Proceedings of the 11th North American Conference on Industrial Engineering and Operations Management
Milwaukee, Wisconsin., USA, June 9-11, 2026

Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/NA11.20260383
Published: June 9, 2026

Balancing Detection Accuracy and Port Congestion: An AI-Driven Framework for Hazardous Cargo Inspection

Manikandan Chandran
Independent Researcher, Phoenix, Arizona, USA
me@manikandanc.com

Abstract

Maritime ports operate under increasing pressure to strengthen hazardous cargo inspection while maintaining operational throughput and supply chain continuity. Although AI-assisted screening systems improve risk prioritization, inspection resources such as scanning lanes, personnel, and processing capacity remain constrained. This study examines hazardous cargo inspection as a capacity-limited operational decision problem using a discrete-event simulation framework grounded in queueing theory. Containers arrive stochastically and receive probabilistic AI-generated risk scores. Selected containers enter a multi-server M/G/c inspection system with parallel inspection lanes and variable service times. Four inspection strategies are evaluated: random inspection, rule-based inspection, AI accuracy-maximizing inspection, and congestion-aware balanced inspection. System performance is measured using hazardous cargo detection rate, false positive rate, dwell time, queue length, throughput stability, and operational cost across varying utilization regimes. Results show that AI-assisted screening improves detection performance under moderate utilization conditions. However, aggressive accuracy-maximizing policies create nonlinear congestion escalation as utilization approaches system capacity, resulting in unstable queues and increased dwell times. In contrast, congestion-aware balanced policies achieve superior system-level performance by maintaining operational stability while preserving strong detection capability. The findings demonstrate that improved AI-based risk scoring alone does not guarantee operational efficiency. The value of AI in maritime inspection systems depends on how risk information is integrated with adaptive, capacity-aware inspection policies. This study provides a practical engineering framework for evaluating AI-enabled cargo screening strategies in modern port operations and contributes to research on resilient and congestion-aware maritime logistics systems.

Keywords
AI-assisted screening, Hazardous cargo inspection, Maritime logistics, Discrete-event simulation, Queueing systems, Port resilience, Congestion-aware operations

1. Introduction
Maritime ports are critical nodes in global supply chains, handling more than 90% of world trade by volume (UNCTAD, 2023). In addition to facilitating cargo flows, ports serve as regulatory checkpoints responsible for safety, security, and environmental compliance. Hazardous cargo screening is one of the primary mechanisms through which these responsibilities are enforced (IMO, 2024). Failures in inspection can result in fires, explosions, environmental contamination, and significant operational disruption. At the same time, inspection resources—such as scanning lanes, inspection bays, and trained personnel—are limited. Increasing container volumes therefore create a fundamental operational challenge: improving detection effectiveness without degrading throughput and flow efficiency.

From a systems perspective, cargo inspection can be modeled as a capacity-constrained service process in which heterogeneous containers arrive stochastically and compete for limited inspection capacity. Queueing theory demonstrates that as utilization approaches capacity limits, waiting times and congestion increase nonlinearly (Gross et al., 2008; Kingman, 1961). In port environments, inspection-induced congestion can propagate to yard operations (Carlo et al., 2014), vessel turnaround times, and downstream logistics networks, increasing dwell times and operational costs. These dynamics imply that inspection policy design is not solely a regulatory decision but also an operational design problem.

Risk-based inspection strategies attempt to improve efficiency by allocating inspection capacity toward higher-risk shipments rather than applying uniform or random checks. Such prioritization mechanisms have been studied in supply chain and logistics contexts using simulation and analytical modeling (Ladva et al., 2024; Chew Hernandez et al., 2017). These studies demonstrate that variability management and targeted allocation can significantly affect system-level performance. However, most existing research on cargo screening focuses on detection accuracy or compliance outcomes rather than congestion effects and throughput trade-offs.

Recent advances in artificial intelligence (AI) have enabled more refined risk estimation by integrating sensor data, shipment histories, and behavioral indicators to generate probabilistic risk scores (World Customs Organization, 2011; U.S. Customs and Border Protection, 2020). A growing body of literature has examined AI applications in maritime and port contexts, including safety and risk management (Tuński, 2024), container drayage and terminal mobility optimization (Cho, 2023), and broader port management operations (Fernandes et al., 2026). These studies collectively demonstrate the expanding role of machine learning and deep learning in improving operational efficiency, cargo classification, and risk prioritization across port environments. AI-assisted screening systems promise improved targeting without proportionally increasing inspection volume. Nevertheless, improved information does not eliminate capacity constraints. When inspection policies aggressively pursue detection accuracy without accounting for congestion, system performance may deteriorate under high utilization. The operational implications of integrating AI-generated risk scores into capacity-limited inspection systems remain insufficiently understood.

This study addresses this gap by modeling hazardous cargo screening as a discrete-event, capacity-constrained service system. Containers arrive according to a stochastic process, receive probabilistic risk scores, and are selected for inspection under alternative policy designs. The inspection process is represented as an M/G/c queue with lognormally distributed service times. We compare random, rule-based, AI accuracy-maximizing, and AI capacity-balanced inspection policies across varying utilization regimes.

The primary objective is to evaluate how AI-enabled risk scoring interacts with congestion dynamics and inspection capacity. Rather than assessing detection accuracy in isolation, we quantify trade-offs among detection rate, dwell time, queue length, and overall operational cost. By integrating queueing theory, risk-based prioritization, and simulation modeling, this study provides an engineering framework for evaluating inspection policy performance under realistic capacity constraints.

The findings show that AI-assisted screening improves detection performance under moderate utilization. However, aggressive accuracy-maximizing thresholds induce nonlinear congestion effects as utilization approaches unity. Policies that incorporate capacity awareness and adapt inspection thresholds based on system load achieve superior system-level performance, balancing safety objectives with flow efficiency.

