15th International Conference

on

Stochastic Models of Manufacturing and Service Operations

June 14 – 19, 2026, Edesheim, Germany

SMMSO 2026

Conference Proceedings

SMMSO 15TH CONFERENCE

KIT Karlsruhe Institute of Technology

IFL

Contents

Committees 2

Welcome message 3

Proceedings 4

Keynote: Manufacturing Physical AI: Architecting Autonomous Factories through VLMs, Sim-to-Real Learning, and Open Testbeds. *Jang, Y. J.* 4

Optimal Dynamic Scheduling Policy for a Two Product Make-To-Stock Production System with Exogenous Returns. *Gandomkar Ghalhar, A., Karaesmen, F.* 6

Optimal Dispatching in Reentrant Systems. *Bosse, T., Hahne, E., Gershwin, S., Furmans, K.* 9

An Optimization Method for Production Line Profit Maximization under Probabilistic Waiting Time Constraints. *Shi, C., Gershwin, S. B.* 12

Joint Optimization of Rates and CONWIP in Serial Lines with Quadratic Energy Costs. *Xiong, P., Liberopoulos, G.* 15

Multi-Period Production Planning with Random Yields: The Value of Yield Prediction. *Akhavan, A., Karaesmen, F.* 18

An analytical method for the performance evaluation of two-stage production systems with queue-time constraints and state-dependent hedging point policies. *Carabelli, M., Tolio, T., Magnanini, M. C.* 21

Why should we care about the retrial time distribution?. *Schmiel, O., Legros, B., Stolletz, R.* 24

From Shelf to Fulfill: An Economic Study of Store Layout and Robotic Picking in Omnichannel Retail. *Koehler, S., Klumpp, M.* 25

Same-day or next-day? Transparent time-dependent shipment pricing for e-fulfillment. *Mohring, U., Drent, M., Adan, I., van Jaarsveld, W.* 30

Energy-efficient production control of a production/inventory system with Deep Q-Networks. *Balpınarlı, D., Doğru, A., Karabağ, O., Khayyati, S., Tan, B.* 33

CO2 emission reduction of serial production lines through energy saving using reinforcement learning. *Rezaee Elyatoo, A., Sylla, A., Di Mascolo, M.* 36

Throughput Modeling for Robotic Compact Storage Systems via Hybrid Queueing Analytics and Machine Learning. *Senger, C., Lehmann, T., Altmann, L., Ernst, A., Furmans, K.* 37

Discussion on Scalable Reinforcement Learning Framework for Control Agents in Semiconductor Manufacturing. *Yeganeh, Y., Shekari, M., Frigerio, N., Matta, A.* 40

A Structure-Informed Convolutional Neural Network Surrogate Approach for Integrated Performance Evaluation & Optimization of Flow Lines via Mixed-Integer Programming. *Kloos, A., Schwarz, J. A.* 43

Designing Efficient Supply Chain Networks under Supply Disruption and Demand Uncertainty. *Azaron, A., Furmans, K.* 47

Executable Multi-Agent Path Finding Considering Practical Hardware Constraints. *Chung, C., Jang, Y. J.* 51

Stochastic Optimization of Patient-to-Room Assignments in Hospital Wards. *Schwarz, J. A., Pelz, P.* 54

Elements of a branch-price-and-cut algorithm to solve a task scheduling and assignment problem. *Helber, S.* 58

Committees

Organising Committee

- Kai Furmans (Karlsruhe Institute of Technology, Germany)
- Raik Stolletz (University of Mannheim, Germany)
- Carolin Brenner (Karlsruhe Institute of Technology, Germany)
- Benedikt Schulz (Karlsruhe Institute of Technology, Germany)
- Marion Baumann (Karlsruhe Institute of Technology, Germany)
- Maximilian Barlang (Karlsruhe Institute of Technology, Germany)
- Ömer Schmiel (University of Mannheim, Germany)

Scientific Committee

- J. Buzacott (York University, Canada)
- M. Di Mascolo (Laboratoire G-SCOP, France)
- K. Furmans (KIT, Germany)
- S. B. Gershwin (MIT, USA)
- S. Helber (Leibniz University Hannover, Germany)
- Y. Jang (KAIST, South Korea)
- F. Karaesmen (Koç University, Turkey)
- G. Liberopoulos (University of Thessaly, Greece)
- J. MacGregor Smith (University of Massachusetts, Amherst, USA)
- A. Matta (Politecnico di Milano, Italy)
- S. Meerkov (University of Michigan, Ann Arbor, USA)
- C. Papadopoulos (Aristotle University of Thessaloniki, Greece)
- R. Stolletz (University of Mannheim, Germany)
- B. Tan (Özyeğin University, Turkey)
- H. Tempelmeier (University of Cologne, Germany)
- T. Tolio (Politecnico di Milano, Italy)

Welcome message

Dear Participants,

It is our great pleasure to welcome you to the 15th International Conference on Stochastic Models of Manufacturing and Service Operations (SMMSO 2026), hosted by the Karlsruhe Institute of Technology (KIT) together with the University of Mannheim.

SMMSO continues a well-established series of biennial conferences that began in 1997 and has since brought together researchers and practitioners from around the world, held across Europe. After its early editions in Greece (1997, 1999, 2001, 2003 and 2005), the conference took place in the Netherlands (2007), Italy (2009, 2017, 2024), Turkey (2011), Germany (2013, 2019), Greece (2015) and France (2022). It is a pleasure to bring SMMSO back to Germany this year, where we come together in Edesheim, Pfalz.

The program has been designed to connect academic research with practical applications in manufacturing and service operations. It includes one keynote speech and 18 presentations, arranged into eight sessions. We are particularly honored to welcome Prof. Young Jae Jang, Department of Industrial and Systems Engineering, KAIST as our keynote speaker, as well as Dr. Georg Laipple from Robert Bosch GmbH as our industrial speaker. We sincerely thank them, together with all contributing authors, for sharing their expertise and insights.

We would also like to express our gratitude to Christos Papadopoulos, who founded this conference series, and to all past organizers whose dedication and commitment have sustained and shaped SMMSO over the years.

Finally, we thank all participants and their accompanying guests for joining SMMSO 2026. We hope you have a rewarding stay in Germany, enjoy the planned activities, and take away both valuable scientific insights and meaningful professional connections.

Sincerely,

The SMMSO 2026 Organizing Committee

Proceedings

Keynote

Prof. Young Jae Jang, Department of Industrial and Systems Engineering, KAIST

Title: Manufacturing Physical AI: Architecting Autonomous Factories through VLMs, Sim-to-Real Learning, and Open Testbeds

Abstract: Manufacturing systems are rapidly evolving toward large-scale autonomous environments in which thousands of heterogeneous robots, automation devices, and software systems must operate as a coherent whole. In this setting, traditional rule-based automation and isolated AI solutions are no longer sufficient to achieve scalability, robustness, and economic efficiency. This keynote introduces Manufacturing Physical AI as a new foundational paradigm for next-generation manufacturing systems. Manufacturing Physical AI is defined as an integrated framework that combines Vision–Language / Vision–Language–Action models (VLM/VLA), Software-Defined Manufacturing (SDM), and Sim-to-Real learning enabled by Digital Twins and Reinforcement Learning. Rather than optimizing individual machines or robots in isolation, this approach treats the entire factory as a single intelligent physical system capable of perception, reasoning, decision-making, and autonomous action. A key technical contribution discussed in this talk is the hybrid control architecture that integrates conventional stochastic network modeling with advanced data-driven learning methods. Stochastic network flow models are employed to capture uncertainty, congestion, and variability in manufacturing traffic, enabling predictive modeling of material and robot flows. These models are combined with reinforcement learning and digital twin–based learning to generate adaptive and anticipatory traffic control policies for large-scale robot fleets. This fusion allows robust robot control under uncertainty while preserving scalability and real-time responsiveness. The keynote presents how this Manufacturing Physical AI framework enables large-scale orchestration and collective control of heterogeneous robotic systems, including AMRs, OHTs, robotic arms, and automated storage systems. Through high-fidelity digital twins, control policies are trained, validated, and optimized in virtual environments and then reliably transferred to real factories using sim-to-real techniques. Real-world industrial deployment cases from semiconductor and battery manufacturing facilities are presented, demonstrating measurable improvements in operational efficiency, scalability, robustness, and deadlock-free performance. These cases illustrate how Manufacturing Physical AI moves beyond academic concepts to production-grade solutions operating in complex, high-volume manufacturing environments. Finally, the talk introduces the newly established Manufacturing Physical AI testbed at KAIST, designed as an open experimental platform for the academic community. This testbed enables reproducible research, large-scale experimentation, and collaborative development of Physical AI technologies, accelerating the transition from theory to real-world manufacturing impact.

Optimal Dynamic Scheduling Policy for a Two Product Make-To-Stock Production System with Exogenous Returns

Arash GANDOMKAR GHALHAR, Fikri KARAESMEN
Dept. of Industrial Engineering, Koç University, Sariyer, Istanbul, Turkey aghalhar22@ku.edu.tr , fkaraesmen@ku.edu.tr

We investigate a two-class make-to-stock production system served by a single shared resource, where finished goods may be returned to inventory. Returns are fully substitutable for newly manufactured items and arrive as an exogenous process independent of demand. The production controller dynamically allocates capacity across product classes to minimize the expected long-run average cost, including linear holding, backorder, and production costs.

We characterize the structure of the optimal policy and propose a scalable index-based approximation that closely matches optimal performance in numerical experiments.

Key words: Manufacturing Flow Control; Remanufacturing; Markov Decision Processes

1. Introduction

Modern supply chains often produce multiple products on shared resources under uncertainty, making it essential to allocate capacity dynamically to control holding and backorder costs. While make-to-stock rules work well in single-product settings, multi-class systems are harder because producing one item reduces capacity available to others, so priorities must adapt to the current inventory and demand state (Ha 1997, Veatch and Wein 1996). At the same time, managing product returns has become increasingly important due to sustainability initiatives and the growing use of remanufacturing and refurbishment. Because returns can significantly change inventory dynamics and costs, incorporating them into production-control models is critical (Zerhouni et al. 2013, Vercraene and Gayon 2013).

In addition, recent works integrate returns and remanufacturing directly into state-dependent make-to-stock control by treating forward production, purchasing, and reverse flows as coupled decisions. Karabağ et al. (2025a) show that leveraging real-time information on products in use (as a predictor of future returns) can improve production and remanufacturing decisions and clarifies when sustainability constraints change the optimal balance between new production and remanufacturing. Karabağ et al. (2025b) consider a system by modeling quality-differentiated returns and demonstrate that optimal purchasing and remanufacturing policies are inherently state-dependent, reflecting trade-offs between inventory positions, conversion options, and service outcomes. Song et al. (2025) similarly develop optimal purchasing and production control for a circular production system, highlighting how the structure of processing and material availability shapes the resulting priority and threshold-type policies.

Motivated by these trends, we study a two product make-to-stock system with returns. Demand for each class is stochastic and unmet demand is backordered. A single server produces one unit at a time with exponential processing times. Returns are independent and always accepted. Returned items first fill backorders and otherwise increase finished-goods inventory. At each decision epoch, the manager either produces one unit of a chosen class or idles, giving $three$ actions. We formulate the problem as an MDP and compute and compare the performance of heuristic policies with respect to the optimal policy.

Our goals are to (i) study the structure of optimal policies and build accurate approximations for the two-class case, (ii) estimate long-run average costs using uniformized relative value iteration on a truncated state space, and (iii) evaluate implementable heuristics (FCFS, base-stock FCFS, and static priorities) against the optimal benchmark. Overall, the results provide practical guidance for state-dependent scheduling in a two-product system with returns.

2. Model

2.1. Formulation

We consider a continuous-time, two-product make-to-stock system with returns and a single shared production resource. There are two item types, indexed by $i \in \{1, 2\}$, and the controller may switch the server at any time among products (or keep it idle). Switching is preemptive and incurs no cost.

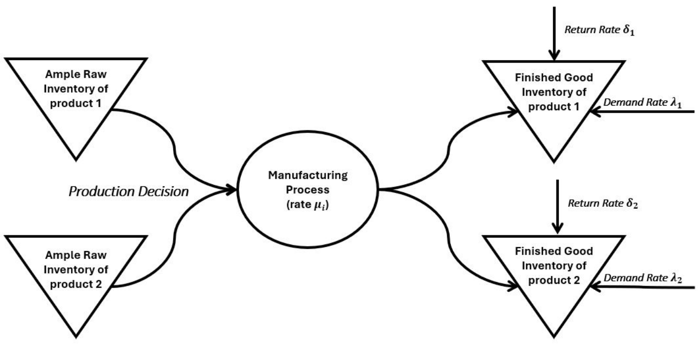


Figure 1. Two Product Make to Stock System with Returns

Let $X_i(t) \in \mathbb{Z}$ denote the net finished-goods inventory level for product i at time t: $X_i(t) > 0$ represents on-hand stock and $X_i(t) < 0$ represents backlog. We write the system state as $X(t) = (X_1(t), X_2(t)) \in \mathbb{Z}^2$, and let e_i denote the ith unit vector. Customer demands for product i arrive according to an independent Poisson process with rate $\lambda_i > 0$, and each demand decreases the state by one unit of that product, i.e., $X(t) \mapsto X(t) - e_i$.

The production server can be idle ($a = 0$) or assigned to a product $a = i \in \{1, 2\}$. When the server works on product i, completion times are exponential with rate $\mu_i > 0$, and each completion increases inventory by one unit, $X(t) \mapsto X(t) + e_i$. We allow a per-completion manufacturing cost $\kappa_i \geq 0$ when producing type i.

Inventory performance is evaluated through linear holding and backorder costs. For a state $x \in \mathbb{Z}^2$, the instantaneous cost rate is

$$c(x) = \sum_{i=1}^{2} \big(h_i x_i^+ + b_i x_i^-\big), \qquad x_i^+ = \max\{x_i, 0\},\ \ x_i^- = \max\{-x_i, 0\}, \tag{1}$$

where $h_i, b_i \geq 0$. Under a stationary Markov policy π and initial state s, the long-run average cost is

$$g^{\pi}(s) := \limsup_{T\to\infty} \frac{1}{T}\,\mathbb{E}_s^{\pi}\left[\int_0^T c\big(X(t)\big)\,dt + \sum_{i=1}^{2} \kappa_i\, N_i^{\pi}(T)\right],$$

where $N_i^{\pi}(T)$ is the number of type-i production completions by time T. Our objective is to minimize $g^{\pi}(s)$ over stationary Markov policies.

In addition to production, finished-goods returns for product i arrive exogenously as an independent Poisson process with rate $\delta_i \geq 0$ (independent of demand and control). Each return adds one unit of product i to the system, i.e., $X(t) \mapsto X(t) + e_i$.

To study policy structure, we work with a uniformized discounted dynamic program and then connect the results to the average-cost problem via vanishing-discount arguments.

3. Concluding Remarks

We studied dynamic scheduling in a two-product make-to-stock system with finished-goods returns. Using discrete-event simulation to evaluate implementable policies and value iteration for optimal benchmarks, we find that state-dependent rules—particularly Index policies—achieve performance close to optimal control, outperform static-priority rules, and reproduce the switching-curve structure of the optimal policy.

Future work includes extending the analysis to more than two classes, likely requiring approximate dynamic programming or reinforcement learning, and developing cost-weighted indices that can be learned online to adapt to realized demand and return patterns.

Acknowledgments

European Union's Horizon 2020 Research and Innovation Programme [AutoTwin project under grant agreement No. 101092021].

References

Ha, Albert Y. 1997. Optimal dynamic scheduling policy for a make-to-stock production system. *Operations Research* **45**(1) 42–53.

Karabağ, Oktay, Fikri Karaesmen, Barış Tan. 2025a. Optimizing production and remanufacturing decisions for make-to-stock hybrid manufacturing systems using real-time information on products in use. *International Journal of Production Research* doi:10.1080/00207543.2025.2559141. Advance online publication.

Karabağ, Oktay, Fikri Karaesmen, Barış Tan. 2025b. Purchasing and remanufacturing decisions with different quality returned material and finished goods. *International Journal of Production Research* **63**(12) 4413–4430. doi:10.1080/00207543.2025.2451419.

