Computer Modeling in Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.085541

ARTICLE

System Reliability Evaluation for Semiconductor Manufacturing Systems Considering Workload and Unscheduled Downtime

Cheng-Ta Yeh[1], Yu-Zhong Kang[2], Louis Cheng-Lu Yeng[2] and Yi-Kuei Lin[2,3,4,5,*]

[1]Department of Information Management, National Taipei University of Business, Taipei City, Taiwan
[2]Department of Industrial Engineering and Management, National Yang Ming Chiao Tung University, Hsinchu, Taiwan
[3]Department of Industrial Engineering and Management, National Chin-Yi University of Technology, Taichung, Taiwan
[4]Department of Business Administration, Asia University, Taichung, Taiwan
[5]Department of Industrial Engineering and Management, Chaoyang University of Technology, Taichung, Taiwan

*Corresponding Author: Yi-Kuei Lin. Email: yklin@nycu.edu.tw

Received: 13 May 2026; Accepted: 16 October 2026; Published: 17 November 2026

ABSTRACT: Rapid advances in artificial intelligence and high-performance computing have increased demand for semiconductors, placing fabrication plants under growing pressure to meet order commitments under uncertain operating conditions. In practice, production planners must evaluate whether a manufacturing system can complete a given customer order pattern within a specified delivery time threshold. However, semiconductor fabrication systems involve reentrant process flows, heterogeneous product routings, sampling operations, and rework flows, while workstation capacities vary because of unscheduled downtime and machine unavailability. These characteristics make deterministic capacity evaluation insufficient for assessing order fulfillment capability. To address this issue, this study models a semiconductor manufacturing system as a multistate semiconductor manufacturing network (MSMN) and proposes a workload-based reliability evaluation procedure. System reliability is defined as the probability that the MSMN can complete the required order pattern within the delivery time threshold. The proposed approach derives the minimal machine number vector required to satisfy the aggregate workload requirements and evaluates reliability under stochastic workstation capacities. The numerical results demonstrate that the proposed framework provides a probabilistic basis for order fulfillment evaluation and supports delivery time assessment and capacity expansion decisions in semiconductor manufacturing systems.

KEYWORDS: Multistate semiconductor manufacturing network; system reliability; workload-based evaluation; reentrant process flow; unscheduled downtime

1 Introduction

In recent years, the emergence of hyperscale artificial intelligence applications has substantially increased the demand for high-performance computing and high-performance processors [1,2]. According to the World Semiconductor Trade Statistics (WSTS), global semiconductor sales reached USD 795.6 billion in 2025, and the market is expected to continue growing toward the USD 1 trillion level in 2026.

Driven by increasing demand, semiconductor fabrication plants must balance high capacity utilization with strict delivery constraints, making on-time delivery a key performance indicator for production planners [3]. Achieving this objective depends on the system's ability to meet the required demand within the specified delivery time threshold. However, evaluating such performance is challenging due to the inherent complexity of semiconductor manufacturing systems. These systems are characterized by

Copyright © 2026 The Authors. Published by Tech Science Press. This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

reentrant processes, numerous processing steps, and heterogeneous product routings. In addition, sampling and rework operations are frequently embedded across production stages, further complicating material flows [4]. From an operational perspective, capacity uncertainty caused by unscheduled downtime (UDT), maintenance, and partial machine unavailability leads to time-varying and state-dependent workstation capacities, represented by the number of operational machines at each workstation [5]. These structural and stochastic characteristics jointly make system-level performance evaluation highly challenging.

To support planning and decision-making under such conditions, it is essential to adopt modeling frameworks capable of capturing both system structure and capacity uncertainty. The reentrant structure, complex routing patterns, and interdependent material flows in semiconductor manufacturing systems naturally lend themselves to a network-based representation, in which nodes and arcs can explicitly represent processing relationships, flow dependencies, and resource interactions across the system [6]. Network-based models provide a structured analytical representation for linking system topology, capacity states, and performance requirements. Accordingly, such approaches have been widely applied to system-level performance evaluation in various domains, including supply chains, transportation, and distribution systems [7–9].

In manufacturing contexts, network models have been further extended to account for stochastic workstation capacities, giving rise to multistate manufacturing network (MMN) formulations [6]. In these models, each workstation typically consists of multiple parallel machines, whose operational states may vary due to failures, maintenance, or partial degradation. Therefore, the effective capacity of a workstation is represented as a discrete random variable, thereby allowing a realistic characterization of manufacturing systems subject to capacity uncertainty. Within MMN frameworks, system reliability is commonly defined as the probability that the system can satisfy specified operational requirements, such as demand fulfillment, delivery time requirements, or resource limitations [10–12]. Compared with local performance measures that focus on individual machines or workstations, system reliability provides a system-level indicator that simultaneously considers network topology and stochastic workstation capacity [8,13]. As a result, reliability-based evaluation has been widely applied to manufacturing systems with rework, buffer interactions, and stochastic processing conditions, as well as to infrastructure and logistics networks operating under uncertainty [6,9,14].

Despite these advances, existing studies primarily focus on manufacturing systems with simplified or acyclic process structures and do not explicitly address the reentrant characteristics inherent in semiconductor manufacturing. Reentrant flows result in repeated visits to the same workstations and introduce cyclic dependencies in material flows, significantly increasing the complexity of system modeling and performance evaluation [15]. Moreover, prior studies [10,15,16] primarily evaluate system feasibility based on minimal-path-based formulations, which determine whether system capacity configurations can satisfy demand. However, these approaches do not explicitly convert the order pattern and delivery time threshold into workload requirements. This limitation restricts their applicability to semiconductor manufacturing environments, where order fulfillment is jointly determined by reentrant processing, heterogeneous product routings, sampling and rework operations, and stochastic workstation capacities.

To address these limitations, this study proposes a reliability evaluation framework for semiconductor manufacturing systems that explicitly considers reentrant process flows, heterogeneous product routings, sampling operations, rework flows, and stochastic workstation capacities. Specifically, the semiconductor manufacturing system is modeled as a multistate semiconductor manufacturing network (MSMN), and a workload-based analytical approach is developed to evaluate system reliability. In this study, system reliability is defined as the probability that the MSMN can complete a given order pattern within the specified delivery time threshold. This reliability measure is intended for tactical capacity planning and

delivery time feasibility assessment rather than detailed operational scheduling. The proposed approach derives the minimal machine number vector required to satisfy the aggregate workload requirements of all workstations and then evaluates the corresponding reliability under stochastic workstation capacities. Such a formulation provides a more realistic and practically relevant representation of order fulfillment performance in semiconductor manufacturing environments. This integration of workload conversion and multistate reliability evaluation distinguishes the proposed framework from conventional MMN reliability models, delivery time reliability models, and workload-based capacity planning approaches.

The remainder of this paper is organized as follows. Section 2 reviews related studies. Section 3 develops the MSMN-based reliability evaluation model. Section 4 presents the proposed workload-based reliability evaluation procedure. Section 5 provides a numerical case study of semiconductor manufacturing and discusses the managerial implications. Finally, Section 6 concludes the study.

2 Literature Review

Network-based modeling provides an analytical approach for representing complex systems with interdependent components and uncertain operating conditions. By abstracting system components as nodes and their interactions as arcs, network representations provide a structured description of system topology and facilitate system-level performance evaluation. Such approaches have been extensively applied in various domains, including supply chain systems [9,12], transportation systems [8,14,17], distribution systems [7,13], and electricity systems [18]. In manufacturing contexts, network representations are commonly used to model production flows, in which nodes typically represent workstations and arcs denote stockers or conveyors [16,19]. These models provide a flexible framework for analyzing manufacturing system structure and performance under uncertainty.

A critical aspect of manufacturing system modeling is the representation of workstation capacity. In practical production environments, workstation capacity is uncertain because machines may become unavailable due to failures, maintenance, or operational disruptions. Rather than assuming a fixed capacity level, many studies [6,10,15,20] represent workstation capacity based on the number and operating states of machines available at each workstation. This leads to the concept of an MMN, in which each node is associated with multiple capacity states and corresponding probabilities. Such modeling frameworks have been applied to evaluate manufacturing systems with rework, buffer interactions, and stochastic processing conditions, demonstrating their effectiveness in capturing the probabilistic nature of manufacturing operations [16,21].

Based on the MMN framework, system reliability has been widely adopted as a key performance indicator for evaluating manufacturing systems. In general, system reliability is defined as the probability that a system can satisfy specified performance requirements, such as meeting demand levels or completing production under given constraints [22,23]. In manufacturing systems, reliability evaluation is typically conducted using network-based approaches that assess whether the available capacity configuration can support the required production performance. Complementing the scheduling-oriented performance assessments in studies [24,25], system reliability integrates network structure and stochastic workstation capacities to quantify the probability of satisfying specified production requirements.

Recent studies have further emphasized the importance of reliability assessment for manufacturing systems from broader planning and decision-making perspectives. Friederich and Lazarova-Molnar [26] reviewed reliability assessment in manufacturing systems and identified several challenges, such as data availability, increasing system complexity, and the integration of different reliability assessment methods. Zhang et al. [27] evaluated due-date reliability for production systems with finite buffers under time-dependent and operation-dependent failures, highlighting the relevance of reliability evaluation to demand satisfaction

and due-time performance. In addition, Renna [28] reviewed game-theory-based decision models for production planning, scheduling, cloud manufacturing, and sustainable production systems, showing the importance of decision-support models in uncertain manufacturing environments. These studies indicate that manufacturing reliability evaluation should not only capture stochastic equipment behavior but also support production planning and delivery-time decision-making under complex system conditions.

Despite these advances, existing studies on manufacturing system reliability have primarily been developed for systems with relatively simple structures, simplified routing assumptions, or deterministic processing requirements. However, semiconductor manufacturing systems exhibit several distinctive characteristics that significantly increase modeling complexity. First, semiconductor manufacturing involves reentrant process flows, in which products may revisit the same workstation multiple times due to process requirements [29,30]. In addition, sampling operations are embedded for quality inspection, and nonconforming units may be routed back to designated processing steps for rework, further complicating material flows [4]. Second, different product types often follow heterogeneous processing routes, leading to uneven workload distribution across the system [5]. Finally, workstation capacity is stochastic because machines may become unavailable due to UDT, maintenance, or partial machine unavailability [5].