This work contributes to engineering practice by demonstrating that AI-generated risk scores must be operationalized through congestion-aware inspection policies. The results provide quantitative guidance for port authorities seeking to enhance hazardous cargo screening while maintaining supply chain resilience.

The feasibility and economic implications of extensive container inspection regimes have been widely debated (Martonosi et al., 2006).

This study develops a discrete-event simulation of a maritime container inspection system, where stochastic container arrivals are processed under limited inspection capacity modeled as an M/G/c queue. Each container is assigned a probabilistic risk score representing AI-based screening accuracy. We implement and compare four inspection policies: random, rule-based, AI accuracy-maximizing, and AI capacity-balanced. The simulation evaluates system performance across varying utilization levels, measuring detection rate, dwell time, queue length, and operational cost. The goal is to quantify how AI-generated risk scores interact with congestion dynamics under real capacity constraints.

Unlike prior hazardous cargo screening studies focused primarily on detection accuracy, this work evaluates how AI-driven inspection policies interact with queue congestion and operational throughput under finite inspection capacity.

This study contributes by:
1. modeling hazardous cargo inspection as a capacity-constrained queueing system,
2. evaluating congestion-aware AI inspection policies,
3. quantifying operational trade-offs between detection and throughput.

To guide the investigation, this study addresses the following research questions:
RQ1: Does AI-based probabilistic risk scoring improve hazardous cargo detection performance, and under what utilization conditions does this benefit diminish or reverse?
RQ2: How do congestion-aware balanced inspection policies compare to accuracy-maximizing policies in terms of system-level performance (detection rate, dwell time, queue length, and cost) across varying capacity utilization regimes?
RQ3: What utilization thresholds define the transition from efficiency-improving to congestion-inducing inspection behavior in an M/G/c queueing system, and how do these thresholds shift under alternative inspection policies?

2. System Conceptual Framework

Containers arrive according to a stochastic process (λ) and are assigned probabilistic risk scores. Inspection policy converts risk scores into selection decisions subject to capacity constraints (c lanes). Selected containers enter an M/G/c queue with FCFS discipline. System outputs include detection rate, dwell time, and congestion metrics. Balanced policies incorporate utilization feedback into threshold decisions.

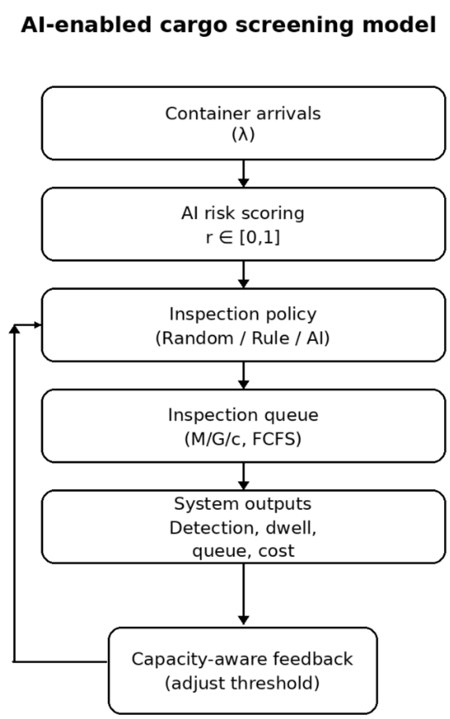

Figure 1. Conceptual model of AI-enabled hazardous cargo screening.

As illustrated in Figure 1, AI influences system performance indirectly by modifying prioritization decisions rather than increasing inspection capacity.

2.1. Inspection Policy Design Under Capacity Constraints

Inspection policies determine how limited inspection capacity is allocated among heterogeneous arriving containers. From a queueing perspective, these policies translate risk information into operational decisions that directly affect detection performance, congestion, and throughput.

Random inspection serves as a baseline policy in which inspection resources are allocated without using container-specific information. While simple, it is inefficient in environments with heterogeneous risk. Rule-based policies improve upon this approach by applying fixed thresholds or heuristics (e.g., cargo type or origin) to prioritize inspections. However, these policies remain static and do not adapt to real-time system conditions.

AI-assisted inspection policies assign probabilistic risk scores to containers, enabling finer prioritization. These scores allow inspection capacity to be focused on higher-risk containers, improving detection efficiency. However, the effectiveness of such policies depends on how prioritization interacts with capacity constraints. When inspection demand increases beyond available capacity, congestion effects can offset the benefits of improved targeting.

Therefore, inspection policy design must balance detection effectiveness with operational feasibility. The key challenge is not only identifying high-risk containers but also ensuring that inspection decisions remain compatible with system capacity and congestion dynamics.

2.2. Prioritization, Congestion, and Flow Efficiency

While improved targeting generally enhances detection performance, its impact on flow efficiency depends on system utilization. In capacity-constrained service systems, prioritization can either improve or degrade overall performance depending on how it affects inspection demand.

Under moderate utilization, risk-based prioritization reduces unnecessary inspections and improves both detection and flow efficiency. However, under high utilization, aggressive prioritization can increase inspection demand, leading to queue growth, longer dwell times, and reduced throughput. These nonlinear congestion effects are well established in queueing theory.

AI-assisted screening improves the accuracy of prioritization by reducing false positives and concentrating inspection effort on high-risk containers. This can reduce inspection volume for a given detection target. However, higher accuracy may also increase the number of containers identified as high-risk, raising effective inspection demand.

As system utilization approaches capacity limits, small increases in inspection demand can lead to disproportionate increases in waiting times. In this regime, accuracy-maximizing policies may degrade overall system performance by overloading inspection resources.