Song, Yun-Lei, Nicla Frigerio, Oktay Karabağ, Jun-Qiang Wang, Barış Tan. 2025. Optimal purchasing and production control for a circular production system. *IISE Transactions* doi:10.1080/24725854.2025.2602827. Advance online publication.

Veatch, Michael H., Lawrence M. Wein. 1996. Scheduling a make-to-stock queue: Index policies and hedging points. *Operations Research* **44**(4) 634–647.

Vercraene, S., J.-P. Gayon. 2013. Optimal control of a production-inventory system with product returns. *International Journal of Production Economics* **142**(2) 302–310.

Zerhouni, Hicham, Jean-Philippe Gayon, Yves Frein. 2013. Influence of dependency between demands and returns in a reverse logistics system. *International Journal of Production Economics* **143**(1) 62–71.

Optimal Dispatching in Reentrant Systems

Tobias Bosse
Robert Bosch GmbH, Reutlingen, Germany

Ellen Hahne
Westfield, New Jersey, USA

Stan Gershwin
Massachusetts Institute of Technology, Cambridge, USA

Kai Furmans
Karlsruhe Institute of Technology, Karlsruhe, Germany

While linear production systems have been studied successfully, manufacturing systems characterized by the reentrant paradigm lack tractable analytical models. An example of such a system is a semiconductor wafer fabrication facility, where parts revisit processing stages multiple times. The cyclic dependencies in reentrant flows cause exponential growth in analytical complexity. We model a minimal two-machine reentrant system with unreliable machines and finite buffers where the revisited machine requires a dispatch policy to select the next operation. We formulate this as a Markov decision process and evaluate exact Markov chains based on their steady-state distribution. Under the throughput-maximizing policy, the system becomes isomorphic to a tandem queue. This structure could be exploited to untangle reentrant flows and suggests a method of decomposition or an applicable heuristic for larger systems.

Key words: Reentrant Systems; Scheduling; Markov Decision Process; Tandem Queue

1. Introduction

In semiconductor wafer fabrication, a single wafer may visit the same processing tool dozens of times (Mönch et al. 2013). Between visits, the wafer undergoes different operations at other tools. This pattern—where parts return to previously visited machines—defines a reentrant manufacturing system. Managing these systems is difficult: when a machine serves multiple operations, which operation proceeds next must be selected.

Tandem production lines - where material flows in one direction - are well understood. Analytical methods decompose long lines into machine pairs connected by buffers (Gershwin 1994). Reentrant flows break this framework. The cyclic structure causes state spaces to grow exponentially with buffer size, and no analogous decomposition method exists.

We study a minimal reentrant system with two unreliable machines and finite buffers. We show that the throughput-maximizing dispatch rule is a Yoked policy that makes the reentrant system isomorphic to a two-machine tandem line.

2. Modeling Approach

The model consists of two unreliable machines, M_1 and M_2, characterized by independent geometric failure and repair times. Parts follow $M_1 \to M_2 \to M_1$ with finite buffers B_1 and B_2 between

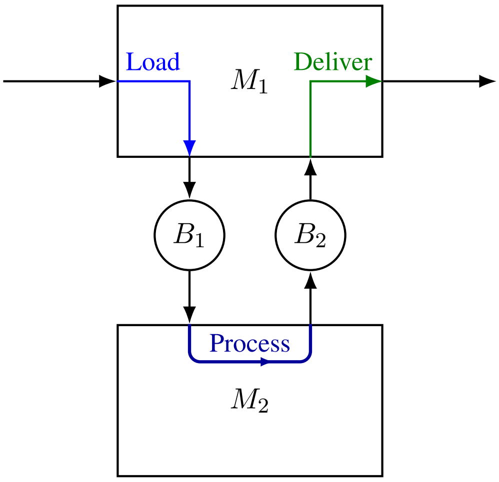


Figure 1. Topology of the reentrant model.

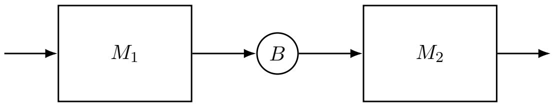


Figure 2. Equivalent two-machine line.

operations (see Figure 1). Processing times are 1, 2, and 1 for the three visits. M_2 is blocked when B_2 is full and starved when B_1 is empty. M_1 is blocked from Loading when B_1 is full and starved from Delivering when B_2 is empty. Unlimited upstream supply is available to Load new parts into the system, and unlimited downstream buffer space is available to Deliver finished parts. When both operations are available, the dispatch policy selects the next operation of M_1. M_2 processes a part if possible.

3. Method

Each deterministic dispatch policy defines a Markov chain. We compute the steady-state distribution to obtain long-term average throughput and work-in-process. The number of possible policies is $2^{(N^2)}$, where N is the capacity of each buffer. For small buffers, we enumerate all policies and compare their throughput. For larger buffers, enumeration becomes intractable, so we apply reinforcement learning with policy iteration to find the throughput-maximizing policy (Sutton and Barto 2018). Independent analysis using a restricted class of policies yields similar results, suggesting Blackwell optimality.

4. Results

To simplify the analysis, we reformulate the system with state transitions every two time steps, eliminating the boolean state variable tracking Machine 2's processing progress. Let b_k denote the current number of parts in buffer B_k, and let N denote the capacity of each buffer. Optimal policies couple Load and Deliver to keep the overall system half full, i.e., to keep the total occupancy $b_1 + b_2$ at N. This balance minimizes blocking and starvation at M_2 and collapses the state space: with total occupancy regulated, b_1 is determined by b_2. Omitting the redundant variable yields a state space equivalent to the two-machine tandem line shown in Figure 2, where buffer B also has capacity N. A simple version of the policy can be stated as follows, which we call the *Yoked policy* because it couples Load and Deliver operations:

$$\pi(b_1, b_2) = \begin{cases} \text{Load, } b_1 + b_2 < N \\ \text{Deliver, else} \end{cases} \tag{1}$$

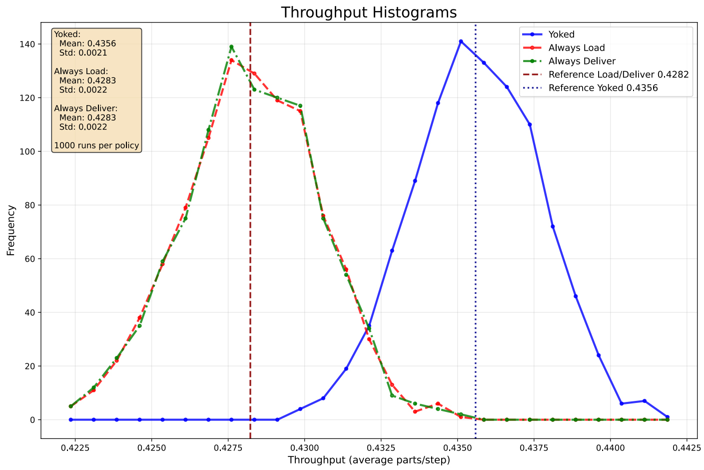


Figure 3. Simulated throughputs of a system with buffer sizes $(12|12)$.

We evaluate three dispatch policies: "Always-Load" loads parts into the system if possible, otherwise parts are delivered. "Always-Deliver" delivers parts from the system if possible, otherwise parts are loaded. "Yoked" applies the policy defined in Equation (1). Figure 3 shows the throughput distribution across these policies. "Yoked" achieves 0.436 parts per time step, exceeding the Always-Load and Always-Deliver baselines at 0.428 parts per time step; a result confirmed by simulation and steady-state analysis. The optimal policy stabilizes average work-in-process at 11.5 parts by regulating total system occupancy, while Always-Load accumulates 16.8 parts on average and Always-Deliver maintains only 7.2 by clearing parts more aggressively. The optimal policy depends on buffer occupancies, not machine states, balancing blocking and starvation at Machine 2 by regulating total buffer occupancy. The reentrant flow can be visualized as two parallel tandem queues: parts flow from M_1 to M_2 with each Load, while empty buffer space "flows" from M_1 to M_2 with each Deliver. M_2 requires both to operate, which explains why the optimal policy couples the two operations. Under the Yoked policy, the reentrant system and the tandem queue produce identical Markov chains. With $b = b_1$ (equivalently $b_2 = N - b_1$), corresponding transitions have identical probabilities. The constraint $b_1 + b_2 = N$ eliminates cyclic complexity and reduces the two-buffer reentrant system to a single-buffer tandem queue. This isomorphism unlocks analytical tractability and suggests decomposition methods for larger reentrant networks.

References

Gershwin, S. B. 1994. *Manufacturing Systems Engineering*. Prentice Hall, Englewood Cliffs, NJ.

Mönch, L., J. W. Fowler, S. J. Mason. 2013. *Production Planning and Control for Semiconductor Wafer Fabrication Facilities: Modeling, Analysis, and Systems*. Springer, New York, NY.

Sutton, R. S., A. G. Barto. 2018. *Reinforcement Learning: An Introduction*. 2nd ed. MIT Press, Cambridge, MA.

An Optimization Method for Production Line Profit Maximization under Probabilistic Waiting Time Constraints

Chuan Shi
School of Data Science, The Chinese University of Hong Kong, Shenzhen

Stanley B. Gershwin
Department of Mechanical Engineering, Massachusetts Institute of Technology, Retired

We study a production line profit maximization problem with a probabilistic waiting time constraint motivated by quality-sensitive manufacturing. The main challenge is that the waiting time distribution in a finite-buffer line with unreliable machines can be evaluated numerically but is not available in closed form, so it cannot be directly embedded in gradient-based buffer optimization. We propose an iterative surrogate framework that replaces the probabilistic constraint with an average waiting time constraint derived from Little's law and then searches over the surrogate parameter to recover feasibility for the original problem. Numerical experiments on 200 randomly generated four-machine lines show that the proposed method produces solutions nearly identical to those obtained by direct surface search.

Key words: production line optimization; buffer allocation; probabilistic waiting time constraint; chance constraints; decomposition method

1. Introduction

In quality-sensitive manufacturing systems, a part may not be allowed to wait too long between successive operations. This issue is especially important in semiconductor manufacturing, where excessive waiting may lead to yield loss, rework, or scrap (Lee and Park 2005, Kitamura et al. 2006, Kim et al. 2003). Such requirements motivate a buffer allocation problem with an explicit waiting time guarantee.

We consider a serial production line with unreliable machines and finite buffers in the discrete-time framework of Gershwin (1994). The decision variable is the vector of buffer capacities $\mathbf{N} = (N_1, \dots, N_{k-1})$. The objective is to maximize

$$J(\mathbf{N}) = AP(\mathbf{N}) - \sum_{i=1}^{k-1} b_i N_i - \sum_{i=1}^{k-1} c_i \bar{n}_i(\mathbf{N}), \tag{1}$$

subject to a production-rate requirement and a probabilistic waiting time requirement at a critical buffer $B_{\hat{i}}$:

$$P(\mathbf{N}) \geq \hat{P}, \tag{2}$$

$$\text{prob}(T(\mathbf{N}) \leq W_{\hat{i}}) \geq 1 - \alpha, \tag{3}$$

together with lower bounds on the buffers. Here $P(\mathbf{N})$ is the line production rate, $\bar{n}_i(\mathbf{N})$ is the average inventory of buffer B_i, and $T(\mathbf{N})$ is the waiting time of a part in $B_{\hat{i}}$.

The key computational difficulty is that the distribution of $T(\mathbf{N})$ can be computed numerically using the waiting time analysis of Shi and Gershwin (2016) combined with decomposition (Gershwin 1987), but it does not have a closed form. Therefore, the chance constraint in (3) cannot be directly embedded in the continuous optimization framework used for buffer design. Our contribution is an iterative surrogate-based method that converts the original problem into a sequence of tractable constrained optimization problems and then recovers a solution satisfying the original probability requirement.

2. Surrogate framework

To bypass the intractable probabilistic constraint, we introduce an average waiting time surrogate using Little's law:

$$\bar{n}_{\hat{i}} = P(\mathbf{N})\,\mathbb{E}[T_{\hat{i}}], \tag{4}$$

which leads to the surrogate constraint

$$\bar{n}_{\hat{i}} \leq \delta W_{\hat{i}} P(\mathbf{N}), \tag{5}$$

where $\delta > 0$ is a scalar controlling the tightness of the approximation. For a fixed δ, we solve

$$\max_{\mathbf{N}} \; J(\mathbf{N}) \tag{6}$$

subject to (2), (5), and the lower bounds on N_i.

This surrogate is not claimed to be an exact analytical reformulation of the chance-constrained problem. Instead, it yields a family of tractable optimization problems indexed by δ. The original probability requirement is enforced through an outer one-dimensional search over δ:

1. choose a value of δ and solve the surrogate problem;
2. compute $\mathrm{prob}(T(\tilde{\mathbf{N}}(\delta)) \leq W_{\hat{i}})$ for the resulting solution $\tilde{\mathbf{N}}(\delta)$;
3. adjust δ until the probability requirement in (3) is met while preserving maximal profit.

Operationally, this procedure links a difficult chance-constrained problem to a sequence of smoother optimization problems that can be handled with decomposition-based performance evaluation and KKT-type analysis.

3. Solving the surrogate problem

For a fixed δ, the surrogate problem remains nonlinear because both $P(\mathbf{N})$ and $\bar{n}_i(\mathbf{N})$ depend nonlinearly on the buffer sizes. Its structure can be classified by the activity of the production-rate and surrogate constraints; Table 1 summarizes the five cases.

Case	Rate constraint	Surrogate constraint
1	infeasible with the other	infeasible with the other
2	active	active
3	active	inactive
4	inactive	active
5	inactive	inactive

Table 1. Constraint activity cases.

Cases 3–5 reduce to simpler or previously studied problems, so the main difficulty is Case 2, where both constraints are active:

$$P(\tilde{\mathbf{N}}) = \hat{P}, \qquad \bar{n}_{\hat{i}}(\tilde{\mathbf{N}}) = \delta W_{\hat{i}} P(\tilde{\mathbf{N}}). \tag{7}$$

Treating $\mathbf{N}$ as continuous, we apply the KKT conditions. The gradient of $P(\mathbf{N})$ is strictly positive (Shi 2012), whereas the gradient of $\bar{n}_{\hat{i}}(\mathbf{N})$ has positive upstream and negative downstream components relative to $B_{\hat{i}}$, implying linear independence of the two active constraint gradients in the generic case.

Let $\mu_0 > 0$ and $\mu_k > 0$ be the multipliers associated with the surrogate and production-rate constraints. The stationarity condition is

$$-\nabla J(\mathbf{N}) + \mu_0 \Big(\nabla \bar{n}_{\hat{i}}(\mathbf{N}) - \delta W_{\hat{i}} \nabla P(\mathbf{N}) \Big) - \mu_k \nabla P(\mathbf{N}) = \mathbf{0}. \tag{8}$$

For a given pair (μ_0, μ_k), this is equivalent to maximizing

$$\bar{J}(\mathbf{N}) = \left(A + \mu_0 \delta W_{\hat{i}} + \mu_k\right) P(\mathbf{N}) - \sum_{i=1}^{k-1} b_i N_i - \sum_{i=1}^{k-1} c_i \bar{n}_i(\mathbf{N}) - \mu_0 \bar{n}_{\hat{i}}(\mathbf{N}), \tag{9}$$

subject to the lower bounds on buffer sizes. The Case 2 solution is obtained by a two-dimensional search over (μ_0, μ_k) until (7) holds.

4. Numerical results

We tested the proposed framework on 200 randomly generated four-machine, three-buffer production lines following Gershwin (2011). Machine isolated efficiencies were between 0.923 and 0.952. The target production rate $\hat{P}$ was set to 0.86, 0.87, or 0.88, the revenue coefficient was $A = 2000$, and the waiting time confidence level was $1 - \alpha = 0.9$. For each line, one buffer was randomly chosen as the critical buffer.

To assess solution quality, we compared the proposed algorithm with a direct surface search that checks feasibility using the original production-rate and probabilistic waiting time constraints. Table 2 summarizes the results.

Metric	Result
Identical buffer allocation to surface search	146 / 200
Average profit error	0.002%
Average production rate error	0.01%
Average maximum buffer size error	1.77%

Table 2. Results for 200 test lines.