These characteristics imply that system performance in semiconductor manufacturing is jointly affected by reentrant process flows, heterogeneous product routings, sampling and rework operations, and stochastic workstation capacities. However, existing manufacturing network reliability studies have rarely considered these factors simultaneously. More importantly, many existing approaches evaluate system feasibility mainly by determining whether a capacity configuration can satisfy demand, without explicitly converting the order pattern and delivery time threshold into workstation-level workload requirements. Different from finite-buffer due-date reliability models and workload-based capacity planning approaches, the proposed framework focuses on reentrant semiconductor manufacturing networks and further incorporates stochastic workstation capacities to evaluate the probability of order fulfillment. Therefore, there remains a need for a workload-based reliability evaluation framework that integrates these characteristics, links order patterns and delivery time thresholds to workstation-level workload requirements, and supports delivery time assessment in semiconductor manufacturing systems.

3 MSMN-Based Reliability Evaluation Model

This section develops the MSMN-based reliability evaluation model. Section 3.1 defines the MSMN structure and the model assumptions. Section 3.2 formulates the probability distribution of operational machines at each workstation. Section 3.3 presents the reliability evaluation framework based on the minimal feasible machine number vector required to complete the order pattern within the delivery time threshold. Sections 3.4 and 3.5 formulate the operation and workload requirements, respectively. Finally, Section 3.6 derives the minimal machine numbers for all workstations.

3.1 MSMN Structure and Model Assumptions

Let $\mathbf{G} \equiv (\mathbf{A}, \mathbf{V}_N, \mathbf{V}_M, \boldsymbol{W})$ be an MSMN structure as shown in Fig. 1, where $\mathbf{A}$ represents the set of stockers (arcs), serving as temporary buffers between workstations. $\mathbf{V}_N = \{v_g \mid g = 1, 2, \ldots, n\}$ denotes the set of n processing workstations (nodes), which are dedicated to the physical fabrication and structural transformation of wafers. $\mathbf{V}_M = \{v_g \mid g = n+1, n+2, \ldots, n+m\}$ denotes the set of m metrology workstations (nodes), which perform measurement and inspection for quality control throughout the production process. The metrology workstations perform either sampling or full inspection. Nonconforming units are then routed to designated workstations for rework. $\boldsymbol{W} = (w_1, w_2, \ldots, w_{n+m})$ is a maximum machine number vector, where w_g denotes the maximum number of machines available at workstation v_g.

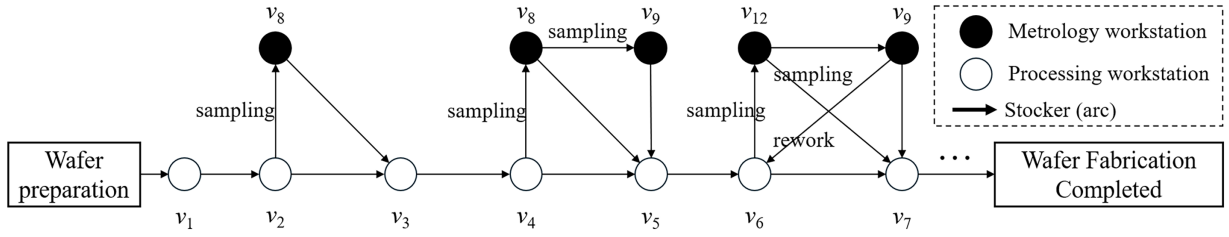

Figure 1: Schematic diagram of the MSMN topology.

The MSMN processes k product types. The number of processing steps for product type i is denoted by s_i for $i = 1, 2, \ldots, k$. The symbol $l_{i,j} \in \{1, 2, \ldots, n + m\}$ denotes the index of the workstation assigned to step j of product type i. Reentrant flows are represented by allowing the same workstation index to appear multiple times in the route of a product type. That is, if $l_{i,j} = l_{i,j'}$ for $j \neq j'$, then product type i revisits the same workstation at different processing steps. Let $u_{i,j}$ denote the processing unit size of step j for product type i, representing the number of wafers processed in one processing operation at the assigned workstation $v_{l_{i,j}}$. Depending on the operation, the processing unit may correspond to a single wafer, a lot, or a batch. The sampling rate, denoted by $\alpha_{i,j}$, represents the proportion of units of product type i selected for inspection or measurement at inspection step j. The rework rate, denoted by $\beta_{i,j}$, represents the proportion of inspected units of product type i that fail inspection and require rework.

Let $D = (d_1, d_2, \ldots, d_k)$ denote a customer order pattern, where d_i represents the demand for product type i, $i = 1, 2, \ldots, k$. The symbol T represents the delivery time threshold. To evaluate system reliability, denoted by $R_{D,T}$, the MSMN model is based on the following assumptions:

I. Each stocker (arc) is perfectly reliable.
II. Nonconforming units identified at metrology workstations can be reworked at most once, and rework is always successful.
III. Within each workstation, machines are assumed to operate independently and have a common availability; the numbers of operational machines at different workstations are statistically independent.

These assumptions are commonly adopted in manufacturing network reliability evaluation to support tractable analytical assessment [10,16,19]. Their implications should be interpreted according to the planning-level scope of the proposed model. Specifically, the assumption of perfectly reliable stockers allows the model to focus on workstation-capacity uncertainty caused by UDT. The assumption of at-most-once, always-successful rework allows the rework quantity to be represented by its expected value without modeling repeated rework cycles. Machines within each workstation are sourced from the same supplier and are modeled with a common availability and independent operating states; operational-machine counts are also assumed independent across workstations, allowing their reliability contributions to be combined analytically. In practice, stocker failures, repeated rework, common-cause failures, or shared maintenance resources may violate these assumptions. Therefore, the proposed model is intended to support tactical capacity planning and delivery time feasibility assessment, while detailed shop-floor dynamics, such as dispatching rules, queueing delays, WIP congestion, batching synchronization, and exact downtime processes, are beyond the scope of this study.

3.2 Probability Distribution of Operational Machines at Each Workstation

Machines in a workstation may become unavailable due to UDT, which can be caused by factors such as component wear, electrical short circuits, or friction overheating. The availability of a machine, denoted by r_g, is defined as the probability that the machine is operational (i.e., not in UDT). This availability can

be estimated using the mean time between failures (MTBF) and the mean time to repair (MTTR) [31], as follows:

$$r_g = \frac{MTBF_g}{MTBF_g + MTTR_g}, \text{for } g = 1, 2, \ldots, n+m \tag{1}$$

where $MTBF_g$ and $MTTR_g$ denote the mean time between failures and the mean time to repair for workstation v_g, respectively.

Let $X = (x_1, x_2, \ldots, x_{n+m})$ be the machine number (capacity) vector, where x_g represents the number of operational machines at workstation v_g and takes values in $\{0, 1, 2, \ldots, w_g\}$. Because the machines within workstation v_g are sourced from the same supplier, they are assigned the same availability r_g and are assumed to operate independently. Therefore, the number of operational machines at workstation v_g, denoted by x_g, follows a binomial distribution, and the probability that $x_g = z$ is given by:

$$\Pr\left(x_g = z\right) = \binom{w_g}{z} \left(r_g\right)^z \left(1 - r_g\right)^{w_g - z}, \text{for } g = 1, 2, \ldots, n+m \tag{2}$$

For illustration, suppose that workstation v_1 consists of three machines (i.e., $w_1 = 3$). Each machine is assumed to have an MTBF of 1 week (10,080 min) and an MTTR of 2 h (120 min). According to Eq. (1), the machine availability r_1 is calculated as $MTBF_1/(MTBF_1 + MTTR_1) = 0.9882353$. The probability distribution of x_1, obtained from Eq. (2), is summarized in Table 1.

Table 1: Distribution of operational machines at workstation v_1.

Workstation v_1	Operational Machines	Probability Pr (x_1)
v_1	3	0.965119
	2	0.034469
	1	0.000410
	0	0.000002

The availability-based formulation should be interpreted as a planning-level representation of workstation capacity. The value r_g is an average machine availability estimated from MTBF and MTTR, and Eq. (2) gives the probability distribution of the number of operational machines at a workstation. This formulation does not track the occurrence time of each UDT event or the corresponding repair duration within the delivery time threshold.

3.3 System Reliability Evaluation Using Minimal Machine Number Vectors

System reliability $R_{D,T}$ is defined as the probability that the MSMN can successfully complete the order pattern D within the delivery time threshold T. Let $\mathbf{\Omega}_{D,T}$ be the set of machine number vectors that can fulfill the order pattern D within the delivery time threshold T. Then, the system reliability is given by

$$R_{D,T} = \sum_{X \in \mathbf{\Omega}_{D,T}} \Pr(X) \tag{3}$$

where $\Pr(X) = \Pr(x_1) \times \Pr(x_2) \times \ldots \times \Pr(x_{m+n})$ due to Assumption III.

However, the number of machine number vectors in $\mathbf{\Omega}_{D,T}$ grows rapidly as the numbers of workstations and machines increase, making exhaustive enumeration impractical. To address this issue, the concept of a minimal machine number vector for D and T, referred to as a (D, T)-MMNV, is introduced.

Definition 1: *A (D, T)-MMNV is the minimal element in the set* $\mathbf{\Omega}_{D,T}$. *That is, if X is a (D, T)-MMNV, then any machine number vector* $Y = (y_1, y_2, \ldots, y_{n+m})$ *such that* $Y < X$ *is infeasible for satisfying the order pattern D within T, where* $Y \leq X$ *means* $y_g \leq x_g$ *for all g, and* $Y < X$ *means* $Y \leq X$ *and* $y_g < x_g$ *for at least one g.*

Since any feasible machine number vector must be no smaller than the (D, T)-MMNV, the feasible region is the upper set generated by this (D, T)-MMNV. Suppose $\underline{X} = \left(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_{n+m}\right)$ is the (D, T)-MMNV. Then, the system reliability $R_{D,T}$ can be equivalently expressed as

$$R_{D,T} = \Pr\left(X \geq \underline{X}\right) = \prod_{g=1}^{n+m} \Pr\left(x_g \geq \underline{x}_g\right) \tag{4}$$

where $X \geq \underline{X}$ indicates $x_g \geq \underline{x}_g$ for $g = 1, 2, \ldots, n + m$ and $\Pr\left(x_g \geq \underline{x}_g\right)$ is calculated by

$$\Pr\left(x_g \geq \underline{x}_g\right) = \sum_{v=\underline{x}_g}^{w_g} \binom{w_g}{v} \left(r_g\right)^v \left(1 - r_g\right)^{w_g - v}, \text{ for } g = 1, 2, \ldots, n + m \tag{5}$$

The concept of (D, T)-MMNV enables a significant reduction in computational complexity by avoiding the enumeration of all feasible machine number vectors.