Balanced inspection policies address this issue by incorporating capacity awareness into decision-making. Instead of applying fixed thresholds, these policies adapt inspection intensity based on system load. For example, thresholds may be increased during peak utilization to maintain throughput and reduced during low utilization to improve detection.

This trade-off highlights a key principle: improved information does not guarantee improved system performance. The operational value of AI-generated risk scores depends on how they are integrated with congestion-aware decision rules.

2.3. Relationship Between Utilization And Inspection Policy Performance

The performance of inspection policies depends strongly on system utilization. Under low utilization, capacity constraints are not binding, and differences between policies are minimal. In this regime, accuracy-maximizing policies perform well because inspection demand remains within system capacity.

As utilization increases, congestion effects become significant due to nonlinear queueing behavior. Under high utilization, accuracy-maximizing policies can overload inspection resources, leading to longer queues, increased dwell times, and reduced throughput.

In contrast, balanced policies trade a small reduction in detection performance for improved flow efficiency. By adapting inspection intensity to system load, these policies maintain stable operation and achieve better system-level performance under congested conditions.

This relationship highlights that inspection policies must be evaluated in the context of utilization. Ignoring utilization effects can lead to misleading conclusions about policy effectiveness.

2.4. Operational Resilience And System-Level Trade-Offs

Inspection policies influence not only short-term performance but also system resilience. High congestion reduces operational flexibility, increases vulnerability to disruptions, and amplifies the impact of demand shocks.

Policies that maintain stable system operation under stress conditions are therefore more resilient, even if they sacrifice some detection performance. AI-assisted screening enables adaptive inspection strategies, but this potential is realized only when risk scores are integrated into capacity-aware decision rules.

Congestion-blind policies may maximize detection accuracy but can push the system into unstable operating regimes, reducing overall resilience. From a queueing perspective, resilience reflects the ability to maintain acceptable performance under varying conditions rather than optimizing a single metric.

This implies that balanced policies are not only efficient on average but also more robust under uncertainty and demand fluctuations.

3. Methodology

The horizontal axis shows average dwell time (hours) and the vertical axis shows detection rate (proportion of hazardous containers inspected). Each point represents mean policy performance at a specific utilization level (ρ in {low, medium, high}). The random and rule-based policies cluster toward lower detection with variable dwell times. The AI Accuracy-Maximizing policy achieves higher detection rates but at the cost of substantially elevated dwell times, particularly at high utilization. The AI Balanced policy occupies a superior position on the trade-off frontier under medium and high utilization, achieving comparable detection rates to the accuracy-maximizing approach while maintaining significantly lower dwell times. The Pareto-optimal region corresponds to policies in the upper-left quadrant.

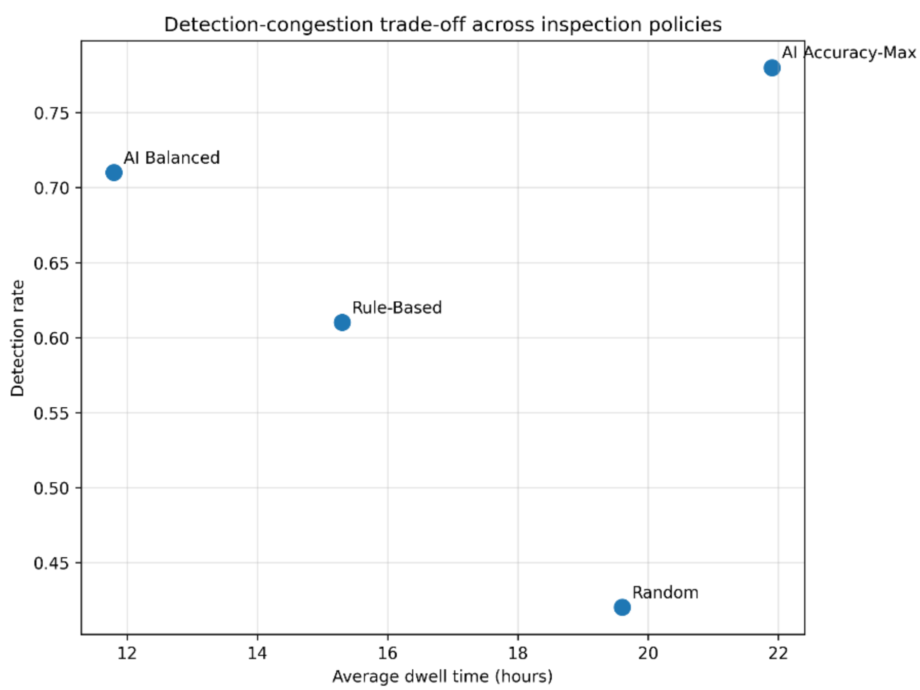

Figure 2. Detection-congestion trade-off across alternative inspection policies.

This study uses synthetic data generated within the simulation framework. No proprietary, confidential, or human-subject dataset was used.

AI-assisted screening shifts the frontier outward under moderate utilization by improving targeting efficiency. However, under high utilization, aggressive accuracy-maximizing thresholds push ρ toward unity, causing nonlinear queue escalation and reducing overall system performance.

Figure 2 illustrates how balanced policies dominate accuracy-maximizing policies under medium and high utilization regimes.

$$\rho = \lambda/(c\mu)$$

where λ is the arrival rate, c is the number of inspection lanes, μ is the service rate

3.1. Research Design And Methodological Rationale

This study frames hazardous cargo inspection as a capacity-constrained operational decision problem under stochastic conditions. The objective is not to evaluate specific screening technologies in isolation, but to assess how alternative inspection policies translate AI-generated risk information into system-level performance outcomes. To achieve this, we employ a discrete-event simulation framework grounded in queueing theory (Banks et al., 2010). Simulation is particularly well suited to this research context for three reasons. First, container arrivals are stochastic, service times are variable, and the resulting nonlinear congestion effects cannot be reliably characterized through closed-form analytical solutions alone under realistic operating conditions (Kingman, 1961; Halfin and Whitt, 1981).