We also compared the decomposition-based estimate of $\mathrm{prob}(T(\mathbf{N}^\star) \leq W_{\hat{i}})$ with discrete-time simulation and found very close agreement across all instances, with only a slight optimistic bias in the analytical approximation.

Acknowledgments
Chuan Shi acknowledges the funding support of the 1+1+1 CUHK-CUHK(SZ)-GDSTC Joint Collaboration Fund (Grant 2025A0505000079).

References

Gershwin, S. B. 1987. An efficient decomposition method for the approximate evaluation of tandem queues with finite storage space and blocking. *Operations Research* **35**(2) 291–305.

Gershwin, S. B. 1994. *Manufacturing Systems Engineering*. Prentice-Hall. Currently available at `http://home.comcast.net/~hierarchy/MSE/mse.html`. For corrections, see `http://web.mit.edu/manuf-sys/www/gershwin.errata.html`.

Gershwin, Stanley B. 2011. Case generation for algorithm testing. *VIII Conference on Stochastic Models of Manufacturing and Service Operations*. Kuşadası, Turkey, 193 – 200.

Kim, Ja-Hee, Tae-Eog Lee, Hwan-Yong Lee, Doo-Byeong Park. 2003. Scheduling analysis of time-constrained dual-armed cluster tools. *IEEE Transactions on Semiconductor Manufacturing* **16**(3) 521–534.

Kitamura, Shoichi, Kazuyuki Mori, Akira Ono. 2006. Capacity planning method for semiconductor fab with time constraints between operations. *Proceedings of the 2006 SICE-ICASE International Joint Conference*. Busan, Korea, 1100–1103.

Lee, Tae-Eog, Seong-Ho Park. 2005. An extended event graph with negative places and tokens for time window constraints. *IEEE Transactions on Automation Science and Engineering* **2**(4) 319–332.

Shi, C. 2012. Efficient buffer design algorithms for production line profit maximization. Ph.D. thesis, Massachusetts Institute of Technology Department of Mechanical Engineering.

Shi, C., S. B. Gershwin. 2016. Part sojourn time distribution in a two-machine line. *European Journal of Operational Research* **248**(1) 146–158. doi:10.1016/j.ejor.2015.07.057.

Joint Optimization of Rates and CONWIP in Serial Lines with Quadratic Energy Costs

Pan Xiong
Department of Industrial Engineering, School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an, Shaanxi, 710072, P.R. China, pacpos.pxiong@gmail.com

George Liberopoulos
Department of Mechanical Engineering, School of Engineering, University of Thessaly, Volos, Magnesia, 38334, Greece, glib@uth.gr

Rising energy costs are driving manufacturing toward improved energy efficiency; however, effective optimization is often hindered by traditional linear models that fail to capture the nonlinear relationship between energy consumption and processing rates. To address this limitation, we jointly optimize the processing rates and the CONWIP level of a CONWIP-controlled serial production line with quadratic energy costs and linear CONWIP costs, subject to a throughput target and rate constraints. We model the line as a closed cyclic queueing network and incorporate two Energy Management during Machine Starvation (EMMS) policies (idling and shutdown), with their optimal selection treated as part of the optimization. We decompose the problem into two hierarchical subproblems. For a fixed CONWIP level, we obtain the processing rates from a nonlinear constrained formulation based on the Bard–Schweitzer approximation to MVA and its associated KKT conditions. We then optimize the CONWIP level via a discrete search procedure supported by a variable transformation that ensures asymptotic convexity of a relaxed cost function. Sensitivity analyses show that increasing energy costs shift the optimal solution toward higher CONWIP levels, enabling more energy-efficient processing rates—a qualitative effect that linear-cost models cannot capture—while numerical experiments confirm millisecond-level computational efficiency.

Key words: Serial production lines; Nonlinear energy consumption; CONWIP; MVA; KKT; MINLP

1. Introduction

Improving energy efficiency has become a critical objective in manufacturing, driven by rising global energy prices and increasingly stringent environmental regulations. Modern manufacturing technologies enable machines to operate at controllable processing rates, allowing operators to dynamically balance throughput and operational costs. However, traditional operations management models usually simplify this relationship by assuming a linear dependence between energy use and processing rates (Gutowski et al. 2006), while empirical and mechanistic studies have shown that energy consumption is intrinsically nonlinear and predominantly quadratic in the rates (Hu et al. 2012, Imani Asrai et al. 2018). This implies that increasing processing speeds to boost throughput leads to a disproportionate rise in energy consumption.

At the system level, production performance is strongly influenced by Work-In-Progress (WIP) inventory. The CONWIP policy is widely used to control WIP, stabilize throughput, and reduce lead times (Spearman et al. 1990). At the same time, buffer allocation and space capacity decisions play a crucial role in determining system performance (Liberopoulos 2020). When nonlinear energy costs are considered alongside linear space capacity costs, a fundamental trade-off emerges. Operating at a low CONWIP level reduces inventory-related costs but requires higher processing rates to meet throughput targets, thereby incurring significant energy penalties. Conversely, reducing processing rates to save energy necessitates a higher CONWIP level to maintain throughput, leading to increased space capacity costs. In addition, machine starvation management policies, such as idling and shutdown, introduce continuous and fixed costs that further complicate system behavior.

Despite the growing body of research on energy-efficient production and CONWIP-based control, most studies treat processing rate decisions and WIP regulation separately, often relying on simplified linear energy cost assumptions (Glock and Grosse 2021). As a result, the interaction between nonlinear energy costs and system-level control policies remains insufficiently understood, particularly in stochastic production environments.

We address this gap by jointly minimizing total cost per unit time, comprising quadratic energy use costs and linear space capacity costs, over processing rates and CONWIP level, subject to a throughput target and rate constraints. We argue that joint optimization under nonlinear energy costs produces qualitatively different operating policies than linear-cost approaches, and develop a hierarchical solution method to demonstrate this.

2. Problem Description

We consider a serial production line of M non-identical single-machine workstations. Processing times are exponentially distributed with controllable rate μ_i, and the system operates under a CONWIP policy maintaining a fixed WIP level K, modeled as a closed cyclic queueing network. System performance is evaluated using Mean Value Analysis (MVA).

Our objective is to minimize expected total cost per unit time, comprising (i) a linear space capacity cost in the CONWIP level $(c_k K)$ and (ii) quadratic energy consumption costs in the processing rates $(\sum_{i=1}^{M} c_{0i} + c_{1i}\mu_i + c_{2i}\mu_i^2)$. We incorporate EMMS policies (idling and shutdown), which introduce fixed operating costs (c_{0i} per unit time if machine i remains idle) and startup costs (c_{si} per startup if machine i is shut down), and treat the choice between them as part of the optimization.

We formulate a joint optimization problem over $\boldsymbol{\mu}$ and K, subject to a throughput target (λ), upper bounds on the processing rates $(\bar{\mu}_i)$, and system dynamics controlled by the MVA equations. Due to the strong coupling between these decisions, we decompose the problem into two subproblems: Problem (P1) optimizes $\boldsymbol{\mu}$ for a fixed K; Problem (P2) determines the optimal integer K.

Problem (P1) is addressed by replacing the exact MVA equations with the Bard-Schweitzer approximation, which is chosen for its accuracy-efficiency balance in closed queueing networks, to derive tractable KKT conditions, solved via a Nested Bisection Algorithm. For Problem (P2), a change of variables reveals that a relaxed version of the cost function is asymptotically convex, enabling a Hybrid Discrete Bisection Search Algorithm to identify the global optimum with guaranteed convergence. This hierarchical decomposition avoids the combinatorial complexity of solving the full MINLP directly while maintaining high solution accuracy.

3. Numerical Results and Conclusions

To evaluate the proposed framework, we conduct numerical experiments assessing solution accuracy and computational efficiency across a range of system configurations, demonstrating millisecond-level computation times throughout.

The results confirm that joint optimization of processing rates and CONWIP level is essential for cost-efficient operation, and that nonlinear energy costs produce qualitatively different optimal policies than linear-cost models would predict. Higher quadratic energy costs lead to increased CONWIP level, enabling lower, energy-efficient processing rates while maintaining throughput; tighter WIP control forces higher processing speeds, causing disproportionate energy consumption.

In a representative configuration ($M = 10, \lambda = 12, c_K = 0.022$), we defined the nominal energy cost coefficients as $c_0 = 0.06$, $c_1 = 0.003$, $c_2 = 0.0009$, and $c_s = 0.004$, and the nominal maximum processing rate as $\bar{\mu} = 18$. To account for heterogeneity, the actual coefficients for each machine i were subjected to independent random noise: specifically, $c_{ki} = c_k(1 + \epsilon_i)$ with $\epsilon_i \sim U[-0.5, 0.5]$ for cost factors, and $\bar{\mu}_i = \bar{\mu} + \epsilon_{\mu,i}$ with $\epsilon_{\mu,i} \sim U[-2, 2]$ for processing rates. The results highlight a critical trade-off that linear-cost models systematically fail to capture: doubling the baseline quadratic cost c_2 increases the optimal CONWIP level by 27% while reducing the average processing rate by 6%; tripling c_2 raises the optimal CONWIP level by 56% and reduces the processing rate by 10%. Furthermore, we compared our joint optimization method against a standard linear energy model, which fixes processing rates at their maximum and optimizes only the CONWIP level. The proposed method significantly reduced unit-time costs as c_2 increased, yielding savings of 6% at $c_2 = 0.0018$, 13% at $c_2 = 0.0045$, and 19% at $c_2 = 0.009$. These results confirm that joint optimization effectively leverages the non-linear relationship between speed and energy use.

The analysis also highlights the impact of EMMS policy selection. High startup costs favor idling over shutdown, while lower startup penalties combined with high energy costs make shutdown more attractive, confirming that optimal EMMS decisions depend critically on the relative magnitude of fixed and variable costs, and validating the importance of treating the EMMS choice as an endogenous optimization decision rather than an exogenous modeling assumption.

From a managerial perspective, these findings offer concrete operational guidance: production managers should raise CONWIP level, rather than simply throttle processing rates, when energy prices increase, and select starvation management policies based on the startup-to-energy cost ratio rather than operational convention. These findings show a fundamental space-for-energy substitution paradigm in sustainable manufacturing, demonstrating that strategically expanding buffer capacity is an effective approach to unlock the economic benefits of energy-efficient machining. Furthermore, the principles of this energy-aware joint optimization transcend manufacturing, offering direct applicability to resource allocation in digital infrastructures like data centers.

Future research could extend this framework to complex topologies and incorporate uncertainties such as time-varying demand, machine failures, and dynamic energy pricing.

References

Glock, Christoph H, Eric H Grosse. 2021. The impact of controllable production rates on the performance of inventory systems: A systematic review of the literature. *Eur J Oper Res* **288**(3) 703–720.

Gutowski, Timothy, Jeffrey Dahmus, Alex Thiriez. 2006. Electrical energy requirements for manufacturing processes. Joost Duflou, Wim Dewulf, eds., *Proc 13th CIRP Int Conf Life Cycle Eng*. Leuven, Belgium, 623–638.

Liberopoulos, George. 2020. Comparison of optimal buffer allocation in flow lines under installation buffer, echelon buffer, and conwip policies. *Flex Serv Manuf J* **32**(2) 297–365.

Spearman, Mark L, David L Woodruff, Wallace J Hopp. 1990. CONWIP: a pull alternative to kanban. *Int J Prod Res* **28**(5) 879–894.

Multi-Period Production Planning with Random Yields: The Value of Yield Prediction

Ahmad Akhavan, Fikri Karaesmen
Dept. of Industrial Engineering, Koç University, Sarıyer, Istanbul, Turkey aastaneh25@ku.edu.tr, fkaraesmen@ku.edu.tr

In many manufacturing environments, production yields are random and may vary over time, creating significant challenges for effective production planning. This paper studies a finite-horizon production planning problem with random, time-varying yields under limited prediction availability. At each decision period, the planner observes the current inventory level together with yield-related features for a finite number of upcoming periods, which can be used to inform yield predictions. Beyond this prediction horizon, such features are unavailable, and yield behavior is therefore approximated using long-run average information. We formulate a dynamic programming model that captures these informational limitations and determines production decisions by balancing production costs, inventory holding costs, and terminal penalties for both excess inventory and inventory shortfalls at the end of the planning horizon. The analysis reveals a fundamental trade-off between producing early to hedge against yield risk and waiting to avoid surplus inventory. We examine how the length of the yield prediction horizon influences the optimal policy and expected costs. Our results demonstrate that leveraging even a short-term yield forecast can substantially improve performance compared to relying on a long-run average yield. These findings highlight the value of aligning production decisions with the limited predictive information typically available in practice and provide insights into effectively integrating short-term yield forecasts into finite-horizon production planning.

Key words: Production Planning; Yield Uncertainty; Limited Lookahead; Dynamic Programming; Inventory Control

---

1. Introduction

Production planning is often complicated by uncertainty in production yields, that is, the probability that a production attempt generates a usable unit. Classical inventory and production models frequently assume deterministic or stationary yields, although prior research has shown that yields are affected by scrap, rework, inspection failures, and process variability (Yano and Lee 1995, Wein 1992). Ignoring this uncertainty can lead to poor production timing, underestimated capacity needs, and excessive inventory costs.

In practice, yield uncertainty is often both stochastic and time-varying. This is particularly important in settings such as semiconductor manufacturing, where yields may change across lots and over time (Kumar et al. 2006, Milor 2013). Learning, process improvement, equipment deterioration, and changing operating conditions can all affect yield behavior (Argote and Epple 1990, Lapré et al. 2000). Recent work also suggests that part of this variability can be explained using observable process features, making short-term yield prediction feasible (Bibak and Karaesmen 2025).

Motivated by these observations, we study a finite-horizon make-to-stock problem with stochastic, time-varying yields and limited lookahead information. At each decision epoch, the planner observes the current inventory and yield estimates for only the next few periods; yields beyond that horizon are represented by a long-run average. We formulate a dynamic program to study how limited forecast visibility affects production decisions and cost performance.

2. Model

We consider a finite-horizon make-to-stock system over periods $t = 0, 1, \ldots, T-1$, with terminal time T. The planner aims to meet a target demand level d by time T by dynamically making production decisions over the horizon.

In each period, the planner chooses

$$a_t \in \{0, 1\},$$

where $a_t = 1$ means production is attempted.

If production is attempted in period t, one usable unit is obtained with probability $\mathbb{P}(Y_t(\theta_t) = 1)$, where θ_t is a vector of observable yield-related features. Based on θ_t, the planner uses a prediction model to estimate

$$\hat{p}_t = \mathbb{P}(Y_t(\theta_t) = 1 \mid \theta_t).$$

Inventory evolves as

$$x_{t+1} = x_t + Y_t(\theta_t). \tag{1}$$

Each production attempt incurs cost c, and inventory held from one period to the next incurs cost h per unit. At the terminal time, shortage and excess inventory are penalized with unit costs b and o, respectively, giving terminal cost

$$hx_T + b(d - x_T)^+ + o(x_T - d)^+. \tag{2}$$

The objective is to minimize expected total cost:

$$\mathbb{E}\left[\sum_{t=1}^{T} hx_t + \sum_{t=0}^{T-1} ca_t + b(d - x_T)^+ + o(x_T - d)^+\right]. \tag{3}$$

Let $V_t(x)$ be the minimum expected cost-to-go from period t onward when inventory is x. Then

$$V_T(x) = hx + b(d - x)^+ + o(x - d)^+,$$

and for $t < T$,

$$V_t(x) = \min_{a \in \{0,1\}} \left\{ca + hx + \hat{p}_t V_{t+1}(x + a) + (1 - \hat{p}_t)V_{t+1}(x)\right\}. \tag{4}$$

We assume limited lookahead information: at time t, the planner observes feature vectors

$$(\theta_t, \theta_{t+1}, \ldots, \theta_{t+L-1}),$$

which generate yield estimates

$$(\hat{p}_t, \hat{p}_{t+1}, \ldots, \hat{p}_{t+L-1}).$$

Beyond this horizon, yield probabilities are approximated by a baseline average $\bar{p}$. The planner resolves the dynamic program each period using the currently available estimates, yielding a rolling-horizon policy.