3.4 Calculation of Expected Operation Count

To determine the (D, T)-MMNV, the workload requirement induced by the order pattern D and the delivery time threshold T must first be quantified for each workstation. Let ρ_i denote the demand-based production requirement of product type i, defined as the average quantity that must be completed per unit time to satisfy demand d_i within T:

$$\rho_i = \frac{d_i}{T}, \text{ for } i = 1, 2, \ldots, k \tag{6}$$

Given the production requirement ρ_i, the expected operation count at step j, denoted by $e_{i,j}$, represents the expected number of operations required at the assigned workstation $v_{l_{i,j}}$. For a step assigned to a processing workstation, i.e., $v_{l_{i,j}} \in \mathbf{V}_N$, the expected operation count consists of the normal processing count required to satisfy ρ_i and the expected rework count caused by nonconforming units returned from a subsequent inspection step. The normal processing count is determined by the demand-based production requirement ρ_i and the processing unit size $u_{i,j}$. For a processing step j, let $f_{i,j}$ denote the subsequent inspection step for product type i, where a product failing inspection is sent back to processing step j for rework. The expected rework count is obtained by multiplying the demand-based production requirement by the sampling rate and rework rate at the subsequent inspection step, and then converting the resulting wafer quantity into the number of rework operations using $u_{i,j}$. Thus, the expected operation count for $v_{l_{i,j}} \in \mathbf{V}_N$ is given by:

$$e_{i,j} = \frac{\rho_i}{u_{i,j}} + \frac{\rho_i \times \alpha_{i,f_{i,j}} \times \beta_{i,f_{i,j}}}{u_{i,j}}, \text{ for } v_{l_{i,j}} \in \mathbf{V}_N, j = 1, 2, \ldots, s_i, \text{ and } i = 1, 2, \ldots, k \tag{7}$$

The first term $\frac{\rho_i}{u_{i,j}}$ represents the normal processing count required to satisfy ρ_i, whereas the second term $\frac{\rho_i \times \alpha_{i,f_{i,j}} \times \beta_{i,f_{i,j}}}{u_{i,j}}$ represents the expected rework count caused by nonconforming units identified at the subsequent inspection step $f_{i,j}$ and returned to processing step j. If no subsequent inspection step triggers rework back to step j, the second term is set to zero.

For a step assigned to a metrology workstation, i.e., $v_{l_{i,j}} \in \mathbf{V}_M$, the expected operation count represents the expected number of inspection or measurement operations generated by sampled products. Unlike processing workstations, metrology workstations do not perform normal fabrication processing or rework processing. Therefore, the expected operation count is determined by the demand-based production requirement ρ_i, the sampling rate $\alpha_{i,j}$ at the current inspection step, and the processing unit size $u_{i,j}$, as follows:

$$e_{i,j} = \frac{\rho_i \times \alpha_{i,j}}{u_{i,j}}, \text{for } v_{l_{i,j}} \in \mathbf{V}_M, j = 1, 2, \ldots, s_i, \text{and } i = 1, 2, \ldots, k \tag{8}$$

where $\alpha_{i,j} = 1$ if full inspection is performed. The expected operation counts should be interpreted as planning-level workload measures. They represent the average workload caused by sampling and rework rather than the random variation in individual sampling and rework operations.

3.5 Determination of Aggregate Workload Time

After the expected operation count $e_{i,j}$ is obtained, the workload time required by each product type at each step can be determined according to the processing mechanism of the assigned workstation. In this study, two processing mechanisms are considered: the single-machine mechanism and the cascading-machine mechanism [5]. Let $p_{i,j}$ be the unit processing time of product type i at step j and $\mathbf{V}^{Sin}$ denote the set of workstations operating under the single-machine mechanism. For a step assigned to such a workstation with $v_{l_{i,j}} \in \mathbf{V}^{Sin}$, each operation requires the full unit processing time $p_{i,j}$. Therefore, the workload time required for product type i at step j is given by:

$$t_{i,j} = p_{i,j} \times e_{i,j}, \text{for } v_{l_{i,j}} \in \mathbf{V}^{Sin}, j = 1, 2, \ldots, s_i, \text{ and } i = 1, 2, \ldots, k \tag{9}$$

Let $\mathbf{V}^{Cas}$ denote the set of workstations operating under the cascading-machine mechanism. For a step assigned to such a workstation with $v_{l_{i,j}} \in \mathbf{V}^{Cas}$, processing operations can overlap. The first group of operations requires the full unit processing time $p_{i,j}$, whereas subsequent operations are completed at the cascading interval $c_{i,j}$, the time interval between two consecutive completed units. Thus, the required workload time is given by:

$$t_{i,j} = p_{i,j} + \left(e_{i,j} - 1\right) \times c_{i,j}, \text{for } v_{l_{i,j}} \in \mathbf{V}^{Cas}, j = 1, 2, \ldots, s_i, \text{ and } i = 1, 2, \ldots, k \tag{10}$$

Let AWT_g denote the aggregate workload time required at workstation v_g. It is obtained by summing the workload time requirements of all product-step pairs assigned to workstation v_g:

$$AWT_g = \sum_{\substack{i=1 \\ I_{i,j}=g}}^{k} \sum_{j=1}^{s_i} t_{i,j}, \text{for } g = 1, 2, \ldots, n + m \tag{11}$$

3.6 Derivation of the Minimal Machine Number Vector for D and T

Let IT denote the ideal operational time for all workstations, representing the maximum available production time within the planning horizon when no downtime is considered. Let PMT_g be the average preventive maintenance time (PMT) of workstation v_g, which is estimated from historical maintenance records during the same planning horizon. After accounting for preventive maintenance, the available operational time for one machine at workstation v_g, denoted by OT_g, is determined as

$$OT_g = IT - PMT_g, \text{for } g = 1, 2, \ldots, n + m \tag{12}$$

This value represents the effective operational time available for one machine when determining the minimal machine number. For example, if the planning horizon is one week, then $IT = 7 \times 24 \times 60 = 10{,}080$ min. If $PMT_g = 480$ min per week, then the available operational time of one machine at workstation v_g becomes $OT_g = 10{,}080 - 480 = 9600$ min.

Once the order pattern, product routes, sampling rates, rework rates, and processing mechanisms are specified, the aggregate workload time is fixed for each workstation. Therefore, the feasibility of a machine number vector X can be evaluated based on whether the available operational time at each workstation is sufficient to complete its aggregate workload. This feasibility condition is defined from a planning-level workload perspective and does not model detailed shop-floor dynamics. Under this workload-based feasibility definition, product routes are fixed, and no alternative routing or dispatching decisions are considered. The following lemma shows an important property of the (D, T)-MMNV.

Lemma 1: *If the machine number vector $X = (x_1, x_2, \ldots, x_{n+m})$ is generated by the following equation, then X is the (D, T)-MMNV.*

$$x_g = \left\lceil \frac{AWT_g}{OT_g} \right\rceil, \text{ for } g = 1, 2, \ldots, n + m \tag{13}$$

Proof of Lemma 1: Since X is derived from Eq. (13), $x_g \times OT_g \geq AWT_g$ for $g = 1, 2, \ldots, n + m$. This means that the available operational time at each workstation is sufficient to complete its aggregate workload. Hence, $X \in \mathbf{\Omega}_{D,T}$. Subsequently, suppose that such X is not the (D, T)-MMNV (i.e., X is not the minimal element in $\mathbf{\Omega}_{D,T}$). Then, there exists the (D, T)-MMNV $Y = (y_1, y_2, \ldots, y_{n+m})$ (i.e., $Y \in \mathbf{\Omega}_{D,T}$) such that $Y < X$ (i.e., $y_h < x_h$ and $y_g \leq x_g$ for all $g \neq h$). Since $x_h = \left\lceil \frac{AWT_h}{OT_h} \right\rceil$, x_h is the smallest integer satisfying $x_h \times OT_h \geq AWT_h$. Therefore, for $y_h < x_h$, we have $y_h \times OT_h < AWT_h$. This means that workstation v_h does not have sufficient available operational time to complete its aggregate workload, so $Y \notin \mathbf{\Omega}_{D,T}$. This contradicts $Y \in \mathbf{\Omega}_{D,T}$. Therefore, no feasible machine number vector smaller than X exists. Since $X \in \mathbf{\Omega}_{D,T}$ and no $Y \in \mathbf{\Omega}_{D,T}$ satisfies $Y < X$, X is a minimal element in $\mathbf{\Omega}_{D,T}$. Hence, X is the (D, T)-MMNV. □

It should be noted that if $x_g > w_g$ for any workstation v_g, the required minimal machine number exceeds the maximum available number of machines at that workstation. In this case, no feasible machine number vector exists under the current system configuration, and $R_{D,T} = 0$.

4 Workload-Based Reliability Evaluation Procedure

Based on the MSMN-based reliability evaluation model developed in Section 3, a workload-based reliability evaluation procedure is presented. The detailed procedure is described as follows.

Workload-based reliability evaluation procedure: Evaluate the system reliability $R_{D,T}$.

Input:

(1) **Network structure: $\mathbf{G} \equiv (\mathbf{A}, \mathbf{V}_N, \mathbf{V}_M, W)$.**

(2) **Process parameters:** number of product types k, number of processing steps s_i of product type i, index $l_{i,j}$ of the workstation assigned to step j of product type i, processing unit size $u_{i,j}$, sampling rate $\alpha_{i,j}$, rework rate $\beta_{i,j}$, unit processing time $p_{i,j}$, and cascading interval $c_{i,j}$, in which $j = 1, 2, \ldots, s_i$ and $i = 1, 2, \ldots, k$.

(3) **Workstation parameters:** $MTBF_g$, $MTTR_g$, set of single-machine workstations V^{Sin}, set of cascading workstations V^{Cas}, and ideal operational time IT, PMT_g, where $g = 1, 2, \ldots, n + m$.