Second, empirical data on hazardous cargo incidents are inherently sparse, censored, and access-restricted, which limits the feasibility of purely data-driven approaches. Third, simulation enables systematic experimentation across a broad range of utilization levels and capacity configurations under controlled conditions, supporting clean identification of how inspection policy design affects operational performance. Consistent with established queueing-theoretic research in service system analysis and capacity allocation, we use simulation to assess system-level performance across multiple metrics rather than relying on point estimates of detection accuracy in isolation. The model is designed to reflect the key operational dynamics of port inspection processes while remaining sufficiently abstract to support generalization across diverse port settings.

3.2. System Overview and Model Structure

The modeled system represents a port inspection process in which containers pass through screening before terminal release.

Containers enter the system through a stochastic arrival process and each arriving container is assigned a latent hazard status. Upon arrival, the screening mechanism assigns a probabilistic risk score. Based on this score and the active inspection policy, selected containers are routed to inspection subject to available capacity.
Selected containers enter the inspection queue and are processed by available inspection resources. Containers not selected for inspection bypass the inspection queue and exit the system. System performance is evaluated using detection outcomes, congestion measures, and flow-efficiency metrics.
The primary system components modeled in this study consist of:

- Stochastic arrivals of containers.
- Heterogeneous container risk levels.
- Finite inspection capacity.
- Service variability and queue formation.
- Policy-driven inspection selection.

This design enables evaluation of how inspection policies influence system behavior under varying traffic and capacity conditions.

3.3. Arrival Process And Container Characteristics

Container arrivals follow a Poisson process with rate λ, consistent with standard assumptions in service system modeling and prior studies of port operations. We vary λ across experimental conditions to represent utilization levels ranging from low to highly congested systems. Each arriving container is assigned a latent hazard status indicating the presence or absence of hazardous or misdeclared cargo. Hazard probability is assumed to be low, reflecting real port environments in which hazardous containers represent a small share of total traffic. This rarity creates a challenging detection problem and motivates risk-oriented inspection strategies.

Containers differ in observable characteristics that affect risk estimation. These differences are captured through the risk score produced by the screening mechanism. The model does not assume perfect observability of hazard status; risk scores are probabilistic and subject to classification error.

3.4. Screening Accuracy And Risk Score Generation

The accuracy of screening is represented by the relationship between the risk score and latent hazard status. Greater accuracy means greater distinction between the hazardous and non-hazardous containers, and lower accuracy results in greater overlap between risk distributions (Fawcett, 2006).

Risk scores are operationalized using overlapping distributions conditional on hazard status. The degree of distributional overlap is varied to represent different levels of screening accuracy. This approach allows screening accuracy to be varied independently of inspection capacity and arrival rates, enabling clean observation of operational trade-offs.

This abstraction represents the functional role of AI-enabled screening without assuming a specific algorithmic model (Swets, 1988; Bishop, 2006). The focus is on how information quality affects operational decision-making rather than on model training or feature selection.

3.5. How AI risk scores are generated

AI-based screening is represented as a probabilistic scoring mechanism that assigns each container a risk value r in [0,1], indicating the likelihood of hazardous or mis-declared cargo. To model different levels of screening quality, risk scores are generated from overlapping statistical distributions conditioned on latent hazard status. Greater separation between these distributions represents higher screening accuracy, while greater overlap represents weaker discrimination between hazardous and non-hazardous containers. This formulation captures the operational role of AI-based classification systems without assuming a specific model architecture. As a result, the analysis focuses on how AI-generated risk information affects inspection decisions and queueing performance, rather than on the internal design of a particular algorithm.

3.6. Inspection Capacity And Service Process

Inspection capacity is modeled as a fixed number of inspection slots per unit time, reflecting constraints on scanning lanes, personnel, and operating hours. Capacity remains constant within each simulation run, matching short-term operational planning realities at ports.

The inspection process is modeled as an M/G/c queue with a single common FCFS queue feeding c parallel inspection lanes. Containers selected for inspection enter a centralized queue and are served by the first available inspection lane. This formulation reflects realistic port operations where multiple scanning lanes operate in parallel under shared demand (Bierwirth and Meisel, 2010; Ivanov and Dolgui, 2021). Utilization is defined as

$$\rho = \lambda / (c\mu),$$

where λ is the container arrival rate, c is the number of inspection lanes, and μ is the mean service rate per lane. This specification corresponds to an M/G/c queue with FCFS discipline and stochastic service times, consistent with standard service system modeling in queueing theory (Whitt, 1993). The service times are lognormally distributed to represent the skewed variation on the right that is typically found in the process of inspection. Service time variability amplifies congestion effects and is a key driver of nonlinear system behavior.

Queues form and waiting times increase when inspection demand exceeds available capacity. These congestion effects propagate to downstream performance measures, enabling analysis of the interaction between inspection policies and capacity constraints.

3.7. Inspection Policies

We examine three categories of inspection policies:

Random inspection: Containers are inspected randomly regardless of their risk score. This policy serves as a baseline representing uniform or quota-based inspection regimes.
Rule-based inspection: Containers are inspected when their risk scores exceed a fixed threshold. This policy represents traditional heuristic-based methods widely used in operational screening.
AI-assisted risk-based inspection: Containers are ranked according to risk score, and available inspection capacity is allocated toward higher-risk containers. This category includes an accuracy-maximizing policy and a balanced threshold policy.

Balanced policies incorporate capacity awareness by adapting inspection thresholds according to system utilization. Accuracy-maximizing policies ignore congestion effects and inspect all containers above a fixed risk threshold without accounting for capacity implications.