3. Conclusions

We examine how limited yield forecast horizons affect production decisions and cost performance in a finite-horizon make-to-stock system under random yield. By explicitly modeling the availability of short-term yield information, our dynamic programming framework sheds light on the interaction between yield uncertainty and planning horizon. A key finding is that acknowledging the limits of forecast visibility can materially improve decision-making. Even a modest lookahead—such as yield information for only a few future periods—enables the planner to better time production in anticipation of yield fluctuations. This leads to meaningful cost reductions relative to policies that rely on stationary yield averages or ignore near-term yield information altogether. In particular, limited-lookahead policies can hedge against anticipated yield deterioration by producing earlier, or defer production when favorable yields are expected, thereby reducing the likelihood of large end-of-horizon shortages or surpluses.

References

Argote, L, D Epple. 1990. Learning curves in manufacturing. *Science* **247**(4945) 920–924.

Bibak, B., F. Karaesmen. 2025. Integration of machine learning and optimization models for a data-driven lot sizing problem with random yield. *International Journal of Production Economics* **282** 109529.

Kumar, N, K Kennedy, K Gildersleeve, R Abelson, CM Mastrangelo, DC Montgomery. 2006. A review of yield modelling techniques for semiconductor manufacturing. *International Journal of Production Research* **44**(23) 5019–5036.

Lapré, M A, A S Mukherjee, L N V Wassenhove. 2000. Behind the learning curve: Linking learning activities to waste reduction. *Management Science* **46**(5) 597–611.

Milor, L. 2013. A survey of yield modeling and yield enhancement methods. *IEEE transactions on semiconductor manufacturing* **26**(2) 196–213.

Wein, A Spence. 1992. Random yield, rework and scrap in a multistage batch manufacturing environment. *Operations Research* **40**(3) 551–563.

Yano, C Arai, H L Lee. 1995. Lot sizing with random yields: A review. *Operations research* **43**(2) 311–334.

An analytical method for the performance evaluation of two-stage production systems with queue-time constraints and state-dependent hedging point policies

Matteo CARABELLI
Department of Mechanical Engineering, Politecnico di Milano, Milano 20156, Italy, matteo.carabelli@polimi.it
Tullio TOLIO
Department of Mechanical Engineering, Politecnico di Milano, Milano 20156, Italy, tullio.tolio@polimi.it
Maria Chiara MAGNANINI
Department of Mechanical Engineering, Politecnico di Milano, Milano 20156, Italy, mariachiara.magnanini@polimi.it

Queue-time (Q-time) constraints arise in production systems where excessive waiting between consecutive processing stages may lead to quality degradation and product scrap or rework. This work presents an approximate analytical model for the performance evaluation of two-stage production systems with parallel machines separated by an intermediate buffer subject to a finite Q-time constraint. The proposed model is based on a steady-state Markovian method for production systems with finite buffer capacity, multiple operational states, and generalized threshold-based control. The generalized thresholds allow modeling of machine state-dependent hedging point policies by linking upstream production decisions to downstream operational conditions and buffer levels. This steady-state representation is coupled with a Markov-modulated fluid process that captures the stochastic cumulative production dynamics governing Q-time violations, enabling joint evaluation of throughput performance and violation-induced yield losses. The explicit representation of hedging point policies allows systematic optimization of threshold parameters and structured exploration of their impact on system behavior.

Keywords: Stochastic Model; Queue-Time Constraints; Analytical Model; Performance Evaluation; Semiconductor

1 Introduction

In production systems subject to process queue-time constraints, an upper bound is imposed on the total time elapsed between specific manufacturing steps to preserve product quality (Wu et al., 2016). In semiconductor manufacturing, for example, if the time constraint is not satisfied, wafers may experience quality degradation, leading to scrap or the need for costly rework (Klemmt and Mönch, 2012). When a lot reaches the starting point of a time-constrained processing sequence, a decision must be made on whether to release it and activate the corresponding time constraint (May et al., 2024). In practice, this decision is often complex and time-critical and is frequently taken manually based on operator experience and heuristic rules (Lima et al., 2017b). Most of the existing literature addresses time constraints at the scheduling level or at the dispatching level. Only a few studies address time constraints at the capacity planning level. Recently, Mastrangelo et al. (2024) introduced a Markovian-based analytical model for the performance evaluation of time-constrained production systems, incorporating hedging point policies that regulate the activation and deactivation of upstream machines depending on the buffer level. Hedging point policies are a well-established class of threshold-based production control strategies used to regulate production rates in stochastic manufacturing systems (Gershwin, 1994). In dynamic and stochastic production environments, analytical performance evaluation models are essential to capture the complex interactions between production variability and control policies (Carabelli et al., 2026). However, in previous work the control policy is limited to switching upstream machines on and off depending on the buffer level, while the throughput of parts violating the time constraint is estimated using a simplified average-based approximation. This study introduces an analytical performance evaluation model for two-stage production systems subject to queue-time constraints. In particular, the proposed approach introduces an explicit analytical formulation for the accurate evaluation of queue-time violations based on a transient Markovian model. This formulation is coupled with a steady-state Markovian analytical model of a two-stage production system with parallel machines and an intermediate buffer. The steady-state model incorporates generalized threshold-based control points that represent hedging point policies regulating upstream capacity. The proposed control points incorporate not only the buffer level but also information about the operational conditions of the downstream stage. Moreover, the control action is not limited to switching upstream machines on or off but can also regulate their production rate, enabling a more flexible modulation of upstream capacity. As a result, the model enables the efficient identification of optimal hedging points by explicitly accounting for both the throughput of conforming parts and the throughput of parts that violate the time constraint.

2 Methodology

A two-stage production system is considered, composed of an upstream stage, a downstream stage, and an intermediate buffer of size N. The upstream stage operates in a finite set of states U. A generic upstream state $S_u \in U$ is associated with production rate $\mu(S_u)$. Similarly, the downstream stage operates in a finite set of states D, where a generic state $S_d \in D$ is characterized by production rate $\mu(S_d)$. Both stages are modeled as continuous-time, discrete-state Markov chains. The overall system state is therefore described by the triple (x, S_u, S_d), with x representing the number of parts in the buffer. A queue-time constraint T_q is imposed between the completion of the upstream processing stage and the start of the downstream stage. For joint states in which the upstream production rate exceeds the downstream processing rate, the buffer level x tends to increase. As the Work in Progress (WiP) level grows, the waiting time experienced by released parts also increases, which may lead to violations of the queue-time constraint. To mitigate this effect, a hedging point $H(S_d)$ is defined for each downstream state S_d, corresponding to a threshold level of the intermediate buffer. When the buffer level reaches this threshold, the upstream stage is regulated so that its production rate matches the downstream processing rate. If the system operates above this threshold, the upstream stage is temporarily halted until the buffer level decreases back to $H(S_d)$.

The steady-state behavior of the system is evaluated using a threshold-based analytical model based on Tolio and Ratti (2018). The model is adapted to incorporate the control logic induced by the hedging point policies. Given the upstream and downstream Markovian states and the corresponding hedging points $H(S_d)$, the analytical formulation provides the state-based internal probabilities $f(x, S_U, S_D)$ and the state-based threshold probabilities $\pi(x^*, S_U, S_D)$, where x^* denotes a generic threshold level. These quantities characterize the steady-state behavior of the controlled production system and allow the total system throughput to be computed directly as in Equation 1.

$$th_{tot} = \sum_{S_u} \sum_{S_d} \mu_d(S_d) \left[\int_0^N f(x, S_u, S_d)dx + \pi(x^*, S_U, S_D)\right] \tag{1}$$

From the steady-state solution it is possible to compute the throughput of parts released into the buffer for each system state by discretizing the buffer level as in Equation 2 and Equation 3.

$$th_r(x_i, S_u, S_d) = \int_{x_i}^{x_i+\Delta x} f(x, S_u, S_d)\mu(S_u) \tag{2}$$

$$th_r(x^*, S_u, S_d) = \pi(x^*, S_u, S_d)\mu(S_u) \tag{3}$$

For each combination (x_i, S_u, S_d), the probability that a released part is processed by the downstream stage within the queue-time limit T_q must then be evaluated. To this end, a transient analytical model is developed to compute the probability $p_v(x_i, S_u, S_d)$ that the waiting time experienced by a part violates the queue-time constraint. The formulation is based on the framework for computing first passage times in Markov-modulated fluid models as in Scheinhardt et al. (2005), where the transient evolution of the system is obtained by solving the Kolmogorov forward equations. From the transient model it is also possible to compute the probability that a part violates the queue-time constraint even when scrapping of parts ahead in the buffer occurs, denoted as $p_s(x_i, S_u, S_d)$. By accounting for this effect, the final expression for the violation throughput can be derived as follows:

$$th_{violation} = \sum_{S_u} \sum_{S_d} \sum_{x_i} th_r(x_i, S_u, S_d) p_v(x_i, S_u, S_d) p_s(x_i, S_u, S_d) \tag{3}$$

An overview of the model is shown below in Figure 1.

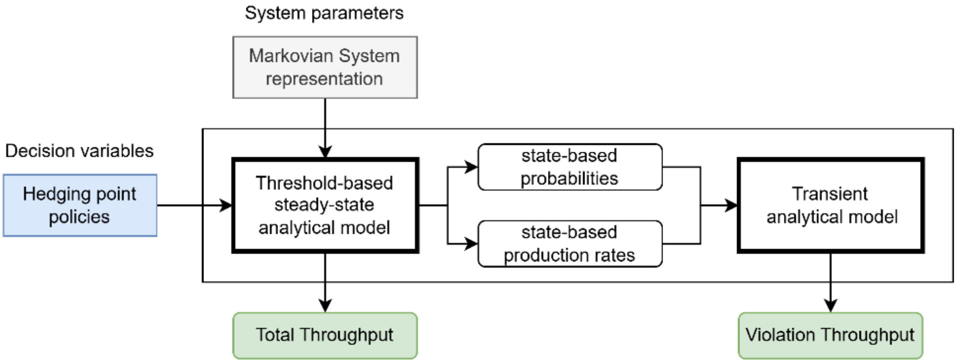


Figure 1. Model Overview

3 Conclusions

This work introduces an analytical framework for the performance evaluation of two-stage production systems subject to queue-time constraints. The proposed model enables the complete analytical evaluation of system performance and allows rapid performance assessment, making it suitable for the optimization of hedging control policies. In particular, the framework enables the identification of hedging points that balance the cost associated with violation throughput and the profit generated by the throughput of conforming parts. Adopting conservative hedging policies may reduce the number of parts violating the queue-time constraint, but may also increase the risk of starvation in the downstream stage, thereby reducing overall system productivity. Identifying an optimal trade-off between these conflicting effects therefore becomes essential for effective system control. The proposed analytical framework provides a systematic tool to support this decision-making process.

Acknowledgements

This project has received funding from the European Climate, Infrastructure and Environment Executive Agency under grant agreement No. 101137954. UK Participants are supported by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee.

References

Carabelli M., Magnanini M.C., Djurdjanovic D., Tolio T., An approximate analytical method for the performance evaluation of semiconductor front-end fabrication with model-based inspection and rework policies in process control, *Journal of Manufacturing Systems*, 86 (2026), 456–471.

Gershwin, S. B. (1994). Manufacturing systems engineering.

Klemmt, A., & Mönch, L. (2012). Scheduling jobs with time constraints between consecutive process steps in semiconductor manufacturing. *Proceedings of the Winter Simulation Conference*, 1–10.

Lima, A., Borodin, V., Dauzère-Pérès, S., & Vialletelle, P. (2017b). A decision support system for managing line stops of time constraint tunnels. In *2017 28th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)* (pp. 309–314).

Mastrangelo, M., Magnanini, M. C., & Tolio, T. A. (2023). Control policy for production capacity modulation with waiting-time-constrained work in process. In *Selected Topics in Manufacturing* (pp. 159–175).

May, M. C., Oberst, J., & Lanza, G. (2024). Managing product-inherent constraints with artificial intelligence: production control for time constraints in semiconductor manufacturing. *Journal of Intelligent Manufacturing*, 35, 4259–4276.

Scheinhardt, W., Van Foreest, N., & Mandjes, M. (2005). Continuous feedback fluid queues. *Operations Research Letters*, 33(6), 551–559.

Tolio, T. A., & Ratti, A. (2018). Performance evaluation of two-machine lines with generalized thresholds. *International Journal of Production Research*, 56(1–2), 926–949.

Wu, C.-H., Chien, W.-C., Chuang, Y.-T., & Cheng, Y.-C. (2016). Multiple product admission control in semiconductor manufacturing systems with process queue time constraints. *Computers & Industrial Engineering*, 99, 347–363.

Why should we care about the retrial time distribution?

Ömer Schmiel
Chair of Production Management, University of Mannheim, Mannheim, Germany, oemer.schmiel@uni-mannheim.de

Benjamin Legros
SKEMA Business School, Université Côte d'Azur, Paris, France, benjamin.legros@skema.edu

Raik Stolletz
Chair of Production Management, University of Mannheim, Mannheim, Germany, stolletz@uni-mannheim.de

Call centers face challenging capacity planning decisions due to the stochastic variability and time-dependency of arrival times. Although various approaches rely on queueing theoretical formulas, the phenomenon of retrials where some customers leave the queue and call back later, is often ignored. This project focuses on the impact of retrial time distribution in call centers. Based on an empirical data analysis, we show that retrial times in call centers are not exponentially distributed. We develop a novel approach to approximate various performance measures for queueing systems with generally distributed retrial times: the stationary coupled flow (SCF) approximation. We present the basic ideas of this approach, as well as analytical and numerical findings on how retrial time distribution impacts performance measures and staffing decisions.

From Shelf to Fulfill: An Economic Study of Store Layout and Robotic Picking in Omnichannel Retail

Sebastian KOEHLER (1), Matthias KLUMPP (2)
(1) Institute for Material Handling and Logistics, Karlsruhe Institute of Technology, Karlsruhe, Germany, (2) Politecnico di Milano, Milan, Italy sebastian.koehler@kit.edu, matthias.klumpp@polimi.it

Recent omnichannel developments are prompting many retailers to re-think their business and operational distribution strategies. Services such as buy-online-pick-up-in-store (BOPS) and buy-online-ship-from-store (BOSS) are becoming increasingly important and prevalent, even for everyday products like groceries or drugstore items. As these services intensify the focus on in-store logistics operations, optimizing fulfillment processes becomes critical. However, since retail stores are not designed for order picking and fulfillment per definition, a fundamental redesign of in-store operations is required. This includes potential process automation steps or store layouts adjustments. In this regard, we analyze the economic impact of two specific operational levers: First, we assess the introduction of a dedicated fast-forward area in a store backroom with respect to throughput capacity and total fulfillment costs. Second, we evaluate the cost-effectiveness of deploying picking robots that operate alongside human workers in the backroom or, potentially, in the shopping area. To this end, we develop an analytical semi-open queuing network (SOQN) model to capture system performance and cost implications under different design configurations. Steady-state probabilities are derived using the matrix-geometric method, and results are validated through an agent-based discrete-event simulation. Furthermore, we present an empirical case study based on real operational data from an example hypermarket store to quantify the economic effects of such design choices. Our results show that both, robotic pickers and/or the introduction of a fast-forward area significantly reduce picking costs, bringing them close to warehouse fulfillment levels. However, the cost-optimal fulfillment configuration varies by store, making a store-specific evaluation necessary for real-world application.

Key words: In-store pickingStore designRobotics in RetailSemi-Open Queuing Networks

1. Introduction

The shift from traditional brick-and-mortar retail to omnichannel services such as buy-online-pick-up-in-store (BOPIS) and ship-from-store (BOSS) places increasing pressure on in-store order fulfillment. Retail stores are not designed for logistics-intensive operations, yet must handle growing online demand under tight margins and persistent labor shortages. While stores are well positioned to meet short customer lead times, order fulfillment must be performed in complex environments shared with in-store customers and competing staff responsibilities. Prior research shows that layout adaptations, such as fast-forward picking areas in the backroom, and the deployment of picking robots can improve efficiency. However, literature lacks a general, computationally efficient evaluation framework that allows retailers to systematically compare alternative in-store fulfillment designs and operating modes. To address this gap, we propose an analytical Semi-Open Queueing Network (SOQN) model to evaluate store-based order fulfillment systems. The model assesses the impact of fast-forward areas and hybrid human–robot picking on throughput, operational costs, and human workload. Using this framework, we analyze (i) cost-optimal fulfillment designs under varying online order volumes and (ii) the performance effects of hybrid human–robot operations compared to fully manual systems. The applicability of the approach is demonstrated using real-world data from a European drugstore chain and a a US grocery hypermarket.