(4) **Order pattern *D* and delivery time threshold *T*.**

Step 1. Calculate the demand-based production requirement ρ_i for each product type as follows:

$$\rho_i = \frac{d_i}{T}, \text{for } i = 1, 2, \ldots, k \tag{14}$$

Step 2. Determine the expected operation count $e_{i,j}$ using Eq. (15) for processing workstations and Eq. (16) for metrology workstations.

$$e_{i,j} = \frac{\rho_i}{u_{i,j}} + \frac{\rho_i \times \alpha_{i,f_{i,j}} \times \beta_{i,f_{i,j}}}{u_{i,j}}, \text{for } v_{l_{i,j}} \in \mathbf{V}_N, j = 1, 2, \ldots, s_i, \text{and } i = 1, 2, \ldots, k \tag{15}$$

$$e_{i,j} = \frac{\rho_i \times \alpha_{i,j}}{u_{i,j}}, \text{for } v_{l_{i,j}} \in \mathbf{V}_M, j = 1, 2, \ldots, s_i, \text{and } i = 1, 2, \ldots, k \tag{16}$$

Step 3. Obtain the aggregate workload time for each workstation.

Step (3.1) Calculate the workload time $t_{i,j}$ for product type i at step j assigned to a single-machine workstation $v_{l_{i,j}}$ using Eq. (17).

$$t_{i,j} = p_{i,j} \times e_{i,j}, \text{for } v_{l_{i,j}} \in \mathbf{V}^{Sin}, j = 1, 2, \ldots, s_i, \text{and } i = 1, 2, \ldots, k \tag{17}$$

Step (3.2) Calculate the workload time $t_{i,j}$ for product type i at step j assigned to a cascading-machine workstation $v_{l_{i,j}}$ using Eq. (18).

$$t_{i,j} = p_{i,j} + \left(e_{i,j} - 1\right) \times c_{i,j}, \text{for } v_{l_{i,j}} \in \mathbf{V}^{Cas}, j = 1, 2, \ldots, s_i, \text{and } i = 1, 2, \ldots, k \tag{18}$$

Step (3.3) Calculate the aggregate workload time AWT_g for each workstation v_g.

$$AWT_g = \sum_{i=1}^{k} \sum_{\substack{j=1 \\ I_{i,j}=g}}^{s_i} t_{i,j}, \text{for } g = 1, 2, \ldots, n + m \tag{19}$$

Step 4. Calculate the available operational time OT_g for a single machine at workstation v_g.

$$OT_g = IT - PMT_g, \text{ for } g = 1, 2, \ldots, \ n + m \tag{20}$$

Step 5. Obtain the (D, T)-MMNV $\underline{X} = \left(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_{n+m}\right)$ using Eq. (21).

$$\underline{x}_g = \left\lceil \frac{AWT_g}{OT_g} \right\rceil, \text{for } g = 1, 2, \ldots, n + m \tag{21}$$

If there exists a workstation v_g such that $x_g > w_g$, no feasible machine number vector exists under the current system configuration, and the system reliability is set to $R_{D,T} = 0$. Otherwise, the procedure continues to Step 6.

Step 6. Calculate machine availability r_g for each workstation via the following equation.

$$r_g = \frac{MTBF_g}{MTBF_g + MTTR_g}, \text{for } g = 1, 2, \ldots, n + m \tag{22}$$

Step 7. Calculate the system reliability $R_{D,T}$ as follows:

$$R_{D,T} = \Pr\left(X \geq \underline{X}\right) = \prod_{g=1}^{n+m} \Pr\left(x_g \geq \underline{x}_g\right) \tag{23}$$

Output: System reliability $R_{D,T}$.

The computational complexity of the proposed procedure can be analyzed in terms of the numbers of processing steps and workstations. Since there are $\sum_{i=1}^{k} s_i$ processing steps across all product types, Steps 1–3 require $O(\sum_{i=1}^{k} s_i)$ operations to calculate the demand-based production requirement, expected operation counts, and aggregate workload time. Steps 4–6 each require $O(n + m)$ operations, so their combined complexity is $O(n + m)$. In Step 7, the probability that each workstation has at least the required number of operational machines is calculated by summing the binomial probabilities from the required machine number to the maximum machine number. This calculation requires $O(\sum_{g=1}^{n+m} w_g)$ operations. The workstation probabilities are then multiplied to obtain the system reliability, which requires $O(n + m)$ operations. Thus, Step 7 requires $O(\sum_{g=1}^{n+m} w_g + n + m)$ operations, which can be simplified as $O(\sum_{g=1}^{n+m} w_g)$. Overall, the time complexity of the procedure is $O(\sum_{i=1}^{k} s_i + (n + m) + \sum_{g=1}^{n+m} w_g)$, which can be simplified as $O(\sum_{i=1}^{k} s_i + \sum_{g=1}^{n+m} w_g)$.

The proposed procedure avoids enumerating all possible machine number vectors. A direct enumeration approach would need to examine up to $\prod_{g=1}^{n+m} \left(w_g + 1\right)$ possible machine number vectors, and this number increases rapidly when the number of workstations is large. In contrast, the proposed procedure first derives the (D, T)-MMNV from the aggregate workload time and available operational time, and then calculates the probability that the machine number vector is no smaller than this minimal vector. Thus, in terms of the number of required calculations, the proposed procedure grows additively with the number of processing steps and workstation machine states, rather than multiplicatively with all possible machine number vectors.

5 Numerical Case Study

This section presents a case study based on Dataset 1 from Kopp et al. [5]. The processing data, sampling and rework data, and workstation data from Dataset 1 are used to describe the semiconductor manufacturing system. The order pattern and delivery time threshold are specified for the numerical case study. Section 5.2 illustrates the application of the proposed procedure, Section 5.3 compares the system reliability obtained by the proposed procedure with an estimate obtained by the simulation-based reliability estimation method, and Section 5.4 examines the managerial implications of system reliability evaluation. All computational experiments are conducted in Google Colaboratory using Python 3.13.15 on a Linux runtime with two logical Intel Xeon CPU cores at 2.00 GHz and 12.67 GB of RAM.

5.1 Description of the Case Study

This subsection describes the semiconductor manufacturing system and the data used in the numerical analysis. Dataset 1 is selected because it provides a publicly available and reproducible high-volume low-mix semiconductor manufacturing setting with two product types, long reentrant processing routes, 105 workstations, different processing mechanisms, and scheduled and unscheduled downtime characteristics. Dataset 1 is used to illustrate the proposed procedure rather than to calibrate it. The procedure contains no parameters fitted to Dataset 1 and can be applied to another semiconductor manufacturing system by replacing the corresponding product-route and workstation data, provided that the stated assumptions are satisfied.

The fabrication plant produces two product types (i.e., $k = 2$), each following a distinct manufacturing route. Product types 1 and 2 involve 583 and 343 manufacturing steps, respectively. Each step is assigned to one of 105 workstations organized into 11 functional areas. Specifically, the system contains 92 processing workstations grouped into eight processing functional areas: diffusion (v_1 to v_{10}), wet etch (v_{11} to v_{24}), dry etch (v_{25} to v_{45}), implant (v_{46} to v_{54}), lithography (v_{55} to v_{65}), dielectric (v_{66} to v_{75}), planarization (v_{76} to v_{81}), and thin films (v_{82} to v_{92}). The remaining 13 workstations are grouped into three metrology functional areas: lithography metrology (v_{93} to v_{96}), thin films metrology (v_{97} to v_{98}), and defectivity metrology (v_{99} to v_{105}).

Because the complete manufacturing routes contain hundreds of steps, displaying every step would make the network diagram difficult to interpret. Therefore, Fig. 2 presents representative portions of the MSMN routes for both product types. The two products follow different processing sequences but may share the same workstations when undergoing the same manufacturing function. For example, both routes use workstations v_4, v_{23}, v_8, v_{97}, and v_{103} in the illustrated portions. Open and filled nodes represent processing and metrology workstations, respectively, while the connecting arcs represent stockers. Metrology operations are performed at dedicated workstations according to predefined sampling rates. When an inspected unit is identified as nonconforming, the corresponding rework route returns it to the designated processing step.

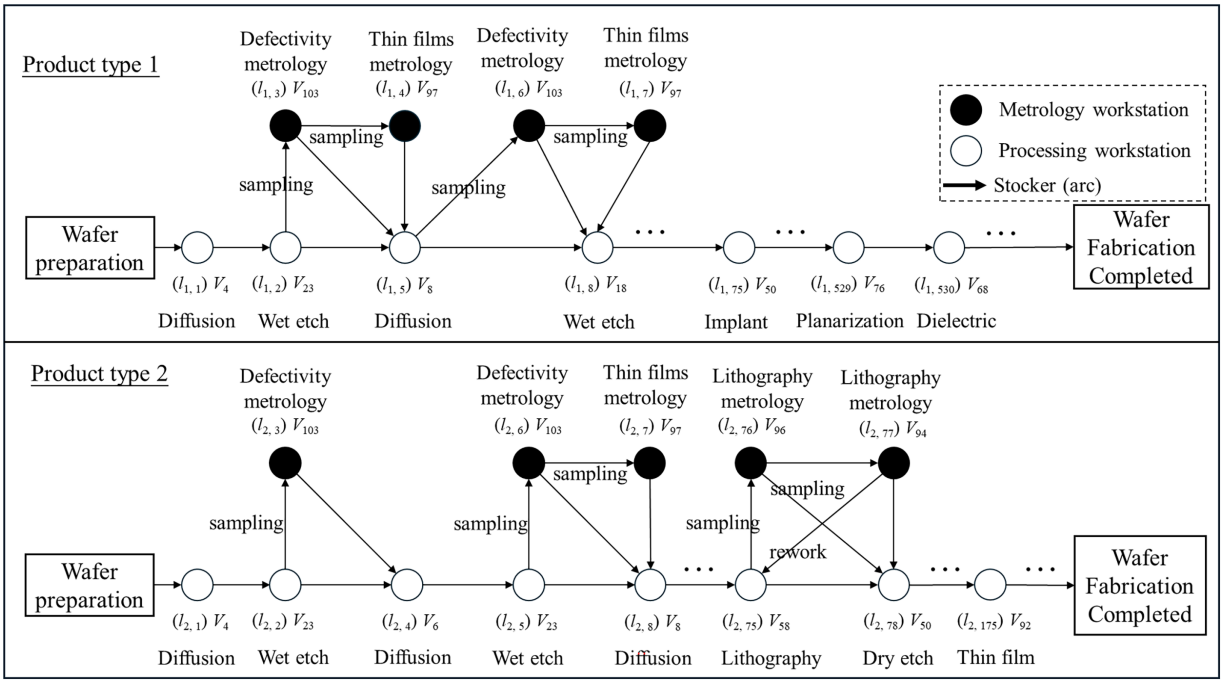

Figure 2: Representative MSMN routes of the semiconductor manufacturing system for two product types.