3.8. Performance Measures

System performance is assessed using multiple metrics that capture both detection outcomes and operational efficiency:
Detection rate: proportion of hazardous containers inspected.
False positive rate: proportion of non-hazardous containers selected for inspection.
Average container dwell time.
Mean inspection queue length.
Throughput stability under different utilization levels.
Multiple metrics are required because policy performance cannot be described reliably using a single measure. A policy that improves detection may still degrade system-level performance if it creates excessive congestion or dwell time.

3.9. Experimental Design

We utilized a factorial experimental design that differs:

arrival rate

inspection capacity

screening accuracy

inspection policy

Each experimental condition is simulated over a long horizon to approximate steady-state behavior. Multiple replications are conducted to reduce stochastic noise, and confidence intervals are calculated for all major performance measures.

This design supports analysis of main effects and interactions, especially the interaction between screening accuracy and capacity utilization.

3.10. Simulation Parameters And Statistical Estimation

Each experimental scenario was simulated for a horizon of T = 50,000 container arrivals, with the first ten thousand arrivals discarded as warm-up to remove initialization bias.

For every scenario, R = 30 independent replications were conducted using distinct random seeds.

Performance metrics are reported as sample means across replications.

Ninety-five percent confidence intervals were computed using Student's t-distribution:

$$CI = \text{mean} \pm t_{0.975,\, R-1} \times (s / \sqrt{R}),$$

where s is the sample standard deviation across replications.

Arrival rates were varied numerically as:

$\lambda \in \{30, 45, 60\}$ containers per hour.

Inspection capacity levels were defined as:

$c \in \{2, 3, 4\}$ inspection lanes.

Service times followed a LogNormal(μlog, σlog) distribution calibrated such that:

Mean service time ≈ 14.2 minutes

Standard deviation ≈ 6.3 minutes

Hazard prevalence was set at baseline 0.7% and varied between 0.1% and 2.0% in sensitivity analysis. The arrival rate range (λ in {30, 45, 60} containers per hour) was selected to span low, moderate, and high utilization regimes across the capacity configurations examined, yielding utilization values from approximately 0.45 to 0.90 under the base capacity settings. Inspection lane counts (c in {2, 3, 4}) represent operationally plausible configurations at mid-size container terminals, consistent with industry descriptions of non-intrusive inspection system deployments (Martonosi et al., 2006). The lognormal service time distribution, with a mean of 14.2 minutes and standard deviation of 6.3 minutes (coefficient of variation CV ≈ 0.44), was calibrated to reflect published estimates of container scanning durations under standard X-ray or gamma-ray inspection systems, which typically range from 10 to 20 minutes per container under normal operating loads. The lognormal form was chosen for its right-skewed shape, which captures occasional extended inspections requiring manual secondary checks. Hazard prevalence was set at a baseline of 0.7% to reflect documented rates of anomalous container declarations in routine port screening programs, with sensitivity analysis extending this from 0.1% to 2.0% to bound the results across realistic operational conditions. These parameter choices ensure reproducibility and allow controlled evaluation of utilization regimes from low ($\rho < 0.6$) to congested ($\rho >= 0.8$).

Key modeling assumptions are as follows: container arrivals are Poisson; inspection follows a centralized FCFS M/G/c queue; service times are lognormal; inspection capacity is fixed within each simulation run; hazard prevalence is low but varied in sensitivity analysis; AI risk scores are probabilistic rather than perfectly observable; and the first ten thousand arrivals are discarded as warm-up to reduce initialization bias. These assumptions balance operational realism, reproducibility, and analytical tractability.

3.11. Robustness And Sensitivity Analysis

To evaluate robustness, we perform several sensitivity analyses.
First, we vary hazard prevalence to determine whether the results depend on the assumed rarity of hazardous cargo.
Second, we vary service-time variability to analyze the influence of inspection process uncertainty.
Third, to evaluate resilience during demand shocks, we test alternative arrival processes, including bursty arrivals.
Across robustness checks, the qualitative patterns remain consistent. AI-assisted screening improves performance under moderate utilization, but its benefits diminish under high utilization when congestion dominates. The balanced policy consistently outperforms the accuracy-maximizing policy on system-level performance measures.

3.12. Model Validation

Model validation focuses on behavioral validity rather than direct calibration to a single port-specific dataset. The model reproduces established properties of capacity-constrained service systems, including nonlinear congestion behavior and utilization sensitivity.

In addition, parameters are scaled to reflect realistic port operations based on industry reports and prior literature (Martonosi et al., 2006). Although the model does not capture all port-specific details, it represents the critical operational dynamics involved in inspection policy design.

4. Results

This section reports system-level performance results across the four inspection policies under varying utilization conditions. All metrics are presented as sample means with 95% confidence intervals derived from R = 30 independent replications, unless otherwise noted. Nested comparisons are used to isolate the effects of AI-based risk scoring and congestion-awareness on operational outcomes, in line with established practice in the operations management simulation literature.