2. Literature

While in-store fulfillment can be more economical than warehouse delivery for small, time-sensitive orders (Calzavara et al. 2023, Pazour and Furmans 2023), the literature remains limited. Existing work addresses store-network-level order assignment (Vazquez-Noguerol et al. 2022, Dethlefs et al. 2022), backroom space allocation (Pires et al. 2020), and fast-forward area design (Seghezzi et al. 2022, Koehler and Theilacker 2025). Picker scheduling has been studied for homogeneous (Difrancesco et al. 2021) and heterogeneous workforces (Mou 2022b,a, Dayarian and Pazour 2022), including robotic resources (Bhowmick et al. 2025b,a). However, no analytical model jointly evaluates fast-forward areas and automation to identify cost-optimal fulfillment configurations. We address this gap using Semi-Open Queueing Networks (SOQNs).

3. Model Description

A retail store consists of a shopping area (SA) and a backroom, with an optional fast-forward area (FFA), a warehouse-style picking zone within the backroom. When an online order arrives, items are picked either from the SA or the FFA, batched, and placed in a pick-up wall for BOPS or BOSS fulfillment. We analyze six fulfillment configurations (see Table 1), varying along two design dimensions: (1) the presence and use of a fast-forward area, and (2) the resource mix. Manual configurations (M1–M3) rely solely on human pickers operating consecutively or in dedicated zones; hybrid configurations (H1–H3) additionally deploy robotic pickers. Zoning-based configurations split orders into suborders per area, which must be consolidated afterwards. Depending on the fulfillment configuration, the set of available resource types $i \in I$ may include consecutive human workers (w_{con}), consecutive robots (r_{con}), backroom workers (w_{br}), shopping area workers (w_{sa}), backroom robots (r_{br}), and shopping area robots (r_{sa}). Each configuration is characterized by a binary availability vector $\boldsymbol{\alpha} = (\alpha_{w_{\mathrm{con}}}, \alpha_{r_{\mathrm{con}}}, \alpha_{w_{\mathrm{br}}}, \alpha_{w_{\mathrm{sa}}}, \alpha_{r_{\mathrm{br}}}, \alpha_{r_{\mathrm{sa}}})$, where $\alpha_i = 1$ indicates that resource type i is active and $\alpha_i = 0$ otherwise. The vector $n_c = (n_{w,\mathrm{con}}, n_{r,\mathrm{con}}, n_{w,\mathrm{br}}, n_{w,\mathrm{sa}}, n_{r,\mathrm{br}}, n_{r,\mathrm{sa}})$ defines the number of resources per type in configuration $c \in C$.

Design Dimension	**Fulfillment Configuration c**					
	Manual			**Hybrid human–robot**		
	M1 Without BR	**M2 With BR (consecutive)**	**M3 With BR (zoning)**	**H1 Without BR**	**H2 With BR (consecutive)**	**H3 With BR (zoning)**
Fast-forward area	no	yes	yes	no	yes	yes
Consecutive picking	no	yes	no	no	yes	no
Dedicated picking zones (BR / SA)	no	no	yes	no	no	yes
Humans operating in fast-forward area	–	yes	yes	no	yes	no
Robots operating in fast-forward area	–	–	–	no	yes	yes
Availability vector $\boldsymbol{\alpha}$	$(0,0,0,1,0,0)$	$(1,0,0,0,0,0)$	$(0,0,1,1,0,0)$	$(0,0,0,1,0,1)$	$(1,1,0,0,0,0)$	$(0,0,0,1,1,0)$

Table 1. Manual and hybrid human–robot in-store order fulfillment configurations

4. The Semi-Open Queuing Network

To evaluate throughput capacity and resource utilization, we model each fulfillment configuration as a Semi-Open Queuing Network (SOQN). Depending on the configuration, (sub-)orders arrive at rate λ_i^{eff}, are batched, and processed by n_i resources of type i. Picking times follow a two-phase Coxian distribution (capturing $\text{SCV} \geq 0.5$), while consolidation and drop-off times are exponentially distributed. The resulting SOQN is depicted in Figure 1.

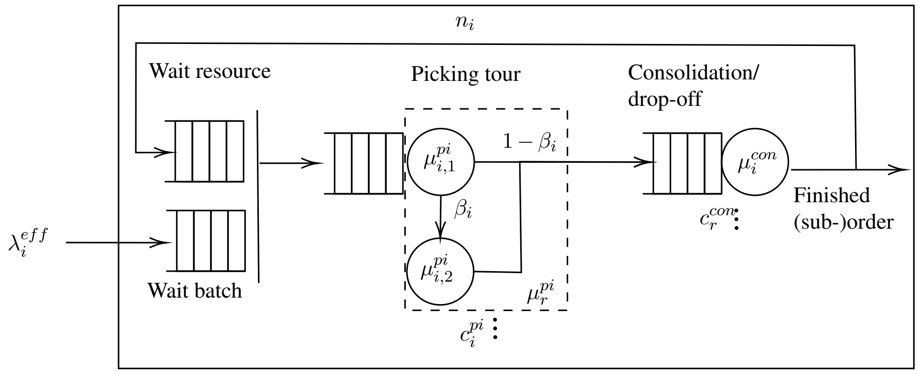


Figure 1. SOQN model of resource type i

The system state is described by the tuple $(s_{o,i}, s_{res,i}, s_{pi,1,i}, s_{pi,2,i}, s_{con,i})$, representing orders at the synchronization point, idle resources, and resources in each processing phase. The SOQN is solved as a quasi-birth–death (QBD) process using the Matrix Geometric Method (Neuts 1994), yielding steady-state probabilities from which key performance indicators, including resource utilization ρ_i, sojourn times ts_i^{pi}, ts_i^{con}, and total throughput TH_i, are derived via Little's Law. The analytical SOQN results are used to identify the cost-optimal fulfillment configuration $c \in \{M1, M2, M3, H1, H2, H3\}$ and the minimum number of resources n_i per type that meets the required throughput. Human workers incur an hourly cost of $c_w = 25$ € (approx. €44,000 annually), while robots incur a fixed annual cost of $c_r = 25{,}000$ €. The optimization problem is:

$$\min_{n,\alpha} C(n,\alpha) = \underbrace{\sum_{i\in\{w_{\text{con}},w_{\text{br}},w_{\text{sa}}\}} \alpha_i\, n_i\, h_i\, c_i}_{\text{Human cost}} + \underbrace{\sum_{j\in\{r_{\text{con}},r_{\text{br}},r_{\text{sa}}\}} \alpha_j\, n_j\, c_j}_{\text{Robot cost}} \tag{1}$$

$$\rho_i \leq \rho_i^{\max} \cdot \alpha_i, \quad \forall i \in I \tag{2}$$

$$TH_i \geq \lambda_i^{\text{eff}} \cdot \alpha_i, \quad \forall i \in I \tag{3}$$

$$n_i \in \mathbb{N}_0,\ \alpha_i \in \{0,1\}, \quad \forall i \in I \tag{4}$$

Utilization is capped at 50% for human workers (who perform tasks beyond picking) and 80% for robots (Eq. 2). Throughput constraints (Eq. 3) ensure the system meets effective order arrival rates.

5. Conclusion

Key findings show that fast-forward areas reduce unit fulfillment costs by up to 13% (drugstore) and 24% (hypermarket), but only beyond a store-specific backroom utilization threshold. Robotic pickers become economically viable above approximately $\lambda^o > 15$ orders per hour, with the cost-optimal deployment zone depending on store format. Under optimal configurations, in-store fulfillment costs approach warehouse-level benchmarks, with reductions of up to 25% (drugstore) and 47% (hypermarket) relative to the manual baseline. Critically, no single configuration dominates across all settings, underscoring the need for store-specific evaluation.

Future work should extend the framework to broader store formats and integrate backroom assignment optimization (Koehler and Theilacker 2025) with fulfillment configuration decisions. Network-level demand balancing across stores (Vazquez-Noguerol et al. 2022) and the analysis of varying robotic capability levels represent further promising directions.

References

Bhowmick, Joyjit, Sebastian Köhler, Gideon Arndt, Georg Fischer, Manmit Padhy, Kai Furmans, Jennifer Pazour. 2025a. Assessing economic and operational feasibility of a designed and lab demonstrated robotic platform for omnichannel logistics. *Computers & Industrial Engineering* 111304.

Bhowmick, Joyjit, Jennifer Pazour, Iman Dayarian. 2025b. Store fulfillment with autonomous mobile robots and in-store customers. *IISE Transactions* 1–18.

Calzavara, Martina, Serena Finco, Alessandro Persona, Ilenia Zennaro. 2023. A cost-based tool for the comparison of different e-grocery supply chain strategies. *International Journal of Production Economics* **262** 108899.

Dayarian, Iman, Jennifer Pazour. 2022. Crowdsourced order-fulfillment policies using in-store customers. *Production and Operations Management* **31**(11) 4075–4094.

Dethlefs, Christian, Manuel Ostermeier, Alexander Hübner. 2022. Rapid fulfillment of online orders in omnichannel grocery retailing. *EURO Journal on Transportation and Logistics* **11** 100082.

Difrancesco, Rita Maria, Isabelle M van Schilt, Matthias Winkenbach. 2021. Optimal in-store fulfillment policies for online orders in an omni-channel retail environment. *European Journal of Operational Research* **293**(3) 1058–1076.

Koehler, Sebastian, Felicia Theilacker. 2025. The backroom assignment problem for in-store order fulfillment. *Computers & Industrial Engineering* 110912.

Mou, Shandong. 2022a. In-store order fulfilment in omni-channel supermarkets with heterogeneous workforce: A bi-objective optimisation approach. *Computers & Industrial Engineering* **171** 108516.

Mou, Shandong. 2022b. Integrated order picking and multi-skilled picker scheduling in omni-channel retail stores. *Mathematics* **10**(9) 1484.

Neuts, Marcel F. 1994. *Matrix-geometric solutions in stochastic models: an algorithmic approach*. Courier Corporation.

Pazour, Jennifer A, Kai Furmans. 2023. Let shoppers shop: why click and collect is not the solution .

Pires, Maria, Ana Camanho, Pedro Amorim. 2020. Solving the grocery backroom sizing problem. *International Journal of Production Research* **58**(18) 5707–5720.

Seghezzi, Arianna, Chiara Siragusa, Riccardo Mangiaracina. 2022. Enhancing in-store picking for e-grocery: an empirical-based model. *International Journal of Physical Distribution & Logistics Management* **52**(4) 301–323.

Vazquez-Noguerol, Mar, Jose Comesaña-Benavides, Raul Poler, J Carlos Prado-Prado. 2022. An optimisation approach for the e-grocery order picking and delivery problem. *Central European Journal of Operations Research* **30**(3) 961–990.

Same-day or next-day? Transparent time-dependent shipment pricing for e-fulfillment

Uta Mohring
Department of Business Administration, University of Zurich, Switzerland, uta.mohring@business.uzh.ch

Melvin Drent
Department of Information Systems and Operations Management, Tilburg University, the Netherlands, m.drent@tilburguniversity.edu

Ivo Adan, Willem van Jaarsveld
School of Industrial Engineering, Eindhoven University of Technology, the Netherlands, {i.adan, w.l.v.jaarsveld}@tue.nl

Key words: Pricing; shipment options; e-commerce; Markov chain; demand management

---

Given the fast growth of e-commerce, fulfilling online orders faster and cheaper becomes more relevant than ever. Offering quick and convenient shipment options is a powerful marketing tool to differentiate from competitors and attract and retain customers. Empirical studies indicate that faster shipment boosts sales and profitability. For example, when promising one day faster shipment, sales for apparel products increase by 1.5% (Fisher et al. 2019). Furthermore, the growing willingness of customers to pay for fast shipment, which has seen a notable rise from 41% in 2023 to 70% in 2024 (Buvat et al. 2024), underscores the customers' impatience and their clear preference for quicker shipment options. In response, e-commerce companies increasingly promise same-day shipment, meaning orders placed until a certain cutoff point in the late afternoon or evening are promised to be shipped the same day to ensure next-day arrival. From an operational perspective, however, fulfilling same-day shipment promises is challenging, especially as the end of the day approaches. Same-day shipment therefore requires a careful coordination between marketing and operations to ensure a smooth and efficient order fulfillment process.

Online retailers typically fulfill orders from so-called fulfillment centers. In these centers, orders are picked from shelves and then prepared and consolidated into batches for shipment. While orders are collected and consolidated continuously throughout the day, the batches are dispatched for shipment only at predefined deadlines – typically at the end of the day – to meet agreed fixed handover times with external parcel delivery companies (e.g., FedEx or UPS) that handle the actual delivery of the orders. Meeting these strict deadlines is crucial for the service quality perceived by customers as missing a deadline by only a few minutes may result in a delay for the customer of at least one day (Ceven and Gue 2017). As e-commerce companies offer increasingly faster service, they risk overpromising their shipment options: fulfillment centers may be unable to process the inflow of orders that require same-day shipment and need to be ready before the next deadline. This risk is particularly high when different shipment options are priced using flat fees, giving customers no incentive to choose next-day over same-day shipment, even when fulfillment center operations would benefit from it.

To mitigate this risk and manage limited capacity effectively, companies need to steer demand so that same-day shipment is selected primarily by customers who value it most and are willing to pay for it. A natural approach is to introduce time- and fulfillment center load-based shipment options and shipment fees, which may provide operational advantages by helping to balance order inflow with available capacity and remaining time until handover. While from an operational perspective appealing, such dynamic shipment policies are difficult to communicate and market effectively in e-commerce (Agatz et al. 2013). Frequent and unexpected price changes are unappealing and appear unfair to customers as they see different prices for the same service at different times without comprehensible reasons for these price differences (Xia et al. 2004). Hence, from the customer's point of view, shipment policies that depend on the load of the fulfillment center are often perceived as opaque and unintuitive.

To address these concerns, we focus on a more practical and customer-focused alternative that has already seen some adoption: transparent time-dependent shipment options and corresponding fees that depend only on the remaining time until the next shipment deadline. These shipment options and fees are consistent each day and can be easily communicated on the ordering platform. Hence, any change in shipment options or fees is immediately visible and predictable to customers. Such a transparent time-dependent shipment policy is operationally relevant, simple to communicate, and perceived as predictable and fair. It thus provides a transparent alternative to opaque dynamic pricing, while still offering a lever to align customer demand with operational capacity.

In this paper, we study this transparent time-dependent shipment policy for online retailers. We do so in the context of products for which the shipment options and shipment fees mostly affect the customer's decision to choose same-day or next-day shipment, and not so much the actual decision to buy the specific product. This makes sense when companies expect little competition or when customers are loyal to the specific products and/or companies. For these settings, we explore how a transparent time-dependent shipment policy can effectively distribute customer demand between same-day and next-day shipment. Our study focuses on cutoff-based shipment options, where all orders placed before a specific cutoff point are eligible for same-day shipment, while orders placed after this point are only eligible for next-day shipment (Mohring et al. 2024). Given the cutoff point decision, fees for same-day and next-day shipment need to be set for each moment in time. Consequently, a transparent time-dependent shipment policy raises the following interrelated questions:

(1) Cutoff point: Until what moment in time should same-day shipment be offered to customers?
(2) Shipment fees: How much should the shipment fees be?
(3) Differentiated fees: When and how should the shipment fees be adapted over time?

