBUTIONS

Not applicable.

ABOUT THE AUTHORS

FRANK, Leede

University of California-Berkeley, 94720, US.

REFERENCES

[1] Subramaniyan, M., Skoogh, A., Salomonsson, H., Bangalore, P., & Bokrantz, J. (2018). A data-driven algorithm to predict throughput bottlenecks in a production system based on active periods of the machines. Computers & Industrial Engineering, 125, 533-544.

[2] Chen, Y. (2025). Leveraging LSTM Networks for Vehicle Stability Prediction: A Comparative Analysis with Traditional Models under Dynamic Load Conditions. Computing and Interdisciplinary Science, 1(2), 15-22.

[3] Wang, H. (2022). Supervised Learning for Complex Data (Doctoral dissertation, The University of North Carolina at Chapel Hill).

[4] Wang, H., Li, Q., & Liu, Y. (2024). Multi-response Regression for Block-missing Multi-modal Data without Imputation. Statistica Sinica, 34(2), 527.

[5] Yin, M. (2025). Predictive Maintenance of Semiconductor Equipment Using Stacking Classifiers and Explainable AI: A Synthetic Data Approach for Fault Detection and Severity Classification. Journal of Industrial Engineering and Applied Science, 3(6), 36-46.

[6] Sun, Y., & Ortiz, J. (2024). An ai-based system utilizing iot-enabled ambient sensors and llms for complex activity tracking. arXiv preprint arXiv:2407.02606.

[7] Yin, M. (2025). Data Quality Control in Semiconductor Manufacturing through Automated ETL Processes and Class Imbalance Handling Techniques. Journal of Industrial Engineering and Applied Science, 3(6), 13-22.

[8] Yin, M. (2025). A Data-Driven Approach for Real-Time Bottleneck Detection and Optimization in Semiconductor Manufacturing Using Active Period Method and Visualization. Academic Journal of Natural Science, 2(4), 19-26.

[9] Yin, M. (2025). Robust Bilevel Network-Flow Scheduling for Semiconductor Wafer Logistics under WLTP Uncertainty.

[10] Huang, S. (2025). Reinforcement Learning with Reward Shaping for Last-Mile Delivery Dispatch Efficiency. European Journal of Business, Economics & Management, 1(4), 122-130.

[11] Chen, Y. (2025). Interpretable Automated Machine Learning for Asset Pricing in US Capital Markets. Journal of Economic Theory and Business Management, 2(5), 15-21.

[12] Rudnitckaia, J., Venkatachalam, H. S., Essmann, R., Hruška, T., & Colombo, A. W. (2022). Screening process

mining and value stream techniques on industrial manufacturing processes: process modelling and bottleneck analysis. IEEE Access, 10, 24203-24214.

[13] Huang, S. (2025). Prophet with Exogenous Variables for Procurement Demand Prediction under Market Volatility. Journal of Computer Technology and Applied Mathematics, 2(6), 15-20.

[14] Cao, S., Wang, J., & Tse, T. K. (2023). Life‐cycle cost analysis and life‐cycle assessment of the second‐generation benchmark building subject to typhoon wind loads in Hong Kong. The Structural Design of Tall and Special Buildings, 32(11-12), e2014.

[15] Pang, F. (2020, November). Research on Incentive Mechanism of Teamwork Based on Unfairness Aversion Preference Model. In 2020 2nd International Conference on Economic Management and Model Engineering (ICEMME) (pp. 944-948). IEEE.

[16] Wang1a, J., Tim, K. T., Li, S., Chan1b, T. K., & Fung3c, J. C. (2023). A systematic comparison of the wind profile codifications in the Western Pacific Region. Wind and Structures, 37(2), 105-115.

[17] Chen, Y. (2025). Generative Diffusion Models for Option Pricing: A Novel Framework for Modeling Volatility Dynamics in US Financial Markets. Journal of Industrial Engineering and Applied Science, 3(6), 23-29.

[18] Wang, J., Cao, S., Tim, K. T., Li, S., Fung, J. C., & Li, Y. (2025). A novel life-cycle analysis framework to assess the performances of tall buildings considering the climate change. Engineering Structures, 323, 119258.

[19] Chen, Y. (2025). A Comparative Study of Machine Learning Models for Credit Card Fraud Detection. Academic Journal of Natural Science, 2(4), 11-18.

[20] Luo, M., Du, B., Zhang, W., Song, T., Li, K., Zhu, H., ... & Wen, H. (2023). Fleet rebalancing for expanding shared e-Mobility systems: A multi-agent deep reinforcement learning approach. IEEE Transactions on Intelligent Transportation Systems, 24(4), 3868-3881.

[21] Wang, J., Tse, T. K., Li, S., & Fung, J. C. (2023). A model of the sea–land transition of the mean wind profile in the tropical cyclone boundary layer considering climate changes. International Journal of Disaster Risk Science, 14(3), 413-427.

[22] Zhang, Z., Li, S., Zhang, Z., Liu, X., Jiang, H., Tang, X., ... & Jiang, M. (2025). IHEval: Evaluating language models on following the instruction hierarchy. arXiv preprint arXiv:2502.08745.

[23] Lee, J. Y. J., Bonab, H., Zalmout, N., Zeng, M., Lokegaonkar, S., Lockard, C., ... & Wang, H. (2025, August). DocTalk: Scalable graph-based dialogue synthesis for enhancing LLM conversational capabilities. In Proceedings of the 26th Annual Meeting of the Special Interest Group on Discourse and Dialogue (pp. 658-677).

[24] Luo, M., Zhang, W., Song, T., Li, K., Zhu, H., Du, B., & Wen, H. (2021, January). Rebalancing expanding EV sharing systems with deep reinforcement learning. In Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence (pp. 1338-1344).

[25] Zhang, K. (2025). Supply Chain Intelligent Rating System: A Study on Its Adaptability and Application in Multiple Industries. Journal of World Economy, 4(5), 20-26.

[26] Ding, Q. (2025). Development and Die-Cutting Process Optimization of High-Temperature-Resistant and Anti-Aging Barcode Substrate for Photovoltaic Modules. Innovation in Science and Technology, 4(8), 22-28.

[27] Wang, J., Golnary, F., Li, S., Weerasuriya, A. U., & Tse, K. T. (2024). A review on power control of wind turbines with the perspective of dynamic load mitigation. Ocean Engineering, 311, 118806.

[28] Jieyang, P., Kimmig, A., Dongkun, W., Niu, Z., Zhi, F., Jiahai, W., ... & Ovtcharova, J. (2023). A systematic review of data-driven approaches to fault diagnosis and early warning. Journal of Intelligent Manufacturing, 34(8), 3277-3304.