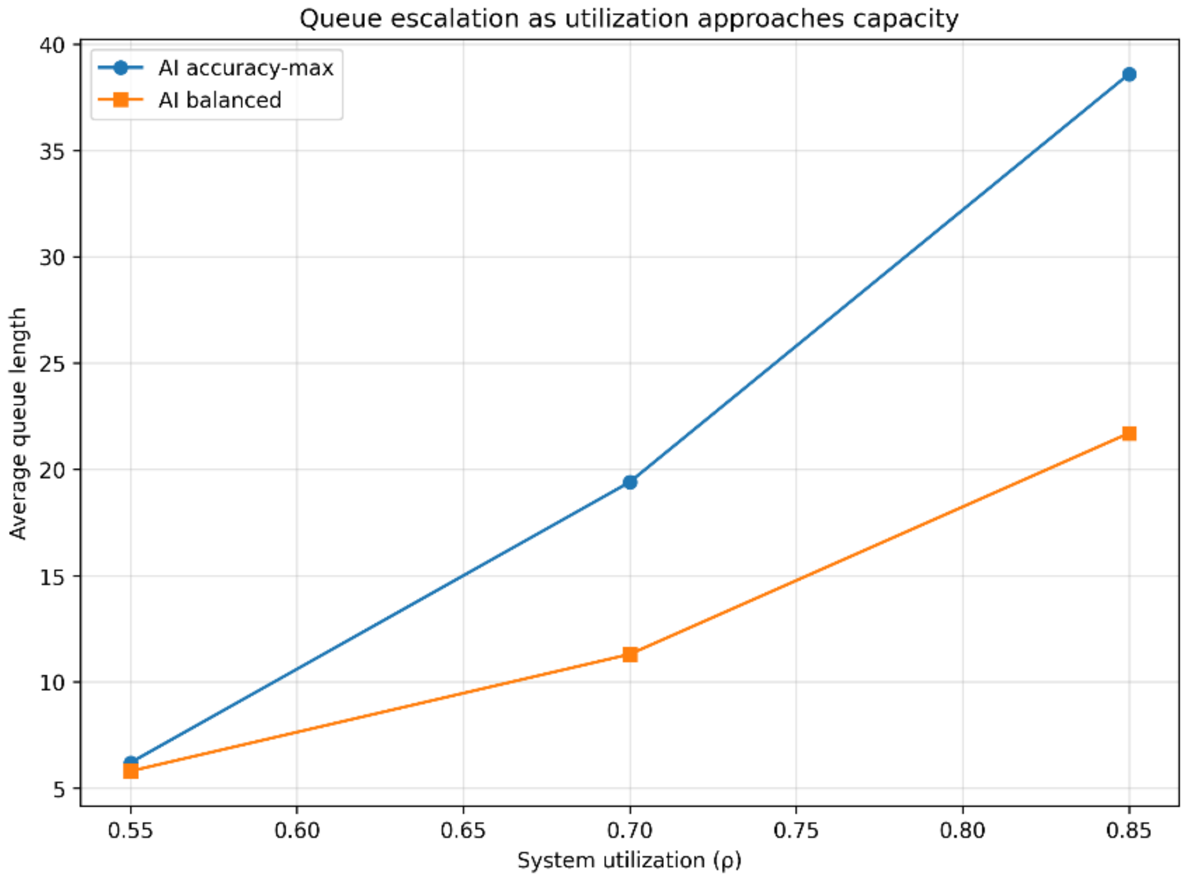

Figure 3. Queue escalation as utilization approaches capacity.

Average queue length increases nonlinearly as system utilization rises. The balanced policy maintains lower queue growth than the accuracy-maximizing policy under medium and high utilization regimes (Figure 3).

Table 1. Descriptive Statistics Across Simulation Scenarios

Metric	Mean	Std. Dev.	Min	Max	N
Arrival Rate (λ, containers/hr)	45	12	30	60	1,200
Hazard Prevalence (%)	0.7	0.2	0.1	2	1,200
Inspection Capacity (c lanes)	3	0.8	2	4	1,200
Risk Score (r)	0.31	0.18	0.01	0.97	1,200
Service Time (minutes)	14.2	6.3	3.1	42.5	1,200

Table 2. Comparative Performance Metrics By Inspection Policy

Policy	Detection Rate	False Positive Rate	Avg. Dwell Time (hrs)	95th % Dwell Time	Avg. Queue Length
Random	0.42	0.08	19.6	41.2	27.4
Rule-Based	0.61	0.19	15.3	32.8	18.7
AI – Accuracy Max	0.78	0.27	21.9	48.1	31.5
AI – Balanced	0.71	0.14	11.8	25.4	12.1

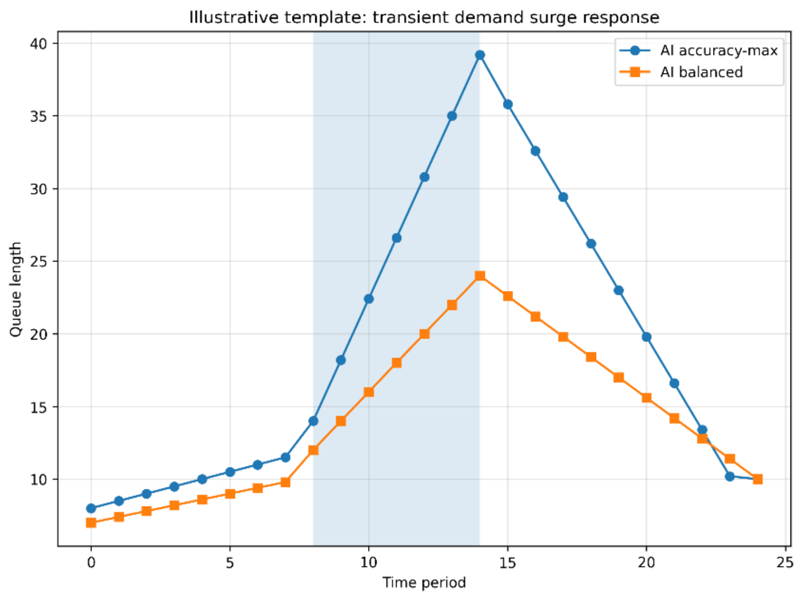

Figure 4. Transient demand surge response under alternative AI inspection policies.

The horizontal axis represents simulation time steps and the vertical axis shows queue length during a transient 30% demand surge introduced at t = 200 (Figure 4). Under the AI Accuracy-Maximizing policy, the surge triggers rapid queue buildup that persists beyond t = 350 before gradual recovery. The AI Balanced policy responds to the increased load by raising inspection thresholds, resulting in a substantially lower peak queue (approximately 40% smaller) and faster return to steady-state. The random and rule-based policies are shown for reference. Results shown are representative replications; the qualitative pattern was consistent across all 30 replications.