We build a parsimonious model that provides answers to the questions posed above. We consider multiple operating cycles, where each cycle consists of a fixed number of time periods and ends with a deadline upon which orders that are due for shipment by this deadline need to be handed over to the parcel delivery company. Throughout each operating cycle, online customers arrive according to a Poisson process. Customers that complete their online transaction before the cutoff point are offered the choice between same-day or next-day shipment together with their corresponding fees (or, equivalently, express shipment or regular shipment). We assume that customers make this choice by trading off the utility that they derive from both shipment options. Customers that arrive after the cutoff point cannot choose same-day shipment, and hence their products are handed over to the parcel company upon the deadline at the end of the following operating cycle. The processing capacity of the fulfillment center in each time period is randomly distributed. If the available capacity is insufficient to process all orders due for shipment in a given operating cycle, then these orders are said to be late and carried over for shipment in the subsequent operating cycle. We are interested in characterizing transparent time-dependent shipment policies that maximize the long-run average profit, i.e., revenue minus cost for late orders. By developing a discrete-time periodic Markov chain model for the steady-state analysis of fulfillment centers described above, we are able to provide closed-form expressions for the relevant performance measures as well as structural properties of the transparent time-dependent shipment policy.

The long-run average profit in our model depends on the entire temporal distribution of same-day orders, as each time-dependent shipment fee influences not only the total demand for same-day shipment but also its timing throughout the day. These interdependencies make the problem analytically challenging and render direct analysis of transparent time-dependent shipment policies in the price domain intractable. Our analysis therefore hinges on a novel problem-domain transformation that represents each policy through the cumulative demand profile it induces for same-day shipment. This transformation forms the basis for our main results: it (i) yields a supermodular reformulation that allows the optimal transparent shipment policy to be computed in polynomial time, and (ii) enables structural characterization of optimal transparent shipment policies. In particular, we find that introducing a cutoff point for same-day shipment and replacing a constant fee with a time-dependent fee improve profit. The optimal same-day fees increase monotonically as the deadline approaches.

Beyond these analytical contributions, we compare specific classes of transparent shipment policies to derive simple, implementable guidelines for e-commerce companies. Our numerical analysis demonstrates that a policy with a two-level fee structure achieves near-optimal performance while remaining fully transparent and easy to communicate to customers. These results show that e-commerce companies can derive substantial benefits from introducing transparent time-dependent shipment options and fees, balancing operational efficiency with customer trust.

The full working paper is available on arXiv: https://arxiv.org/abs/2307.10119.

References

Agatz, N., A. M. Campbell, M. Fleischmann, J. Van Nunen, M. Savelsbergh. 2013. Revenue management opportunities for internet retailers. *Journal of Revenue and Pricing Management* **12**(2) 128–138. doi: 10.1057/rpm.2012.51.

Buvat, Jerome, K V J Subrahmanyam, Sumit Cherian, Siva Chidambaram. 2024. What matter to today's consumer? 2025 consumer behavior trackter for the consumer product and retail industries. Capgemini Research Institute, Available at https://www.capgemini.com/insights/research-library/what-matters-to-todays-consumer-2025/ checked on 5/28/2025.

Ceven, E., K.R. Gue. 2017. Optimal wave release times for order fulfillment systems with deadlines. *Transportation Science* **51**(1) 52–66. doi:10.1287/trsc.2015.0642.

Fisher, Marshall L, Santiago Gallino, Joseph Jiaqi Xu. 2019. The value of rapid delivery in omnichannel retailing. *Journal of Marketing Research* **56**(5) 732–748. doi:10.1177/0022243719849940.

Mohring, Uta, Christoph Jacobi, Kai Furmans, Raik Stolletz. 2024. Managing cutoff-based shipment promises for order fulfilment processes in warehousing. *OR Spectrum* **46**(2) 513–543. doi:10.1007/s00291-023-00739-7.

Xia, Lan, Kent B Monroe, Jennifer L Cox. 2004. The price is unfair! a conceptual framework of price fairness perceptions. *Journal of marketing* **68**(4) 1–15. doi:10.1509/jmkg.68.4.1.42733.

Energy-efficient production control of a production/inventory system with Deep Q-Networks

Duru Balpınarlı
Department of Industrial Engineering, Özyeğin University, Istanbul, Türkiye, email: duru.balpinarli@ozu.edu.tr

Anıl Doğru
Department of Computer Science, Özyeğin University, Istanbul, Türkiye, email: anil.dogru@ozu.edu.tr

Oktay Karabağ
College of Engineering, İzmir University of Economics, İzmir, Türkiye, Email: oktay.karabag@ieu.edu.tr

Siamak Khayyati
HEC Liege, the business school of University of Liege, Liege, Belgium, email: s.khayyati@uliege.be

Barış Tan
Industrial Engineering and Business Administration, Özyeğin University, Istanbul, Türkiye, email: baris.tan@ozyegin.edu.tr

Manufacturing systems can improve energy efficiency by dynamically adjusting production and energy modes based on demand and system conditions. A key challenge is determining optimal machine mode transitions (e.g., producing, idling, or shutting down) to maintain production and minimize energy consumption, particularly under uncertainty in demand and processing times. This study explores a make-to-stock production/inventory system where a machine operates in multiple modes, each with different costs and transition delays, where threshold-type policies have been shown to perform well. Although these policies offer structured decision rules, they struggle with adapting to complex and dynamic environments due to processing time and stochastic modelling challenges. To overcome these challenges, we propose a reinforcement learning-based approach using Deep Q-Networks (DQN) to automate machine operation management adaptively. To benchmark the performance of the system, we model the system as a network of two time-dependent queues connected to each other and analyze them using pre-trained machine learning models. Our method discovers policies that maximize revenue while maintaining energy efficiency by learning from interactions within a developed simulation environment. Evaluation results demonstrate that our approach adjusts flexibly to changing conditions, offering a more adaptive and automated solution than static decision rules.

Keywords: Energy control; Deep Q-Networks; Production Control

1 Introduction

Energy efficiency has become a major priority in manufacturing due to increasing pressure to reduce both environmental impact and energy-related operating costs. This is particularly important because the industrial sector accounts for approximately 50% of global primary energy consumption and 38% of total greenhouse gas emissions (Tan et. al, 2023). Within this broader context, manufacturing processes are often highly energy intensive. For example, steel manufacturing alone represents about 6% of global industrial energy consumption, with nearly 30% of that energy used for heating processes (Tan et. al, 2024). Similar energy-intensive patterns are observed across industries, including automotive, chemical, textile, and metalworking, in operations such as refining, casting, and dyeing. These figures highlight the substantial potential for improving both economic and environmental performance through more effective production and energy control policies. Reinforcement learning has been investigated as a method for energy control in production systems e.g. to control the energy consumption for a cell of homogeneous parallel machines (Loffredo et. al, 2023). In this work, we aim to use Deep Q-Networks to control a production/inventory system with stochastic demand and production processes.

2 Problem definition

In this context, this study considers the joint production and energy mode control problem in a make-to-stock manufacturing system subject to stochastic demand. The system consists of a single machine that produces inventory items and can operate in multiple energy modes. Customer demand arrives randomly over time and is satisfied immediately when inventory is available; otherwise, it is lost, and a penalty is incurred. The machine can operate in four distinct modes: on, idle, off, and warm-up. In the on mode, the machine is actively producing and consuming energy. In the idle mode, it is not producing but remains ready to resume production while still consuming some energy. In the off mode, the machine is shut down and consumes no energy. The warm-up mode captures this transition from off to an operational state, during which both time and energy are expended. Each mode is associated with a different operating cost, and switching between modes incurs an additional cost reflecting the energy and time required for the transition.

The objective is to determine, at each decision epoch, whether the machine should continue producing, remain idle, or be turned off, while balancing production benefits against energy and switching costs under demand uncertainty. These decisions must be made dynamically, since system performance depends critically on the interaction between the current inventory level, the machine's operating mode, and future demand. Poorly timed transitions may result in either excessive energy consumption or insufficient inventory, leading to lost sales. The goal is therefore to develop a control policy that adapts to the evolving state of the system and improves long-run operational performance.

3 Mathematical Model

If the complete stochastic model of the system were fully available, the problem could, in principle, be formulated as a Markov decision process and solved using the classical dynamic programming approach. In that case, methods such as value iteration or policy iteration could be employed to derive an optimal control policy through the Bellman equations. In the present setting, however, as in many practical applications, this requirement may be restrictive, because the system dynamics are driven by stochastic demand and production processes and may involve continuous-time features that are difficult to represent explicitly. These challenges make exact dynamic programming difficult to implement in practice and therefore motivate the use of reinforcement learning, which can learn control policies directly from interaction with the environment without requiring explicit transition probabilities, that is, in a model-free setting (Watkins and Dayan, 1992).

Among reinforcement learning methods, temporal-difference learning, which was originally introduced by Sutton (1988), provides a particularly suitable foundation for this problem. Rather than relying on a fully specified model, temporal-difference methods update value estimates directly from observed transitions. In particular, Q-learning estimates the action-value function through the temporal-difference error:

$$\delta = r + \gamma Q(s', a') - Q(s, a),$$

where r denotes the observed reward, γ is the discount factor, and $Q(s, a)$ is the estimated value of taking action of a in state s. The corresponding update rule is given by

$$Q(s, a) \leftarrow Q(s, a) + \alpha\big(r + \gamma Q(s', a') - Q(s, a)\big),$$

where α is the learning rate. While standard Q-learning is effective when the state-action space is discrete and moderate in size, it relies on a tabular representation and therefore becomes impractical in settings with large or continuous state spaces.

This limitation is central in the present problem. An effective control policy must account not only for the machine's current operating mode and the current inventory level, but also for time-related information describing the evolution of the system. Such variables may include, for example, the timing of recent or upcoming demand arrivals and the progress of production-related events. Representing this richer state space through a finite Q-table is generally not feasible due to the curse of dimensionality. For this reason, we adopt the Deep Q-Network (DQN) approach for the production and energy mode control problem. DQN replaces the tabular action-value function with a neural-network approximation, thereby allowing the method to accommodate high-dimensional or continuous state representations and making it particularly suitable for the problem considered here, especially when the state space must incorporate the timing of recent demand arrivals and the progress of production-related events. At the same time, the action space remains discrete, since the controller selects among a finite set of operational decisions such as producing, idling, or shutting the machine down. This combination makes DQN a natural methodological choice for the problem under consideration.

The design of the observation space is therefore critical. The state representation provided to the neural network must contain the information necessary for effective decision making while remaining sufficiently structured for stable learning. In our formulation, the observation vector includes the current machine mode, the current inventory level, and relevant temporal variables characterizing the system state. The machine mode is encoded categorically, for example through one-hot encoding, while the inventory level is represented numerically. The time-related variables are normalized to improve generalizability and to ensure that the resulting representation remains comparable across different operating conditions. This observation structure allows the learning algorithm to capture both the operational status of the machine and the temporal features required to anticipate future demand and production opportunities.

The neural network takes this observation vector as input and outputs an estimated Q-value for each action. However, in this problem, rewards are not determined solely by the immediate action; rather, they often reflect the accumulated consequences of earlier decisions. For example, the current cost or benefit may depend on a machine mode selected in the recent past or on an earlier decision that affected the current inventory position. As a result, directly associating the observed reward with the current action alone may lead to unstable learning and inaccurate value estimation. To address this issue, we employ a Dueling DQN architecture, which decomposes the action-value function into a state-value component and an advantage component:

$$Q(s,a) = V(s) + A(s,a).$$

This allows for the contribution of the current state being separated from the incremental effect of a specific action. In other words, the state-value function captures the value inherent in the system's current condition, including the delayed effects of past decisions, while the advantage function measures the relative benefit of each available action in that state. By structuring the value approximation this way, the model can better distinguish state-dependent rewards from action-dependent improvements, thereby improving learning stability and Q-value estimation accuracy.

$$Q(s,a) = V(s) + A(s,a) - \frac{1}{|\boldsymbol{A}|}\sum_{a'} A(s,a')$$

To assess the performance of the system considered here with general inter-event times, we consider using pre-trained machine learning models to predict the behavior of the system modeled as a time-dependent network of queues. In particular, the finished goods buffer and the lost sales buffer form two queues, and although we consider the system in steady state, the control policy changes the machine's status, so the incoming stream to the inventory is time dependent. Here, the performance of each node in each mode is modeled using a pre-trained machine learning model, and these models are then connected to analyze the system as a whole.

4 Conclusion

Overall, the proposed reinforcement learning framework provides a practical approach for deriving production and energy mode control policies in stochastic manufacturing systems when explicit model-based optimization is difficult to implement. Moreover, the learned policy exhibits a threshold-type structure that closely aligns with the policy pattern reported in Tan et al. (2024), suggesting that the proposed method is able to recover the essential decision logic of the underlying control problem while remaining applicable in more complex settings.

Acknowledgements

This research has been supported by funding from the Belgian National Fund for Scientific Research (FNRS).

References

Loffredo, A., May, M. C., Schäfer, L., Matta, A., & Lanza, G. 2023. Reinforcement learning for energy-efficient control of parallel and identical machines. CIRP Journal of Manufacturing Science and Technology, 44, 91-103.

Sutton, R. S. 1988. Learning to predict by the methods of temporal differences. Machine learning, 3(1), 9-44.

Tan, B., Karabağ, O., Khayyati, S. 2023. Production and energy mode control of a production-inventory system. European Journal of Operational Research, 308(3), 1176-1187.

Tan, B., Karabağ, O., Khayyati, S. 2024. Energy-efficient production control of a make-to-stock system with buffer-and time-based policies. International Journal of Production Research, 62(16), 5809-5827.

Watkins, C. J., Dayan, P. 1992. Q-learning. Machine learning, 8(3), 279-292.

CO2 emission reduction of serial production lines through energy saving using reinforcement learning

Ahmad REZAEE ELYATOO
Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38000 Grenoble, France, email: ahmad.rezaee-elyatoo@grenoble-inp.fr

Abdourahim SYLLA
Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38000 Grenoble, France, email: abdourahim.sylla@grenoble-inp.fr

Maria DI MASCOLO
Univ. Grenoble Alpes, CNRS, Grenoble INP, G-SCOP, 38000 Grenoble, France, email: Maria.Di-Mascolo@grenoble-inp.fr

Global warming makes energy-efficient industrial production essential. In serial production lines, controlling when machines are switched on or off can reduce energy use while maintaining throughput, but this is challenging due to uncertainties like variable processing times and failures.

This problem, known as Machine State Control, is modeled as a sequential decision-making problem. A model-free Reinforcement Learning approach is proposed to learn effective control policies through interaction with a simulation model. The goal is to determine whether reinforcement learning agents can significantly reduce energy consumption while maintaining production production throughput in automated serial production lines.

Keywords: Energy Saving, Machine State Control, Reinforcement Learning.

Global warming is driving countries to adopt ambitious climate targets aimed at reducing greenhouse gas emissions. As industry is responsible for a significant share of global CO2 emissions, improving energy efficiency in energy-intensive industries has become a key priority. Consequently, reducing energy consumption while maintaining production throughput has emerged as a key factor for sustainable production.

One approach to address this challenge in serial production lines is to dynamically and continuously control machines' state, and to make the right decision between "switching on" and "switching off" machines. This decision is subject to various uncertainties, including stochastic processing times, random time between failures, and dynamically changing buffer levels. The main challenge is to achieve energy savings while preserving production throughput.

Modeling these stochastic and dynamic behaviors using classical optimization methods is difficult due to the inherent uncertainty and variability within production lines. Additional phenomena such as blocking, starvation, and shifting bottlenecks further increase system complexity. Developing accurate analytical models that capture these dynamics typically requires deep domain knowledge, extensive data extraction, and continuous model updates, which can be difficult to maintain in real production environments.

In this paper, this problem, known in the literature as Machine State Control problem, is formulated as a sequential decision-making problem that aims to balance energy savings and production throughput. To address this challenge, a model-free Reinforcement Learning (RL) approach is proposed to learn effective control policies through interaction with a simulated production line. The proposed method is evaluated using a discrete-event simulation of a serial production line that captures key dynamics of real manufacturing systems and is benchmarked against an Always-On policy. The objective is to investigate whether RL agents can achieve significant reductions in energy consumption while maintaining production throughput in automated serial production lines.

Acknowledgements

This work was supported by a French government grant managed by the Agence Nationale de la Recherche (ANR) under the France 2030 program: PEPR Spleen's DCARBO project, reference 22-PESP-0002.