To distinguish the data obtained from Kopp et al. [5] from the numerical results calculated in the numerical case study, the assigned workstation, processing unit size, unit processing time, cascading interval, sampling rate, rework rate, rework return workstation, PMT, MTBF, MTTR, and maximum number of machines presented in Tables 2–4 are based on Dataset 1 from Kopp et al. [5]. For steps with a batch processing unit size, $u_{i,j}$ is calculated as the average of the specified minimum and maximum batch sizes. In contrast, the order pattern D = (36,000, 48,000) and the delivery time threshold T = 12 weeks are specified for the numerical case study. The aggregate workload times in Table 5, machine availabilities in Table 6, (D, T)-MMNV, system reliability, and sensitivity analysis results are calculated using the proposed procedure. The simulation-based reliability estimate reported in Section 5.3 is obtained separately using the simulation-based reliability estimation method.

Table 2: Processing data for the two product types based on Dataset 1 from Kopp et al. [5].

Product-Step Pair (i, j)	Assigned Workstation ($v_{l_{i,j}}$)	Processing Unit Size ($u_{i,j}$)	Unit Processing Time ($p_{i,j}$) (min/processing unit)	Cascading Interval ($c_{i,j}$) (min)
(1, 1)	v_4	Batch: 125/150	501.33	
(1, 2)	v_{23}	Single wafer: 1	0.852	0.639
(1, 3)	v_{103}	Lot: 25	17.994	
⋮	⋮	⋮	⋮	⋮
(1, 65)	v_{61}	Wafer: 1	1.464	
(1, 66)	v_{96}	Lot: 25	17.91	
(1, 67)	v_{94}	Lot: 25	20.634	
⋮	⋮	⋮	⋮	⋮
(1, 583)	v_{13}	Wafer: 1	0.648	0.486
(2, 1)	v_4	Batch: 125/150	501.33	
(2, 2)	v_{23}	Wafer: 1	0.852	0.639
⋮	⋮	⋮	⋮	⋮
(2, 343)	v_{16}	Wafer: 1	0.786	0.5895

Note: $u_{i,j}$ is measured by the number of wafers. $p_{i,j}$ is measured in minutes per processing unit, and $c_{i,j}$ is measured in minutes. A blank entry for $c_{i,j}$ indicates that product type i is processed under the single-machine mechanism at step j.

Table 3: Sampling and rework data for the two product types based on Dataset 1 from Kopp et al. [5].

Product-Step Pair (i, j)	Assigned Workstation ($v_{l_{i,j}}$)	Sampling Rate ($\alpha_{i,j}$)	Rework Rate ($\beta_{i,j}$)	Rework Return Workstation ($v_{l_{i,j}}$)
(1, 3)	v_{103}	0.56		
(1, 4)	v_{97}	0.36		
⋮	⋮	⋮	⋮	⋮
(1, 67)	v_{94}	1	0.017	$v_{l_{1,65}}$
⋮	⋮	⋮	⋮	⋮
(1, 580)	v_{93}	1		
(2, 3)	v_{103}	0.56		
(2, 6)	v_{103}	0.59		
⋮	⋮	⋮	⋮	⋮

(Continued)

Table 3 (continued)

Product-Step Pair (i, j)	Assigned Workstation $(v_{l_{i,j}})$	Sampling Rate $(\alpha_{i,j})$	Rework Rate $(\beta_{i,j})$	Rework Return Workstation $(v_{l_{i,j}})$
(2, 77)	v_{94}	0.41	0.01	$v_{l_{2,75}}$
⋮	⋮	⋮	⋮	⋮
(2, 340)	v_{95}	0.42		

Table 4: Workstation data based on Dataset 1 from Kopp et al. [5].

Workstation v_g	PMT_g	$MTBF_g$	$MTTR_g$	Maximum Number of Machines (w_g)
v_1	393.15	10,080	151.2	12
v_2	393.15	10,080	151.2	2
⋮	⋮	⋮	⋮	⋮
v_{105}	206.4	10,080	35.28	6

Note: PMT_g is measured in minutes per week, whereas $MTBF_g$ and $MTTR_g$ are measured in minutes.

Table 5: Aggregate weekly workload times of workstations.

v_g	Aggregate Workload Time (AWT_g) (min/week)	v_g	Aggregate Workload Time (AWT_g) (min/week)	v_g	Aggregate Workload Time (AWT_g) (min/week)	v_g	Aggregate Workload Time (AWT_g) (min/week)
v_1	63,893.86667	v_{28}	52,397.28	v_{55}	158,335.2972	v_{82}	15,194.4846
v_2	8489.323636	v_{29}	46,203.84	v_{56}	13,784.816	v_{83}	93,261.3
v_3	34,832.52364	v_{30}	90,422.64	v_{57}	16,744	v_{84}	20,716.3722
v_4	57,483.53455	v_{31}	94,929.6	v_{58}	121,998.4244	v_{85}	70,430.1081
v_5	28,616.28121	v_{32}	50,119.2	v_{59}	6186	v_{86}	100,747.6416
v_6	20,132.57143	v_{33}	7975.44	v_{60}	187,261.931	v_{87}	17,716.4874
v_7	12,650.56	v_{34}	37,865.52	v_{61}	27,968.704	v_{88}	23,546.1795
v_8	46,998.18667	v_{35}	90,701.52	v_{62}	25,852	v_{89}	8078.022
v_9	57,617.44	v_{36}	19,707.84	v_{63}	266,080	v_{90}	20,953.3662
v_{10}	52,677.85558	v_{37}	27,389.04	v_{64}	245,028	v_{91}	8209.7955
v_{11}	22,007.2095	v_{38}	8640.72	v_{65}	5592	v_{92}	15,394.6566
v_{12}	46,732.5405	v_{39}	10,612.8	v_{66}	19,389.447	v_{93}	42,162.804
v_{13}	9486.969	v_{40}	82,479.12	v_{67}	36,954.96	v_{94}	76,606.5624
v_{14}	1552.6725	v_{41}	65,586.24	v_{68}	66,726.5874	v_{95}	72,333.6912
v_{15}	11,665.2225	v_{42}	69,666	v_{69}	17,021.28	v_{96}	100,599.348
v_{16}	51,642.5445	v_{43}	148,402.8	v_{70}	17,453.751	v_{97}	1928.808
v_{17}	35,310.4275	v_{44}	48,554.64	v_{71}	46,226.88	v_{98}	7333.0704
v_{18}	179,630.7405	v_{45}	716,865.84	v_{72}	18,765.462	v_{99}	2836.392
v_{19}	25,036.0035	v_{46}	1957.7175	v_{73}	11,188.8	v_{100}	2519.904
v_{20}	26,386.233	v_{47}	2430.258	v_{74}	5306.3586	v_{101}	9558.4104
v_{21}	9860.439	v_{48}	5036.0475	v_{75}	15,502.32	v_{102}	186.696
v_{22}	37,650.576	v_{49}	38,690.4525	v_{76}	77,895.7596	v_{103}	9550.4328
v_{23}	93,964.6995	v_{50}	35,391.3045	v_{77}	11,074.2846	v_{104}	7540.2936
v_{24}	8415.774	v_{51}	1849.7055	v_{78}	5263.6056	v_{105}	28,263.3768
v_{25}	50,400.24	v_{52}	4536.432	v_{79}	4284.126		
v_{26}	47,612.88	v_{53}	6912.24	v_{80}	20,503.326		
v_{27}	109,789.68	v_{54}	22,722.405	v_{81}	3333.78		

Table 6: Machine availability values of all workstations.

v_g	Machine Availability (r_g)	v_g	Machine Availability (r_g)	v_g	Machine Availability (r_g)	v_g	Machine Availability (r_g)
v_1	0.985221675	v_{28}	0.977517107	v_{55}	0.934580306	v_{82}	0.956937799
v_2	0.985221675	v_{29}	0.977517107	v_{56}	0.934580306	v_{83}	0.956937799
v_3	0.985221675	v_{30}	0.977517107	v_{57}	0.934580306	v_{84}	0.956937799
v_4	0.985221675	v_{31}	0.977517107	v_{58}	0.934580306	v_{85}	0.956937799
v_5	0.985221675	v_{32}	0.977517107	v_{59}	0.934580306	v_{86}	0.956937799
v_6	0.985221675	v_{33}	0.977517107	v_{60}	0.934580306	v_{87}	0.956937799
v_7	0.985221675	v_{34}	0.977517107	v_{61}	0.934580306	v_{88}	0.956937799
v_8	0.985221675	v_{35}	0.977517107	v_{62}	0.934580306	v_{89}	0.956937799
v_9	0.985221675	v_{36}	0.977517107	v_{63}	0.934580306	v_{90}	0.956937799
v_{10}	0.985221675	v_{37}	0.977517107	v_{64}	0.934580306	v_{91}	0.956937799
v_{11}	0.978473581	v_{38}	0.977517107	v_{65}	0.934580306	v_{92}	0.956937799
v_{12}	0.978473581	v_{39}	0.977517107	v_{66}	0.943396226	v_{93}	0.996512207
v_{13}	0.978473581	v_{40}	0.977517107	v_{67}	0.943396226	v_{94}	0.996512207
v_{14}	0.978473581	v_{41}	0.977517107	v_{68}	0.943396226	v_{95}	0.996512207
v_{15}	0.978473581	v_{42}	0.977517107	v_{69}	0.943396226	v_{96}	0.996512207
v_{16}	0.978473581	v_{43}	0.977517107	v_{70}	0.943396226	v_{97}	0.996512207
v_{17}	0.978473581	v_{44}	0.977517107	v_{71}	0.943396226	v_{98}	0.996512207
v_{18}	0.978473581	v_{45}	0.977517107	v_{72}	0.943396226	v_{99}	0.996512207
v_{19}	0.978473581	v_{46}	0.943396226	v_{73}	0.943396226	v_{100}	0.996512207
v_{20}	0.978473581	v_{47}	0.943396226	v_{74}	0.943396226	v_{101}	0.996512207
v_{21}	0.978473581	v_{48}	0.943396226	v_{75}	0.943396226	v_{102}	0.996512207
v_{22}	0.978473581	v_{49}	0.943396226	v_{76}	0.980392157	v_{103}	0.996512207
v_{23}	0.978473581	v_{50}	0.943396226	v_{77}	0.980392157	v_{104}	0.996512207
v_{24}	0.978473581	v_{51}	0.943396226	v_{78}	0.980392157	v_{105}	0.996512207
v_{25}	0.977517107	v_{52}	0.943396226	v_{79}	0.980392157		
v_{26}	0.977517107	v_{53}	0.943396226	v_{80}	0.980392157		
v_{27}	0.977517107	v_{54}	0.943396226	v_{81}	0.980392157		

Table 2 presents the processing data for both product types, including the assigned workstation, processing unit size, unit processing time, and cascading interval for each product-step pair. For steps with a batch processing unit size, $u_{i,j}$ is calculated as the average of the specified minimum and maximum batch sizes. Table 3 presents the sampling and rework data for both product types. In Table 3, rework data are listed only for inspection steps with a defined rework route, and the rework return workstation $v_{l_{i,j}}$ indicates the workstation assigned to the processing step j to which units of product type i are returned for rework. Table 4 presents the workstation data, including PMT, MTBF, MTTR, and the maximum number of machines, where the maximum machine number vector is $\boldsymbol{W} = (12, 2, \ldots, 6)$. Tables 2–4 present representative entries because the complete routing, processing, sampling, rework, and workstation data are extensive. The complete dataset, including the routing and workstation-assignment data required to reconstruct the full MSMN, is provided in the Supplementary Material.