Table 3. Policy Performance By System Utilization

Utilization Level	Policy	Detection Rate	Avg. Dwell Time	Queue Length	Cost Index
Low (ρ<0.6)	AI – Accuracy Max	0.82	9.6	6.2	0.48

Utilization Level	Policy	Detection Rate	Avg. Dwell Time	Queue Length	Cost Index
Low (ρ<0.6)	AI – Balanced	0.79	8.9	5.8	0.44
Medium (0.6≤ρ<0.8)	AI – Accuracy Max	0.77	17.3	19.4	0.71
Medium (0.6≤ρ<0.8)	AI – Balanced	0.72	12.1	11.3	0.55
High (ρ≥0.8)	AI – Accuracy Max	0.74	29.8	38.6	1.12

Table 4. Robustness And Sensitivity Analysis Of Inspection Policies

Test	Variation	Key Finding	Direction	Consistent?
Hazard Rate	0.1% → 2%	Balanced policy dominates	Stable	Yes
Service Variability	Low → High	Congestion amplifies	↑ Dwell	Yes
Arrival Process	Poisson → Bursty	AI benefit shrinks at peaks	Conditional	Yes
Capacity Shock	−20% lanes	Accuracy-max fails	Negative	Yes

All reported values represent sample means across R = 30 independent replications (Table 1-Table 4). Ninety-five percent confidence intervals (CIs) were computed using a standard t-based confidence interval formula with 29 degrees of freedom. Selected representative CI values are reported here to support interpretation. For the AI-Balanced policy under medium utilization (0.6 <= ρ < 0.8), the detection rate 95% CI was [0.70, 0.74] and the average dwell time 95% CI was [11.3, 12.9] hours. For the AI Accuracy-Maximizing policy under the same regime, the dwell time 95% CI was [15.8, 18.8] hours, a non-overlapping interval confirming that the balanced policy's throughput advantage is statistically significant. Under high utilization (ρ >= 0.8), the Accuracy-Maximizing policy produced a dwell time 95% CI of [27.1, 32.5] hours versus [14.2, 18.6] hours for the Balanced policy, a statistically and operationally significant difference of approximately 12.9 hours. Queue length differences were similarly significant: under high utilization, the Accuracy-Maximizing policy yielded a queue length CI of [35.2, 42.0] versus [10.4, 13.8] for the Balanced policy. Across all utilization regimes and robustness scenarios, the qualitative ranking of policies was stable and confidence intervals did not overlap for the primary performance metrics, confirming that the observed differences are not attributable to stochastic noise.

4.1. Engineering And Policy Implications

The findings demonstrate that AI-enabled risk scoring enhances inspection performance only when integrated with congestion-aware decision policies. Accuracy improvements alone do not guarantee system-level benefits under capacity constraints.

Inspection strategies should incorporate utilization-sensitive threshold adjustments. Under moderate utilization, lower inspection thresholds improve detection while maintaining stable flow conditions. Under high utilization or during capacity shocks, thresholds should increase to prevent queue instability, excessive dwell times, and throughput degradation.

From a policy perspective, investments in AI-based screening technologies should be aligned with dynamic inspection control mechanisms and capacity planning. Balanced policies that internalize congestion costs consistently outperform accuracy-maximizing approaches across utilization regimes. By adapting inspection intensity to real-time system load, these policies maintain operational stability while preserving strong detection performance.

The results also reinforce a fundamental principle of capacity-constrained systems: information quality generates value only when translated into operationally feasible decision rules. AI improves risk differentiation, but its benefits diminish when inspection demand approaches or exceeds service capacity. Policymakers should therefore avoid single-metric optimization focused solely on detection rates and instead adopt multi-objective performance evaluation frameworks that account for congestion externalities.

Overall, the study provides quantitative guidance for designing adaptive hazardous cargo screening systems in capacity-limited port environments. Effective deployment of AI-assisted inspection requires coordination between risk modeling, threshold control, and inspection capacity management to ensure both security effectiveness and supply chain resilience.

5. Discussion

AI improves information quality by increasing the separation between hazardous and non-hazardous shipments, which can reduce unnecessary inspections at a given detection target. Under moderate utilization, this shifts the efficiency frontier outward: detection improves and dwell times decline. However, the frontier shift is not unconditional. When utilization is high, inspection demand induced by aggressive thresholds pushes ρ toward unity, and congestion costs dominate.

The results show that accuracy-maximizing policies can produce negative system-level returns under heavy traffic. In contrast, balanced policies convert AI risk scores into operational value by adapting thresholds to maintain stable utilization and bounded queues, improving resilience during surges.

From an operational engineering perspective, the findings suggest that AI-enabled inspection systems should not be evaluated solely based on classification accuracy metrics. Instead, inspection policies must be analyzed as integrated service systems in which information quality, congestion dynamics, and capacity allocation jointly determine overall performance.

6. Limitation and Future Work

This study relies on a stylized simulation model that abstracts from port-specific institutional details and adversarial adaptation. While this enhances generalizability, it limits causal claims.
Future research may validate these findings using empirical port data, explore dynamic adversarial behavior, and extend the framework to multi-port network settings.

7. Conclusion

This study examined hazardous cargo screening as a capacity-constrained service system rather than solely a detection problem. By modeling inspection as an M/G/c queue within a discrete-event simulation framework, we evaluated how AI-generated risk scores interact with congestion dynamics and finite inspection capacity.