Throughput Modeling for Robotic Compact Storage Systems via Hybrid Queueing Analytics and Machine Learning

Christophe SENGER, Timo LEHMANN, Laura ALTMANN, Alexander ERNST, Kai FURMANS
Institute for Material Handling and Logistics, Karlsruhe Institute of Technology, Karlsruhe, Germany, christophe.senger@kit.edu, timo.lehmann@kit.edu, laura.altmann@student.kit.edu, alexander.ernst@kit.edu, kai.furmans@kit.edu

Robotic Compact Storage and Retrieval Systems (RCS/RS) are gaining popularity for their exceptional space efficiency, which is achieved by stacking bins and utilizing autonomous robots. A significant operational drawback, however, is the need for time-consuming 'reshuffling' operations when a target bin is not accessible at the top of a stack. This characteristic complicates the development of accurate analytical throughput models, forcing system planners to choose between time- and cost intensive simulations and oversimplified analytical models that lack accuracy. To resolve this dilemma, we propose a hybrid throughput model that synergistically combines a closed queueing network for the deterministic elements of robot cycle times with a machine-learning surrogate model for the complex, state-dependent reshuffling components. This machine-learning element was trained on a comprehensive dataset of 2,000 simulated warehouse configurations. Validation against detailed simulations demonstrates that our hybrid model achieves high accuracy while enabling the planning and analysis of large, complex RCS/RS in a matter of seconds, providing a powerful tool for rapid system design and optimization.

Key words: RCSRS; Throughput capacity; travel time model; performance estimation; machine learning; surrogate model

---

1. Introduction

Robotic compact storage and retrieval (RCS/R) systems are an important warehousing technology due to their high storage density and scalable throughput. Since bins are stacked vertically and retrieval bins are often blocked by other bins, reshuffling is required, which makes throughput difficult to predict. For storage planners, throughput is the key design measure, yet estimation still relies mainly on simulation because analytical models are scarce. This creates a need for fast and reliable planning methods, especially under uncertain demand and product access patterns. To address this gap, we propose a hybrid analytic and machine-learning approach for throughput estimation. The system is modeled as a closed queueing network, while state-dependent cycle-time components are learned from simulation data. This combines the speed of analytical models with the realism of simulation and enables accurate, fast throughput estimation.

2. Literature

Research on throughput estimation for robotic compact storage and retrieval (RCS/R) systems is still limited. Existing work can be grouped into simulation-based, analytical, and machine-learning (ML) -based approaches. Simulation studies such as Trost et al. (2023), Chen et al. (2022), Galka and Scherbarth (2021), and Yue and Smith (2023) capture system behavior in detail, but are computationally expensive. Analytical models are less common and often rely on simplifying assumptions that limit their representation of real system operations, as shown by Lehmann and De Koster (2026), Tutam et al. (2024), Trost and Eder (2024), and Zou et al. (2018). Machine learning approaches remain scarce; Wang et al. (2024), for example, use reinforcement learning for reshuffling decisions in overhead RCS/R systems. Overall, the literature reveals a gap between detailed but slow simulation models and fast but simplified analytical approaches.

3. Model Description

Following Lehmann and De Koster (2026), the throughput capacity is calculated as

$$TP = \frac{3600 \cdot PS \cdot K}{t_{cycle} + t_{pick} + t_{wait}}, \tag{1}$$

with PS the number of pick stations, K the number of robots per pick station, t_{pick} the average processing time at the pick station, t_{cycle} the robot cycle time in the storage system, and t_{wait} the expected waiting time in front of the pick station. To determine t_{wait}, the system is modeled as a closed queueing network (CQN) with two nodes: an infinite-server node representing the robot cycle in the rack and a single pick-station server with queue. Once t_{cycle} is known, mean value analysis can be applied to the closed queueing network to obtain t_{wait}.

The robot cycle time t_{cycle} is composed of a sequence of independent movements and handling operations. In a standard dual-command cycle, a robot (i) travels from the pick station to a storage location carrying a bin, (ii) stores the bin, (iii) travels to the next retrieval location, (iv) retrieves the requested bin, and (v) returns it to the pick station. If the target (retrieval) bin is blocked, additional reshuffle operations are required: the robot repeatedly retrieves blocking bins, relocates them to alternative storage locations, and returns to the retrieval stack; after the target bin is retrieved, the relocated bins are moved back to their original stack. Several components of t_{cycle} can be computed analytically with high precision, e.g., the handling time for pick/drop operations and travel times between stations and storage/retrieval locations (see Lehmann and De Koster (2026)). In contrast, other components are strongly state-dependent and therefore difficult to model in closed form. In this work, such terms, in particular the number of reshuffles and stack access times (lift down/up times), are approximated using machine-learning models trained on simulation data.

For the smoother and more deterministic state-dependent components, we use standard feed-forward multilayer perceptrons (MLPs) to predict the lift travel time during a storage operation, the horizontal travel time from a storage channel to a retrieval channel, and the lift travel time during a retrieval operation. For components with higher variability and more complex conditional behavior, we employ mixture density networks (MDNs) Bishop (1994) with a transformer backbone Vaswani et al. (2017). Specifically, Gaussian transformer-MDNs are used to model the lift travel time during a reshuffle retrieval and the number of reshuffles per retrieval, while a Laplace transformer-MDN is used to model the travel time between a retrieval channel and a reshuffle destination. Unlike deterministic regressors, the MDN outputs a conditional mixture distribution rather than a single point estimate, which better captures noisy or potentially multi-modal outcomes and yields more robust cycle-time estimates. Overall, the learned predictors replace the computationally expensive state-dependent elements of the robot cycle time, while the queueing model maps the resulting cycle-time estimate to throughput.

4. Evaluation

For model evaluation, the dataset comprises 4,539 instances, of which 2,000 were used for training. The training set was generated by Latin Hypercube Sampling (LHS), ensuring a well-distributed coverage of the input space and thus supporting model generalizability. As a benchmark, we additionally trained a Gaussian Mixture Density Network (MDN) to directly predict the full cycle time. The hybrid approach consistently outperforms this direct prediction benchmark across all five validation rounds. It achieves lower throughput deviations (0.89%–0.97% vs. 1.44%–1.66%) and tighter upper bounds of the 95% confidence intervals (5.12%–5.50% vs. 6.57%–7.70%). These results show that combining analytical structure with data-driven approximation improves both accuracy and robustness. The results indicate that decomposing the cycle time into analytically tractable and separately learned state-dependent components is more accurate and robust than predicting the complete cycle time with a single neural model.

5. Conclusion

Overall, the proposed hybrid CQN-ML model closes the gap between slow high-fidelity simulation and fast but simplified analytics, delivering accurate and robust throughput estimates.

References

Bishop, Christopher M. 1994. Mixture density networks .

Chen, Xingwei, Peng Yang, Zixin Shao. 2022. Simulation-based time-efficient and energy-efficient performance analysis of an overhead robotic compact storage and retrieval system. *Simulation Modelling Practice and Theory* **119** 102560.

Galka, Stefan, Christoph Scherbarth. 2021. Simulationsbasierte Untersuchung der Grenzproduktivität von Robotern in einem AutoStore-Lagersystem - simulation-based analysis of the marginal productivity of robots in an autostore storage system. *Simulation in Produktion und Logistik* 197.

Lehmann, Timo, René De Koster. 2026. Digging deep: Finding and maximizing the throughput capacity of multideep storage systems. *Transportation Science* .

Trost, Philipp, Michael Eder. 2024. An analytical approach for the performance calculation of an rcs/rs with several picking stations. *The International Journal of Advanced Manufacturing Technology* **134**(1) 627–644.

Trost, Philipp, Georg Kartnig, Michael Eder. 2023. Simulation study of rcs/r-systems with several robots serving one picking station. *FME Transactions* **51**(2) 201–210.

Tutam, Mahmut, Jingming Liu, John A White. 2024. Consideration of skewness in designing robotic compact storage and retrieval systems. *Expert Systems with Applications* **248** 123361.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, Illia Polosukhin. 2017. Attention is all you need. *Advances in neural information processing systems* **30**.

Wang, Rong, Peng Yang, Yeming Gong, Cheng Chen. 2024. Operational policies and performance analysis for overhead robotic compact warehousing systems with bin reshuffling. *International Journal of Production Research* **62**(14) 5236–5251.

Yue, Anni, Stephen L Smith. 2023. Minimizing robot digging times to retrieve bins in robotic-based compact storage and retrieval systems. *arXiv preprint arXiv:2312.05338* .

Zou, Bipan, René De Koster, Xianhao Xu. 2018. Operating policies in robotic compact storage and retrieval systems. *Transportation Science* **52**(4) 788–811.

Discussion on Scalable Reinforcement Learning Framework for Control Agents in Semiconductor Manufacturing

Yavar Yeganeh*, Mahsa Shekari*, Nicla Frigerio, Andrea Matta
Politecnico di Milano
Department of Mechanical Engineering, Politecnico di Milano, Milan, Lombardy, Italy,
yavar.taheri@polimi.it, mahsa.shekari@polimi.it, nicla.frigerio@polimi.it, andrea.matta@polimi.it

Semiconductor manufacturing systems pose a challenging control problem due to their scale, stochasticity, delayed feedback, and tightly coupled long-horizon dynamics. We discuss a deep reinforcement learning framework for multi-objective operational optimization in such environments, with emphasis on its practical implications for fab control. The framework formulates decision-making as a centralized single-agent problem and models system evolution as an event-driven temporal process, allowing control to remain aligned with the structure of semiconductor operations. A key insight is that effective learning depends not only on reward design, but also on the temporal abstraction used to propagate feedback across dense sequences of interdependent actions. Within this formulation, agents can be trained using different model-free algorithms both in offline and online settings. Evaluations on a high-fidelity simulator with industry-realistic data show consistent gains over conventional dispatching heuristics in throughput and utilization. Overall, the discussion highlights scalability, data efficiency, and the importance of problem formulation as the main enablers of reinforcement learning for realistic semiconductor production control.

Key words: Reinforcement Learning; Decision-Making; Optimization; Semiconductor Manufacturing

---

This extended abstract discusses the main lessons emerging from the reinforcement learning framework proposed in Yeganeh et al. (2026). While the journal paper develops the complete framework, algorithms, and evaluation protocol, the present version focuses on interpretation: what the numerical evidence suggests about scalable fab control. Since this extended abstract refers to the original study, Table 1 serves as the main quantitative reference for the discussion by summarizing the most relevant results across training regimes and algorithms.

Case study and performance. The industrial case concerns a front-end wafer fab based on STMicroelectronics Catania within the HiCONNECTS project. The fab is organized into multiple interconnected buildings, areas, and bays, with tools distributed across process families and non-negligible transport times for inter-building moves. Production is lot-based: wafers are grouped into carriers and processed through long, re-entrant routes with hundreds of operations, shared qualified equipment, batching requirements, and different constraints. The product mix is high, with many technologies and products sharing partially overlapping resources, so real-time dispatching decisions must account for WIP distribution, tool status, qualification, setup, processing time, batchability, sequence position, and process-sensitive constraints such as corrosion. The proposed framework achieves consistent improvements in throughput, saturation, and load across both offline and online learning. In the offline regime (Kostrikov et al. 2021), agents trained solely from previously collected data already learn effective control policies, with DQL and CQL (Kumar et al. 2020) providing the strongest results. In the online regime, performance improves further, with SAC achieving the best overall gains over FIFO as a deterministic baseline. These results indicate that the framework is effective and data-efficient, enabling useful policies to be learned from operationally realistic experience. Figure 1 compares the best offline and online agents, DQL and SAC (Christodoulou 2019), respectively. The results show that the learned policies perform well across the three KPIs.

Scalability and practical feasibility. These results are important because they are obtained in a large-scale semiconductor setting with long re-entrant flows, heterogeneous products, and tightly coupled system dynamics. The main contribution is therefore not only higher KPI values, but the

Regime	Agent	Throughput	Saturation	Load
Offline	IQL	14.7 ± 4.4	13.0 ± 3.1	10.6 ± 2.6
	CQL	17.5 ± 5.1	15.7 ± 4.0	12.3 ± 3.3
	DQL	18.0 ± 4.4	16.0 ± 3.5	12.8 ± 2.9
Online	PPO	16.3 ± 5.3	15.3 ± 3.9	12.1 ± 3.2
	DQL	19.7 ± 3.9	16.6 ± 3.2	13.2 ± 2.8
	SAC	20.7 ± 4.2	17.5 ± 3.5	13.8 ± 2.9

Table 1. KPI gains over the FIFO baseline (%) for offline and online agents.

fact that these gains are achieved with a modular and computationally feasible RL design. By combining centralized learning, modular state–action handling, and event-driven temporal abstraction, the framework improves all three KPIs jointly. This is notable because throughput, saturation, and load are strongly coupled in practice. The results therefore support the claim that RL can be scaled beyond localized dispatching rules toward system-level control in realistic fab environments.

Importance of temporal formulation. A central lesson is that success depends strongly on how the learning problem is formulated. Semiconductor manufacturing is characterized by delayed feedback, long horizons, and dense decision sequences, which make standard temporal-difference learning difficult. This is clearly reflected in the temporal-difference comparison. For DQL, the truncated discounted-sequence formulation performs poorly, with only -3.4% throughput gain, whereas the event-aggregated formulation reaches 18.0%. A similar pattern appears for IQL, where truncated discounting yields -5.1% throughput gain, while event aggregation reaches 14.7%. Even the stacked event-wise and event-averaged formulations perform substantially better than the truncated alternative. This shows that reward shaping alone is not enough; the way rewards are grouped and propagated across events is equally important for solving long-horizon credit assignment.

Role of event aggregation. Among the tested formulations, event aggregation is the most important because it aligns the learning signal with the delayed and coupled nature of fab dynamics while retaining event-level structure. Rather than treating each decision independently or collapsing an entire segment too early into a single scalar, it provides an intermediate abstraction between local event feedback and global system response. This matters even more online, where the agent no longer uses event-level shaped rewards and instead learns from segment-level rewards over 360-minute windows containing thousands of actions. The online results suggest that this abstraction is what makes the reward learnable in practice. In this sense, the main challenge is not simply specifying the objective, but constructing a temporal representation that enables objective to be optimized.

Algorithmic insights. The algorithmic comparison in Table 1 suggests several consistent trends. In offline learning, some degree of conservatism is beneficial. CQL and IQL remain competitive with DQL while offering additional protection against value overestimation under fixed data. At the same time, DQL remains attractive because of its simplicity and strong performance. For online, off-policy methods are especially effective: DQL & SAC are more sample-efficient than PPO (Schulman et al. 2017), which is desirable when simulator interaction is limited. SAC achieves the best result, suggesting that entropy-based exploration is particularly suitable in the investigated manufacturing setting. PPO shows promise, but its performance depends strongly on the architecture and advantage formulation. Overall, the results do not point to a universally dominant algorithm; instead, they show different algorithmic families become effective when combined with the tailored event-driven formulation.

Evaluation and model selection. Another important lesson is that offline model selection is non-trivial. Lower temporal-difference loss does not necessarily imply better downstream control performance. The checkpoint analysis supports this directly: for DQL, intermediate checkpoints with lower TD loss do not always outperform later ones in all three KPIs, and performance evolves non-monotonically during training. This makes robust evaluation essential. Aggregating results across 10 industrially realistic instances and 3 seeds reduces the risk of overinterpreting a favorable trajectory from a single scenario and instead evaluates whether the learned policies generalize across temporally separated and operationally diverse conditions.