5.2 Illustration of the Solution Procedure

This subsection illustrates how the proposed procedure is applied to evaluate the system reliability $R_{D,T}$. The input data described in Section 5.1 are used to obtain the (D, T)-MMNV and the corresponding system reliability.

Input:

(1) **Network structure: G** ≡ (**A**, $\mathbf{V}_N$, $\mathbf{V}_M$, $\boldsymbol{W}$). The MSMN structure is described in the first paragraph of Section 5.1 and illustrated in Fig. 2, while the maximum number of machines $\boldsymbol{W}$ is obtained from Table 4.

(2) **Process parameters:** As described in the first paragraph of Section 5.1, the number of product types is $k = 2$. The numbers of processing steps are $s_1 = 583$ and $s_2 = 343$, as obtained from Table 2. For each product-step pair (i, j), the assigned workstation index $l_{i,j}$, processing unit size $u_{i,j}$, unit processing time $p_{i,j}$, and cascading interval $c_{i,j}$ are obtained from Table 2. The sampling rate $\alpha_{i,j}$ and rework rate $\beta_{i,j}$ for all i and j are obtained from Table 3.

(3) **Workstation parameters:** PMT_g, $MTBF_g$, and $MTTR_g$, for all g, are obtained from Table 4. The sets $\mathbf{V}^{Sin}$ and $\mathbf{V}^{Cas}$ are identified according to the processing mechanism indicated by the cascading interval $c_{i,j}$ in Table 2. The ideal operational time is set to $IT = 10{,}080$ min, corresponding to one week.

(4) **Order pattern and delivery time threshold:** $D = (36{,}000, 48{,}000)$ and $T = 12$ weeks.

Step 1. Based on Eq. (14), the demand-based production requirement $\rho_i = d_i/T$ is calculated for each product type as follows:

$$\rho_1 = \frac{d_1}{T} = \frac{36000}{12} = 3000\ (\text{unit: wafers/week}) \tag{24}$$

and

$$\rho_2 = \frac{d_2}{T} = \frac{48000}{12} = 4000\ (\text{unit: wafers/week}) \tag{25}$$

Step 2. The expected operation count $e_{i,j}$ is then determined for each product-step pair. Representative calculations for processing and metrology workstations are shown below, and the same procedure is applied to all remaining product-step pairs.

$$\begin{aligned}
&v_{l_{1,1}} = v_4 \in \mathbf{V}_N \rightarrow e_{1,1} = \frac{3000}{137.5} = 21.81818182\ (\text{batches/week})\\
&v_{l_{1,2}} = v_{23} \in \mathbf{V}_N \rightarrow e_{1,2} = \frac{3000}{1} = 3000\ (\text{wafers/week})\\
&\vdots\\
&v_{l_{1,65}} = v_{61} \in \mathbf{V}_N \rightarrow e_{1,65} = \frac{3000}{1} + \frac{3000 \times 1 \times 0.017}{1} = 3051\ (\text{wafers/week})\\
&v_{l_{1,66}} = v_{96} \in \mathbf{V}_M \rightarrow e_{1,66} = \frac{3000 \times 0.47}{25} = 56.4\ (\text{lots/week})\\
&v_{l_{1,67}} = v_{94} \in \mathbf{V}_M \rightarrow e_{1,67} = \frac{3000 \times 1}{25} = 120\ (\text{lots/week})\\
&\vdots\\
&v_{l_{1,583}} = v_{13} \in \mathbf{V}_N \rightarrow e_{1,583} = \frac{3000}{1} = 3000\ (\text{wafers/week})\\
&v_{l_{2,1}} = v_4 \in \mathbf{V}_N \rightarrow e_{2,1} = \frac{4000}{137.5} = 29.09090909\ (\text{batches/week})\\
&v_{l_{2,2}} = v_{23} \in \mathbf{V}_N \rightarrow e_{2,2} = \frac{4000}{1} = 4000\ (\text{wafers/week})\\
&\vdots\\
&v_{l_{2,343}} = v_{16} \in \mathbf{V}_N \rightarrow e_{2,343} = \frac{4000}{1} = 4000\ (\text{wafers/week})
\end{aligned} \tag{26}$$

Step 3. Obtain the aggregate workload time AWT_g for each workstation.

Step (3.1) Calculate the workload time $t_{i,j}$ for a product-step pair (i, j) assigned to a single-machine workstation $v_{l_{i,j}}$. For example, product type 1 at Step 1, i.e., $(i, j) = (1, 1)$, is assigned to workstation $v_4 \in \mathbf{V}^{Sin}$. Therefore, its workload time is calculated as follows:

$$t_{1,1} = p_{1,1} \times e_{1,1} = 501.33 \times 21.81818182 = 10,938.10909 \tag{27}$$

Step (3.2) Calculate the workload time $t_{i,j}$ for a product-step pair assigned to a cascading-machine workstation $v_{l_{i,j}}$. For example, product type 1 at Step 2, i.e., $(i, j) = (1, 2)$, is assigned to workstation $v_{23} \in \mathbf{V}^{Cas}$. Therefore, its workload time is calculated as follows:

$$t_{1,2} = p_{1,2} + (e_{1,2} - 1) \times c_{1,2} = 0.852 + (3000 - 1) \times 0.639 = 1917.213 \tag{28}$$

Step (3.3) Calculate the aggregate workload time AWT_g for each workstation v_g. Here, workstation v_4 is used as an example. Since product-step pairs (1, 1), (1, 100), (2, 1), and (2, 96) are all assigned to workstation v_4, i.e., $l_{1,1} = l_{1,100} = l_{2,1} = l_{2,96} = 4$, the aggregate workload time of v_4 is calculated as follows:

$$AWT_4 = t_{1,1} + t_{1,100} + t_{2,1} + t_{2,96} = 10,938.10909 + 13,697.69143 + 14,584.14545 + 18,263.58857 = 57,483.53 \tag{29}$$

Applying the same calculation to all product-step pairs yields the aggregate weekly workload time of each workstation, as summarized in Table 5.

Step 4. Calculate the available operational time OT_g as follows:

$$\begin{aligned} &OT_1 = IT - PMT_1 = 10,080 - 393.15 = 9686.85 \\ &OT_2 = IT - PMT_2 = 10,080 - 393.15 = 9686.85 \\ &\vdots \\ &OT_{105} = IT - PMT_{105} = 10,080 - 206.4 = 9873.60 \end{aligned} \tag{30}$$

Step 5. Based on Eq. (21), the ((36,000, 48,000), 12)-MMNV $\underline{X} = \left(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_{105}\right)$ is calculated as follows:

$$\begin{aligned} &\underline{x}_1 = \left\lceil \frac{AWT_1}{OT_1} \right\rceil = \left\lceil \frac{63893.86667}{9686.85} \right\rceil = 7 \\ &\underline{x}_2 = \left\lceil \frac{AWT_2}{OT_2} \right\rceil = \left\lceil \frac{8489.323636}{9686.85} \right\rceil = 1 \\ &\vdots \\ &\underline{x}_{105} = \left\lceil \frac{AWT_{105}}{OT_{105}} \right\rceil = \left\lceil \frac{28263.3768}{9873.6} \right\rceil = 3 \end{aligned} \tag{31}$$

Because $\underline{X} = (7, 1, \ldots, 3) \leq \boldsymbol{W} = (12, 2, \ldots, 6)$, the required minimal machine numbers are feasible under the current system configuration. Therefore, $\underline{X}$ is the ((36,000, 48,000), 12)-MMNV.

Step 6. Calculate machine availability r_g. We use workstation v_1 as an example. Because $MTBF_1 = 10{,}080$ and $MTTR_1 = 151.2$, the machine availability r_1 is calculated as follows:

$$r_1 = \frac{MTBF_1}{MTBF_1 + MTTR_1} = \frac{10080}{10080 + 151.2} = 0.985221675 \tag{32}$$

Applying the same calculation to all workstations yields the machine availability values summarized in Table 6.

Step 7. According to the ((36,000, 48,000), 12)-MMNV $\underline{X} = (7, 1, \ldots, 3)$ obtained from Step 5 and the machine availabilities obtained from Step 6, the system reliability is calculated as follows:

$$R_{(36000,48000),12} = Pr\left(X \geq (7,1,\ldots,3)\right) = \prod_{g=1}^{105} \Pr\left(x_g \geq \underline{x}_g\right) = 0.8040 \tag{33}$$

Output: System reliability $R_{(36,000,\ 48,000),12} = 0.8040$.

The system reliability $R_{(36,000,\ 48,000),12}$ is computed as 0.8040, indicating that the probability of completing the order pattern consisting of 36,000 wafers of product type 1 and 48,000 wafers of product type 2 within the 12-week delivery time threshold is 80.40%.

The above result is obtained under the assumption that all stockers are perfectly reliable. To illustrate the effect of relaxing this assumption, a simple scenario is considered in which one representative stocker, denoted by ξ, is subject to failure, whereas the remaining stockers are still assumed to be perfectly reliable. Let r_ξ denote the probability that stocker ξ operates normally over the planning horizon. Under the commonly adopted independent-component assumption in MMN reliability evaluation [6,10,16,19], the successful completion of the order pattern requires both sufficient workstation capacities and the normal operation of stocker ξ. Therefore, the adjusted reliability in this scenario can be written as

$$R'_{D,T} = R_{D,T} \times r_\xi \tag{34}$$

For example, when $R_{(36,000,\ 48,000),12} = 0.8040$, the adjusted reliability becomes 0.7960, 0.7799, and 0.7638 if r_ξ is set to 0.99, 0.97, and 0.95, respectively. Such a scenario analysis shows that the perfectly reliable stocker assumption provides an upper bound on system reliability when stocker failures are non-negligible.