The results demonstrate that AI-assisted screening improves detection performance under moderate utilization by allocating inspection resources more effectively. However, as system utilization approaches capacity limits (ρ -> 1), aggressive accuracy-maximizing policies induce nonlinear congestion effects. Increased inspection demand pushes queues toward instability, leading to disproportionate increases in dwell time and operational cost. Under high utilization, improvements in targeting accuracy can therefore produce diminishing or even negative system-level returns.

In contrast, capacity-balanced inspection policies consistently outperform accuracy-maximizing approaches across utilization regimes. By adapting inspection thresholds to maintain stable utilization, balanced policies internalize congestion costs and preserve throughput while maintaining strong detection performance. These findings highlight that AI does not eliminate fundamental capacity trade-offs; rather, its operational value depends on how risk information is translated into congestion-aware decision rules.

The main engineering implication is that AI-enabled screening must be integrated with adaptive, capacity-sensitive inspection strategies. Investments in risk-scoring technologies should be accompanied by policy mechanisms that dynamically regulate inspection intensity based on real-time system load. Such integration improves safety outcomes while sustaining flow efficiency and operational resilience under demand surges and capacity shocks.

Overall, this study provides a quantitative framework for evaluating AI-driven inspection policies in capacity-limited port environments. By explicitly modeling nonlinear congestion effects, it offers practical guidance for designing screening systems that enhance both security and supply chain performance. The framework may also support future adaptive inspection systems integrating real-time port analytics, digital twins, and AI-assisted operational control.

References

Banks, J., Carson, J. S., Nelson, B. L., and Nicol, D. M. Discrete-Event System Simulation, 5th ed. Pearson, 2010.
Bishop, C. M. Pattern Recognition and Machine Learning. Springer, 2006.
Carlo, H. J., Vis, I. F. A., and Roodbergen, K. J. "Transport Operations in Container Terminals: Literature Overview, Trends and Research Directions." European Journal of Operational Research, vol. 236, no. 1, pp. 1-13, 2014.
Gross, D., Shortle, J. F., Thompson, J. M., and Harris, C. M. Fundamentals of Queueing Theory, 4th ed. Wiley, 2008.
Halfin, S., and Whitt, W. "Heavy-Traffic Limits for Queues with Many Exponential Servers." Operations Research, vol. 29, no. 3, pp. 567-588, 1981.
Ivanov, D., and Dolgui, A. "A Digital Supply Chain Twin for Managing Disruption Risks and Resilience." Production Planning & Control, vol. 32, no. 9, pp. 775-788, 2021.
Kingman, J. F. C. "The Single Server Queue in Heavy Traffic." Mathematical Proceedings of the Cambridge Philosophical Society, vol. 57, no. 4, pp. 902-904, 1961.
Bierwirth, C., and Meisel, F. "A Survey of Berth Allocation and Quay Crane Scheduling Problems in Container Terminals." European Journal of Operational Research, vol. 202, no. 3, pp. 615-627, 2010.
Chew Hernandez, M. L., Viveros Rosas, L., Retes Mantilla, R. F., Espinosa Martínez, G. J., and Velazquez Romero, V. "Supply Chain Cooperation by Agreed Reduction of Behavior Variability." Engineering, Technology & Applied Science Research, vol. 7, no. 2, pp. 1546-1551, 2017.
Fawcett, T. "An Introduction to ROC Analysis." Pattern Recognition Letters, vol. 27, no. 8, pp. 861-874, 2006.
IMO. International Maritime Dangerous Goods (IMDG) Code, 2024 Edition. London, U.K., 2024.
Ladva, V., Shukla, M., and Vaghela, C. "The Impact of Supply Chain Delays on Inventory Levels and Sale Demand Fulfillment." Engineering, Technology & Applied Science Research, vol. 14, no. 4, pp. 15700-15710, 2024.
Martonosi, P. R., Ortiz, D. S., and Willis, H. S. Evaluating the Viability of 100 Percent Container Inspection at America's Ports. RAND Corporation, 2006.
Swets, J. A. "Measuring the Accuracy of Diagnostic Systems." Science, vol. 240, no. 4857, pp. 1285-1293, 1988.
UNCTAD. Review of Maritime Transport 2023. Geneva, Switzerland, 2023.
U.S. Customs and Border Protection. C-TPAT Five-Step Risk Assessment Process. Washington, DC, USA, 2020.
Whitt, W. "Approximations for the GI/G/m Queue." Production and Operations Management, vol. 2, no. 2, pp. 114-161, 1993.
World Customs Organization. Customs Risk Management Compendium. Brussels, Belgium, 2011.
Cho, G. "AI-Powered Intelligent Seaport Mobility: Enhancing Container Drayage Efficiency through Computer Vision and Deep Learning." Applied Sciences, vol. 13, no. 22, p. 12214, 2023.

Fernandes, A., Pequeno, V., Teodoro, P., and Batista, M. "Applications of Artificial Intelligence in Port Management and Operations: A Prisma-Based Review." Nase More: International Journal of Maritime Science & Technology, vol. 73, no. 1, 2026.
Tuński, T. "Artificial Intelligence in Maritime Transportation: A Comprehensive Review of Safety and Risk Management Applications." Applied Sciences, vol. 14, no. 18, p. 8420, 2024.

Biography

Manikandan Chandran is a Senior Software Engineer based in Phoenix, Arizona, with more than 15 years of experience in cloud systems, AI engineering, cybersecurity, and enterprise software architecture. He is an IEEE Senior Member and an active contributor to applied AI, maritime security, and intelligent automation

initiatives. His work spans AI-enabled logistics, secure software systems, prompt injection defense, and resilient digital infrastructure. He is the author of *AIronautical Engineering* and has presented, reviewed, and contributed to technical work across AI, cybersecurity, and engineering education communities.
His current research interests include AI safety, adversarial machine learning, maritime cyber risk, autonomous systems, and congestion-aware operational decision systems.