5.3 Comparison with Simulation-Based Reliability Estimation

To further compare the proposed procedure with an alternative reliability evaluation method, a simulation-based reliability estimation method is considered. Following the simulation concept of Ramirez-Marquez and Coit [32], a machine number vector $X = (x_1, x_2, \ldots, x_{n+m})$ is randomly generated in each simulation trial according to the probability distributions of workstation machine states. The generated machine number vector is then compared with the (D, T)-MMNV $\underline{X} = \left(\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_{n+m}\right)$. If the generated machine number vector is no smaller than the (D, T)-MMNV, the trial is regarded as successful; otherwise, it is regarded as unsuccessful. After a large number of simulation trials are completed, the system reliability is estimated as the ratio of successful trials to total trials.

For $D = (36{,}000,\ 48{,}000)$ and $T = 12$, the exact reliability obtained by the proposed procedure is $R_{(36,000,\ 48,000),12} = 0.8040$. To evaluate the sampling variability of the simulation-based reliability estimation method, 30 independent simulation replications are conducted, with each replication consisting of 5000 trials. Across the 30 replications, the mean reliability estimate is 0.8021, with a standard deviation of 0.0057 and a 95% confidence interval of [0.7999, 0.8042] for the mean estimate. The exact reliability obtained by the

proposed procedure lies within this confidence interval. The absolute difference between the mean estimate and the exact value is 0.0019, corresponding to a relative error of approximately 0.24%. The mean absolute error and root mean squared error across the 30 replications are 0.0051 and 0.0059, respectively.

The average CPU time required by the simulation-based method across the 30 replications is 69.3786 s, with a standard deviation of 0.7739 s, whereas the CPU time required by the proposed procedure is 71.1727 s in a single execution. Although the simulation-based method shows a slightly lower average CPU time in this numerical case, its reliability estimates vary across the independent replications because of simulation sampling. In contrast, the proposed procedure directly obtains the exact reliability value under the same model assumptions without simulation sampling error.

5.4 Managerial Implications of Reliability Evaluation

Several parameters may affect system reliability. For example, the order pattern D, sampling rates, and rework rates determine the expected operation counts and therefore affect the aggregate workload time at each workstation. The MTBF and MTTR values determine machine availability, which affects the probability distribution of workstation machine states. In addition, the PMT reduces the available operational time and may increase the required minimal machine number. Therefore, larger order quantities, higher sampling or rework rates, lower MTBF, higher MTTR, and longer PMT generally decrease system reliability.

In this study, the numerical sensitivity analysis focuses on two planning-oriented factors: the delivery time threshold T and capacity expansion. The unit processing times, MTBF, MTTR, and PMT are treated as system characteristics and retained at the values provided in Dataset 1. The delivery time threshold T represents the time requirement for completing the given order pattern, whereas capacity expansion represents the allocation of additional machine resources. These two analyses are used to demonstrate how the proposed framework can support tactical capacity planning and delivery time feasibility assessment.

5.4.1 Delivery Time Threshold Sensitivity

To examine the effect of the delivery time threshold on system reliability, a sensitivity analysis is conducted from the perspective of due date setting. In this analysis, the given order pattern is fixed at D = (36,000, 48,000), and the target reliability threshold is set to 0.85. The delivery time threshold T is then varied, and the corresponding system reliability is recalculated for each candidate value of T. Specifically, nine delivery time thresholds, ranging from 9 to 17 weeks in one-week increments, are considered. For D = (36,000, 48,000), these values correspond to demand-based production requirements ranging from 4000 to 2117.65 wafers/week for product type 1 and from 5333.33 to 2823.53 wafers/week for product type 2. Through this procedure, the minimum value of the delivery time threshold required to satisfy the prescribed reliability threshold can be identified.

The results are presented in Fig. 3. It can be observed that system reliability increases as the delivery time threshold is extended. When T = 12 weeks, the system reliability is 0.8040, which remains below the target reliability threshold of 0.85. When the delivery time threshold is extended to T = 14 weeks, the system reliability increases to 0.8526, thereby exceeding the target value. Therefore, under the current accepted order pattern and system configuration, the minimum delivery time threshold required to satisfy the target reliability threshold is 14 weeks, implying that a two-week extension is necessary.

From a managerial perspective, this analysis provides a practical basis for order commitment and due date negotiation. Before accepting an order, the production planner can evaluate whether the proposed delivery time threshold satisfies the required reliability level. If the evaluated reliability falls below the target threshold, the planner can determine the minimum extension of the delivery time threshold needed to ensure

a reliable commitment. In the present case, a 12-week delivery time threshold results in insufficient reliability, whereas a 14-week threshold satisfies the target reliability requirement. Therefore, the proposed analysis can support production planners in setting feasible due dates while balancing customer responsiveness and delivery reliability.

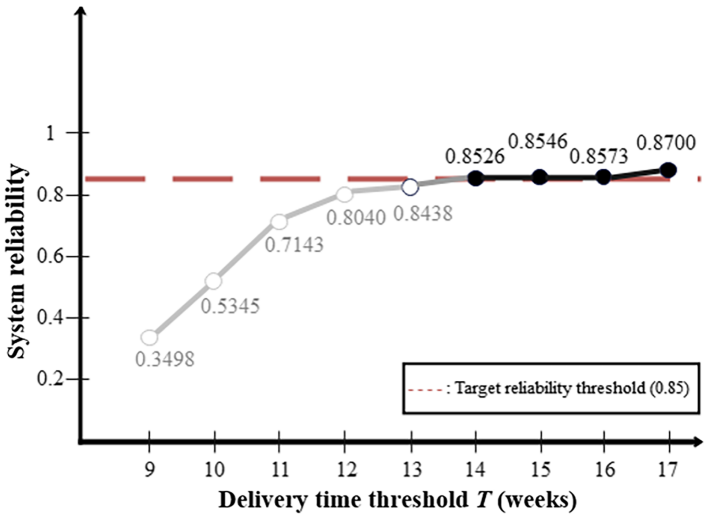

Figure 3: System reliability under different delivery time thresholds.

5.4.2 Capacity Expansion Sensitivity

While the preceding analysis supports due date setting by adjusting the delivery time threshold, another important managerial decision concerns how limited capacity expansion resources should be allocated. To examine the impact of resource allocation on system reliability, a sensitivity analysis is conducted under the original numerical setting in Section 5.2, where the given order pattern is fixed at D = (36,000, 48,000) and the delivery time threshold is fixed at T = 12 weeks. In this analysis, the maximum machine number of one workstation is increased by one at a time, while the maximum machine numbers of the remaining workstations are kept unchanged. The system reliability is then recalculated to measure the marginal improvement generated by the additional machine at each workstation. This procedure allows the contribution of each workstation to overall system reliability to be quantitatively compared.

A total of 105 capacity expansion cases are evaluated, with one case corresponding to each workstation. Fig. 4 presents the ten workstations with the largest reliability improvements. The results show that the marginal benefit of adding one machine varies substantially across workstations, indicating that the effect of capacity expansion is highly workstation-dependent. Among all workstations, v_{47} provides the largest improvement, where the addition of one machine increases the system reliability by 0.046. This suggests that v_{47} is the most critical workstation under the current production setting. The second-largest improvement is observed at v_{21}, followed by v_{14}, v_{56}, and v_{57}. In contrast, the improvements obtained from adding one machine to other workstations are relatively limited.

From a managerial perspective, these results provide a quantitative basis for prioritizing capacity expansion decisions. When only a limited number of machines can be added, resources should be allocated first to the workstation with the highest marginal reliability improvement. Therefore, under the current system configuration, adding capacity to v_{47} should be given the highest priority, followed by v_{21} and other high-ranking workstations shown in Fig. 4. Although the analysis focuses on capacity expansion, the

identified high-ranking workstations can also serve as reference points for preventive maintenance planning because their capacity loss may strongly affect system reliability. This targeted allocation strategy can improve system reliability more effectively than allocating capacity expansion resources without considering marginal reliability improvements.

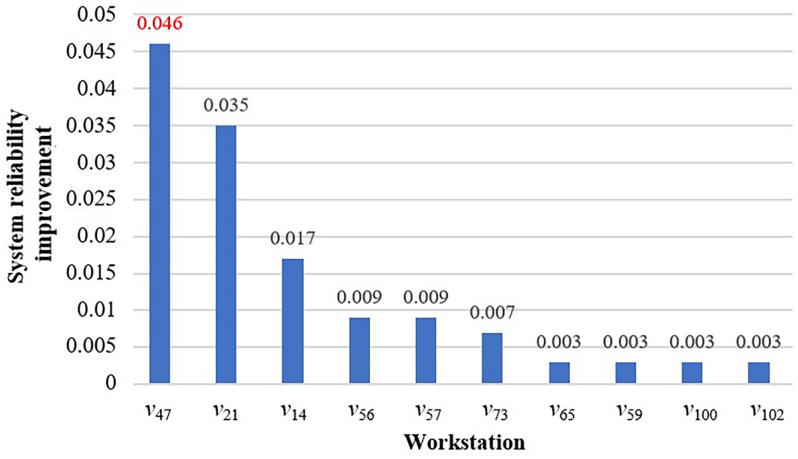

Figure 4: Top 10 workstations ranked by marginal improvement in system reliability after adding one machine.

6 Conclusions

This study developed a system reliability evaluation framework for semiconductor manufacturing systems with stochastic workstation capacities. By modeling such a system as an MSMN, the proposed framework evaluates the probability that a given customer order pattern can be completed within the delivery time threshold. The main contributions of this study are summarized as follows.

(1) **A reliability-based modeling framework for semiconductor manufacturing systems:** This study formulates semiconductor manufacturing systems as MSMNs with stochastic workstation capacities. Unlike conventional approaches that mainly evaluate production performance under fixed-capacity assumptions, the proposed framework explicitly considers the stochastic number of operational machines at each workstation. This enables system reliability to be used as a performance index for evaluating whether customer orders can be fulfilled within the required delivery time threshold.

(2) **A workload-based reliability evaluation procedure:** This study develops a workload-based procedure for deriving the minimal machine number vector required to complete a given order pattern within the delivery time threshold. The procedure incorporates characteristics of semiconductor manufacturing, including reentrant processing flows, product routes, sampling inspection, rework operations, and different processing mechanisms. By identifying the (D, T)-MMNV, the procedure provides a systematic basis for system reliability evaluation without exhaustively enumerating all feasible machine number vectors.

(3) **Managerial support for capacity and delivery time decisions:** The proposed framework provides decision support for production planning in semiconductor manufacturing systems. The numerical case study demonstrates that system reliability can be used to assess order fulfillment capability, identify critical workstations for capacity improvement, and evaluate whether the delivery time threshold should be extended to satisfy a target reliability level.

Overall, this study provides a reliability-based evaluation approach for semiconductor manufacturing systems with stochastic workstation capacities. The proposed MSMN framework offers a probabilistic

measure for evaluating whether a given customer order pattern can be completed within the delivery time threshold under capacity uncertainty. Future research may extend the proposed planning-level reliability evaluation framework by incorporating random variation in sampling and rework operations, as well as time-dependent downtime and repair processes. In addition, the framework may be linked with operational-level simulation or scheduling models to examine how detailed shop-floor dynamics affect actual order completion performance.

Acknowledgement: Not applicable.

Funding Statement: This work is partially supported by National Science and Technology Council, Taiwan, under Grant No. NSTC 114-2221-E-030-015.

Author Contributions: The authors confirm contribution to the paper as follows: conceptualization, Cheng-Ta Yeh, Yu-Zhong Kang and Yi-Kuei Lin; methodology, Cheng-Ta Yeh and Yu-Zhong Kang; software, Yu-Zhong Kang; validation, Cheng-Ta Yeh and Yi-Kuei Lin; formal analysis, Cheng-Ta Yeh; investigation, Cheng-Ta Yeh and Louis Cheng-Lu Yeng; resources, Yi-Kuei Lin; data curation, Yu-Zhong Kang; writing—original draft preparation, Cheng-Ta Yeh; writing—review and editing, Yi-Kuei Lin; visualization, Yu-Zhong Kang; supervision, Yi-Kuei Lin; funding acquisition, Cheng-Ta Yeh and Yi-Kuei Lin. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The complete dataset used in the numerical case study is provided in the Supplementary Material accompanying this article.

Ethics Approval: Not applicable.

Conflicts of Interest: Given his role as Guest Editor of this journal, Yi-Kuei Lin had no involvement in the peer review of this article and had no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to another journal editor. The authors declare no other conflicts of interest.

Supplementary Materials: The supplementary dataset accompanying this article contains the complete data used in the numerical case study, including product routings, workstation assignments, processing parameters, sampling and rework data, and workstation parameters. The supplementary material is available online at https://www.techscience.com/doi/10.32604/cmes.2026.085541/s1.

References

1. Rashidi E, Bhuiyan TH, Mason SJ. Production planning for semiconductor manufacturing under demand and yield uncertainty. Comput Ind Eng. 2024;196:110403. doi:10.1016/j.cie.2024.110403.
2. Jeon W, Lee MY, Lee JH, Lyuh CG. XEM: tensor accelerator for AB21 supercomputing artificial intelligence processor. ETRI J. 2024;46(5):839–50. doi:10.4218/etrij.2024-0141.
3. Gu J, Guo P, Wang Y, An Y, Zhu D, Wen K, et al. Adaptive reinforcement learning for dynamic semiconductor wafer fabrication scheduling: a hybrid reward approach with online bootstrap. Comput Ind Eng. 2026;212:111666. doi:10.1016/j.cie.2025.111666.
4. Ruberti M. The chip manufacturing industry: environmental impacts and eco-efficiency analysis. Sci Total Environ. 2023;858(Pt 2):159873. doi:10.1016/j.scitotenv.2022.159873.
5. Kopp D, Hassoun M, Kalir A, Monch L. SMT2020—a semiconductor manufacturing testbed. IEEE Trans Semicond Manufact. 2020;33(4):522–31. doi:10.1109/tsm.2020.3001933.
6. Chang PC. A simulation analysis of the impact of finite buffer storage on manufacturing system reliability. Simul Model Pract Theory. 2017;70:149–58. doi:10.1016/j.simpat.2016.10.006.
7. Wu Y, Lin Z, Liu C, Huang T, Chen Y, Ru Y, et al. Resilience enhancement for urban distribution network via risk-based emergency response plan amendment for ice disasters. Int J Electr Power Energy Syst. 2022;141:108183. doi:10.1016/j.ijepes.2022.108183.

8. Wu Y, Chen S. Resilience modeling and pre-hazard mitigation planning of transportation network to support post-earthquake emergency medical response. Reliab Eng Syst Saf. 2023;230:108918. doi:10.1016/j.ress.2022.108918.
9. Xu XZ, Zeng X, Niu YF. Performance assessment of a multi-state supply chain network under supply capacity, emission and reliability considerations. Reliab Eng Syst Saf. 2026;266:111651. doi:10.1016/j.ress.2025.111651.
10. Lin YK, Chang PC, Yeng LC, Shih PS. Reliability evaluation for an intermittent production system with stochastic number of normal machines. J Manuf Syst. 2017;45:222–35. doi:10.1016/j.jmsy.2017.10.002.
11. Kozyra PM. A dictionary algorithm for the solution to the generalized quickest path reliability problem. Reliab Eng Syst Saf. 2026;268:112035. doi:10.1016/j.ress.2025.112035.
12. Chang PC, Yeh CT, Cheng TY, Yu VF. Reliability assessment for multistate semiconductor global logistics networks with freight charges and carbon taxes Considerations. Reliab Eng Syst Saf. 2026;269:112027. doi:10.1016/j.ress.2025.112027.
13. Niu YF, Zhao X, Xu XZ, Zhang SY. Reliability assessment of a stochastic-flow distribution network with carbon emission constraint. Reliab Eng Syst Saf. 2023;230:108952. doi:10.1016/j.ress.2022.108952.
14. Niu YF, Gao ZY, Lam WHK. Evaluating the reliability of a stochastic distribution network in terms of minimal cuts. Transp Res Part E Logist Transp Rev. 2017;100:75–97. doi:10.1016/j.tre.2017.01.008.
15. Yeh CT, Lin YK, Yeng LC, Chao YL. Hybrid flow-shop manufacturing network reliability optimization using genetic algorithm and absorbing Markov chain. Appl Math Model. 2023;120:595–611. doi:10.1016/j.apm.2023.04.005.
16. Chang PC, Lin YK, Chen JC. System reliability for a multi-state manufacturing network with joint buffer stations. J Manuf Syst. 2017;42:170–8. doi:10.1016/j.jmsy.2016.12.009.
17. Zhang J, Li Y, Yu G. Emergency relief network design under ambiguous demands: a distributionally robust optimization approach. Expert Syst Appl. 2022;208:118139. doi:10.1016/j.eswa.2022.118139.
18. Huang CF, Huang DH, Chang PC. Assessing reliability for multi-state electricity networks with loss rate under cost constraint. Int J Rel Qual Saf Eng. 2025;32(6):2550036. doi:10.1142/s0218539325500366.
19. Lin YK, Huang CF, Chang PC. System reliability evaluation of a touch panel manufacturing system with defect rate and reworking. Reliab Eng Syst Saf. 2013;118:51–60. doi:10.1016/j.ress.2013.04.007.
20. Aghaei M, Amiri M, Sharifi M, Dehshiri SJH, Taghavifard MT. Redundancy allocation and optimal buffer size determination in series-parallel production systems: a hybrid approach. J Eng Res. 2025;13(4):3568–84. doi:10.1016/j.jer.2025.03.008.
21. Huang T, Zhang Q, Beer M, Liu Y, Huang HZ. A dynamic reliability assessment method for multi-state manufacturing system by merging imprecise observational information. Reliab Eng Syst Saf. 2026;266:111722. doi:10.1016/j.ress.2025.111722.
22. Hao Z, Yeh WC. GE-MBAT: an efficient algorithm for reliability assessment in multi-state flow networks. Reliab Eng Syst Saf. 2025;260:110916. doi:10.1016/j.ress.2025.110916.
23. Huang DH, Huang CF, Chang PC. Quickest path flow network reliability evaluation with stochastic transmission time. Simul Model Pract Theory. 2026;149:103271. doi:10.1016/j.simpat.2026.103271.
24. Geurtsen M, Adan J, Akçay A. Integrated maintenance and production scheduling for unrelated parallel machines with setup times. Flex Serv Manuf J. 2024;36(3):1046–79. doi:10.1007/s10696-023-09511-z.
25. Lee YH, Lee S. Deep reinforcement learning based scheduling within production plan in semiconductor fabrication. Expert Syst Appl. 2022;191:116222. doi:10.1016/j.eswa.2021.116222.
26. Friederich J, Lazarova-Molnar S. Reliability assessment of manufacturing systems: a comprehensive overview, challenges and opportunities. J Manuf Syst. 2024;72:38–58. doi:10.1016/j.jmsy.2023.11.001.
27. Zhang D, Luo Y, Liu Q. Reliability evaluation of production systems with finite buffers subject to time-dependent and operation-dependent failures. Ann Oper Res. 2024;340(1):671–91. doi:10.1007/s10479-024-05891-z.
28. Renna P. A review of game theory models to support production planning, scheduling, cloud manufacturing and sustainable production systems. Designs. 2024;8(2):26. doi:10.3390/designs8020026.
29. Mönch L, Fowler JW, Mason SJ. Production planning and control for semiconductor wafer fabrication facilities: modeling, analysis, and systems. Berlin/Heidelberg, Germany: Springer; 2012.

30. Uzsoy R, Lee CY, Martin-Vega LA. A review of production planning and scheduling models in the semiconductor industry part i: system characteristics, performance evaluation and production planning. IIE Trans. 1992;24(4):47–60. doi:10.1080/07408179208964233.
31. Chen HY, Lin YK. System reliability evaluation for simultaneous production of orders with varying tolerances. Reliab Eng Syst Saf. 2025;261:111090. doi:10.1016/j.ress.2025.111090.
32. Ramirez-Marquez JE, Coit DW. A Monte-Carlo simulation approach for approximating multi-state two-terminal reliability. Reliab Eng Syst Saf. 2005;87(2):253–64. doi:10.1016/j.ress.2004.05.002.