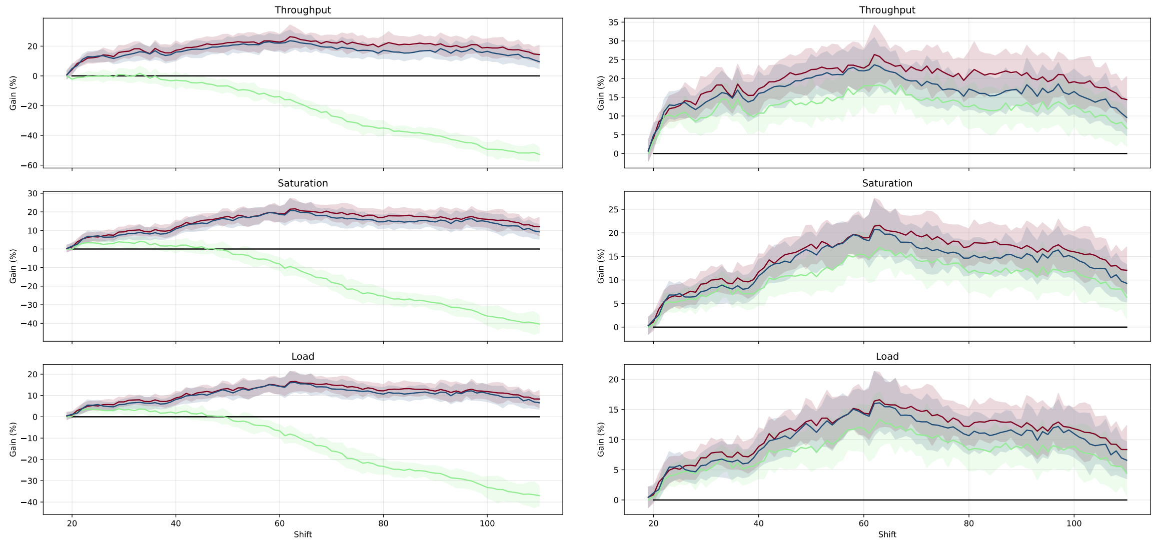


Figure 1. KPI gains (%): Black: FIFO zero-gain baseline. Burgundy: SAC. Blue: DQL. Green: Random (left) and SPT (right).

Importance of encoding In general, state and action representations for fab control can be built using different representation-learning approaches, such as transformers, graph neural networks, generative models, LLM-based encoders, or engineered features. In this study, we used engineered features to keep the framework interpretable, computationally feasible, and focused on the reinforcement-learning formulation. However, this choice may introduce limitations: the accuracy of the learned policy depends on how well the manually designed features preserve the relevant routing, equipment, WIP, setup, constraint, and temporal dependencies of the fab. If important dependencies are omitted or compressed too strongly, the agent may not learn a policy that is effective or is less generalizable under different product mixes, WIP distributions, or operating conditions.

Limitations and conclusion. The study emphasizes a general and modular RL framework rather than exhaustive hyperparameter tuning, and the experiments cover approximately one month of production data. Longer horizons, richer representations, and broader objectives such as due-date compliance remain important directions for future work. Still, Table 1 makes the main conclusion clear: RL can be practically deployed for semiconductor production control when the formulation respects the event-driven, delayed, and tightly coupled structure of fab operations. From this perspective, scalable RL is less about choosing a single best algorithm and more about combining suitable temporal formulations with conservative offline learning and carefully designed online adaptation.

Acknowledgments
The work presented in this paper was supported by HiCONNECTS, which has received funding from Chips Joint Undertaking under grant agreement number 101097296.

References

Christodoulou, Petros. 2019. Soft actor-critic for discrete action settings. *arXiv preprint arXiv:1910.07207* .

Kostrikov, Ilya, Ashvin Nair, Sergey Levine. 2021. Offline reinforcement learning with implicit q-learning. *arXiv preprint arXiv:2110.06169* .

Kumar, Aviral, Aurick Zhou, George Tucker, Sergey Levine. 2020. Conservative q-learning for offline reinforcement learning. *Advances in neural information processing systems* **33** 1179–1191.

Schulman, John, Filip Wolski, Prafulla Dhariwal, Alec Radford, Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* .

Yeganeh, Yavar, Mahsa Shekari, Nicla Frigerio, Daniele Pagano, Andrea Matta. 2026. Scalable reinforcement learning framework for decision-making in semiconductor manufacturing. Under Review.

A Structure-Informed Convolutional Neural Network Surrogate Approach for Integrated Performance Evaluation & Optimization of Flow Lines via Mixed-Integer Programming

Andreas KLOOS
University of Regensburg, Regensburg, Bavaria, Germany, andreas.kloos@ur.de

Justus Arne SCHWARZ
University of Regensburg, Regensburg, Bavaria, Germany, justus.schwarz@ur.de

Simulation is widely used for evaluating stochastic manufacturing systems, due to its adaptable modeling capabilities and the absence of exact or even approximate analytical solutions. However, solving optimization problems with performance evaluation via simulation can be time-consuming, due to the need for many replications and the combinatorial complexity of the problems. To increase the speed of performance evaluation, the existing literature suggests replacing simulation with Artificial Neural Networks pretrained with simulation results. We train a Convolutional Neural Network (CNN) with flow line feature data to predict flowline throughputs for different system configurations. The proposed CNN architecture enforces domain specific structural properties of the underlying system. We then solve the Buffer Allocation Problem (BAP) by linearizing the trained CNN, thereby transforming the non-linear problem into a mixed-integer linear problem (MILP), which is tractable by standard solvers. We introduce a structured sampling strategy for training data generation and investigate the effects of regularization methods on the MILP. Furthermore, we demonstrate the flexibility of our approach by applying it to different production networks and optimization problems.

Key words: Optimization, Manufacturing, Machine Learning

1. Introduction

We propose an integrated approach for solving optimization problems in stochastic manufacturing systems by combining the predictive power of artificial neural networks with the optimization power of mixed integer programming. Machine learning (ML) models excel at approximating complex, functional relationships from data. If these functional relationships are driven by stochastic influences they are typically non-linear. Linearized versions of the ML model can replace these non-linear components in optimization problems. Thereby transforming the non-linear optimization problem into a Mixed-Integer Linear Program (MILP), tractable by standard solvers. This approach was recently introduced by Anderson et al. (2020) and Maragno et al. (2025). However, applications within the manufacturing domain remain scarce (Tremblet et al. 2024, Grimstad and Andersson 2019, Camponogara et al. 2024) and the optimization of stochastic manufacturing systems remains unexplored. Our contribution is threefold: (i) We propose an architecture-aware sampling approach for generating training data, that outperforms generic sampling techniques. (ii) We develop a CNN that is able to make predictions for large production lines while being trained only on short lines.

Moreover, the network architecture guarantees known structural properties of stochastic flow lines. (iii) We demonstrate that a MILP which integrates a linearized version of the CNN solved with standard solvers outperforms the other tested approaches.

2. Illustrative Optimization Problem: Buffer Allocation Problem

We illustrate our approach using the well-studied Buffer Allocation Problem (BAP) for stochastic flow lines (Weiss et al. 2019). The flow line consists of J stations, with processing rates $\boldsymbol{\mu} = (\mu_1, ..\mu_j, ...\mu_J)$ and exponentially distributed processing times. Each station j has an upstream buffer of capacity $\boldsymbol{C} = (C_1 = \infty, C_2, ..., C_j, ..., C_J)$. Buffers mitigate the effects of blocking after service and starving caused by the stochastic processing times within the flow line. We assume the first station is never starved and the last station never blocked. The primary BAP (1) aims to minimize the total buffer capacity (1a) while ensuring that the expected throughput of the flow line $E[TH(\boldsymbol{\mu}, \boldsymbol{C})]$ reaches a certain target throughput th^* (Constraint (1b)) and respects individual limits $C_{j,max}$ for the buffer capacities (Constraints (1c))

$$\text{Minimize: } C_{\text{tot}} = \sum_{j=2}^{J} C_j \tag{1a}$$

$$TH(\boldsymbol{\mu}, \boldsymbol{C})] \geq th^* \tag{1b}$$

$$0 \leq C_j \leq C_{j,max}. \tag{1c}$$

3. Solution approach

We describe the data generation process, the development of the neural network and its integration into a MILP in the Subsections 3.1, 3.2, and 3.3, respectively.

3.1. Date generation

For dataset generation (training, validation and testing), we propose a space-filling sampling approach across the critical parameters that primarily govern flow line behavior. As an critical parameter we identify C_{tot}, which drives the overall mitigation of blocking and starving within the flow line. We further introduce two balance criteria for buffer and processing rate allocations

$$C_{\text{bal}} = \frac{max(\boldsymbol{C}) - min(\boldsymbol{C})}{C_{tot}} \tag{2}$$

$$\mu_{\text{bal}} = \frac{max(\boldsymbol{\mu}) - min(\boldsymbol{\mu})}{\mu_{tot}}, \text{ with } \mu_{tot} = \sum_{j=1}^{J} \mu_i \tag{3}$$

that are also used as critical parameters in the sampling process. The expected steady-state throughput of the sampled flow line configurations is obtained via discrete-event simulation. We compare our proposed sampling approach against standard Sobol Sampling and Latin Hypercube Sampling (LHS) by evaluating the predictive performance of hyperparameter-tuned CNNs trained on each respective dataset.

3.2. Network architecture

Our approach employs a CNN trained on the generated simulation data to estimate the expected throughput of a given buffer allocation. The obtained filters allow the extrapolation to predict long flow lines with CNNs trained on short flow lines. We utilize a Siamese CNN architecture (Bromley et al. 1993) to enforce the structural property of reversibility, obtained by Muth (1979). This guarantees identical throughput predictions for a flow line and its reversed counterpart.

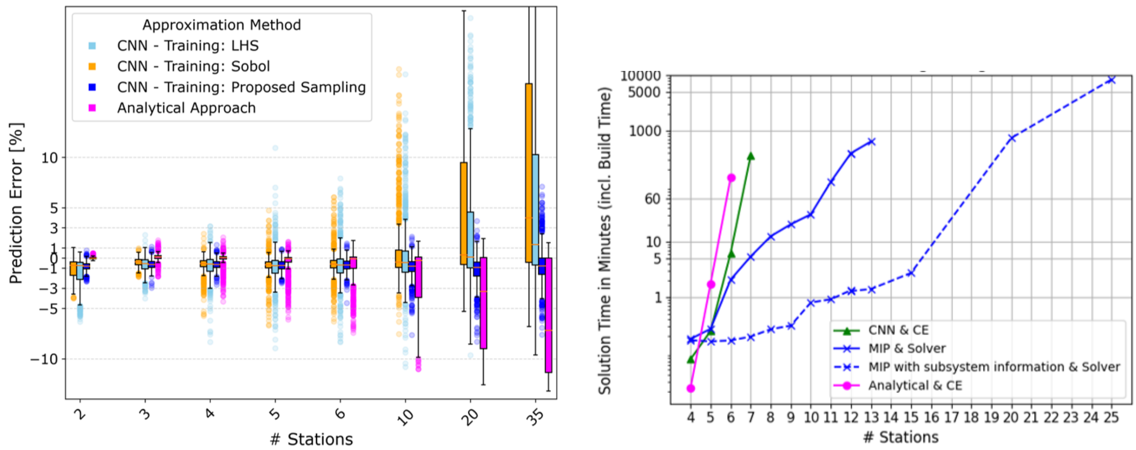
3.3. Mixed-Integer Linear Program

The trained CNN is linearized via binary variables and big-M constraints to replace Constraint (1b). The result is a MILP formulation of the BAP (1) that is implemented with GAMSPy and solved with CPLEX 22.1.2.

4. Numerical Results

4.1. Performance evaluation quality

Figure 1a depicts boxplots of the relative error for throughput predictions across the three CNNs and the analytical approximation approach by Buzacott and Shanthikumar (1993). The CNNs, trained on flow lines with four to ten stations, demonstrate good extrapolation to shorter lines, though evaluating longer lines ($J \geq 10$) proves to be more challenging. The CNN trained on our proposed sampling approach achieves the lowest mean absolute percentage error and variance for longer flow lines. Intentionally, the CNNs systematically underestimate the expected throughput. This loss-induced bias is advantageous as it reduces the likelihood of infeasible solutions due to approximation errors when solving the BAP (1).



(a) Prediction Error; Test: Proposed Sampling

(b) Solution Time Comparison

Figure 1. Performance Evaluation of Prediction Quality and Solution Speed

4.2. Solution Speed and Quality

The optimization performance is assessed by comparing our MILP approach against the analytical approximation paired with complete enumeration. Re-evaluating the obtained solutions via our simulation model confirms comparable solution quality between the two methods. However, the MILP formulation provides significantly faster computation times and greater flexibility. Compared to a Complete Enumeration (CE) combined with the analytical approach or combined with a forward pass using the trained CNN, the MILP solver achieves significantly faster solution times (Figure 1b). Information about lower or upper bounds such as proposed by Weiss and Stolletz (2015) can be integrated easily into the MILP by introducing new constraints. They lead to even faster solution times. To illustrate the flexibility of the approach we also apply it to a cost-weighted BAP and a joint optimization of processing rate and buffer allocation. In both cases our approach ranges from 0.63% to 0% optimality gaps for 4- and 5-station flow lines.

5. Conclusions

Our research demonstrates that CNNs trained on systematically choosen training data are a competitive alternative to traditional performance evaluation methods for stochastic flow lines such as simulation or analytical approximations. Furthermore, linearizing CNNs allows their integration within a MILP framework. The proposed approach combines the predictive power of neural networks with the optimization power and flexibility of mixed integer programming.

Acknowledgments
We like to acknowledge helpful discussion with GAMS Software GmbH concerning the efficient linearization of the CNN via GAMSPy.

References

Anderson, Ross, Joey Huchette, Will Ma, Christian Tjandraatmadja, Juan Pablo Vielma. 2020. Strong mixed-integer programming formulations for trained neural networks. *Mathematical Programming* **183**(1-2) 3–39.

Bromley, Jane, James W. Bentz, Léon Bottou, Isabelle Guyon, Yann Lecun, Cliff Moore, Eduard Säckinger, Roopak Shah. 1993. Signature verification using a "siamese" time delay neural network. *International Journal of Pattern Recognition and Artificial Intelligence* **07**(04) 669–688.

Buzacott, John A., J. George Shanthikumar. 1993. *Stochastic Models of Manufacturing Systems*. Prentice Hall, Englewood Cliffs, NJ.

Camponogara, Eduardo, Laio Oriel Seman, Eduardo Rauh Müller, Luis Kin Miyatake, Eduardo Ferreira Gaspari, Bruno Ferreira Vieira, Bruno Machado Pacheco. 2024. A ReLU-based linearization approach for maximizing oil production in subsea platforms: An application to flow splitting. *Chemical Engineering Science* **295** 120165.

Grimstad, Bjarne, Henrik Andersson. 2019. ReLU networks as surrogate models in mixed-integer linear programs. *Computers & Chemical Engineering* **131** 106580.

Maragno, Donato, Holly Wiberg, Dimitris Bertsimas, Ş. İlker Birbil, Dick den Hertog, Adejuyigbe O. Fajemisin. 2025. Mixed-integer optimization with constraint learning. *Operations Research* **73**(2) 1011–1028.

Muth, Eginhard J. 1979. The reversibility property of production lines. *Management Science* **25**(2) 152–158.

Tremblet, David, Simon Thevenin, Alexandre Dolgui. 2024. Neural networks to approximate capacity consumption in lot-sizing models. *2024 IEEE 20th International Conference on Automation Science and Engineering (CASE)*. IEEE, 1916–1921.

Weiss, Sophie, Justus Arne Schwarz, Raik Stolletz. 2019. The buffer allocation problem in production lines: Formulations, solution methods, and instances. *IISE Transactions* **51**(5) 456–485.

Weiss, Sophie, Raik Stolletz. 2015. Buffer allocation in stochastic flow lines via sample-based optimization with initial bounds. *OR Spectrum* **37**(4) 869–902.

Designing Efficient Supply Chain Networks under Supply Disruption and Demand Uncertainty

Amir Azaron
Sauder School of Business, University of British Columbia, Vancouver, Canada, email: amir.azaron@ubc.ca

Kai Furmans
Institute for Material Handling and Logistics, Karlsruhe Institute of Technology, Karlsruhe, Germany, email: kai.furmans@kit.edu

This study develops a two-stage stochastic optimization framework for supply chain network design under simultaneous supply disruption and demand uncertainty. Strategic decisions determine the location of manufacturing facilities, while second-stage decisions optimize procurement, production, inventory, and logistics. A distinguishing feature of the proposed framework is the explicit modelling of unreliable suppliers using continuous-time Markov chains. Suppliers transition between operational and failed states over time, where lifetime and repair durations follow exponential distributions. Demand uncertainty is incorporated through normally distributed market demand and chance-constrained programming to satisfy service-level requirements. The framework is further extended to incorporate multiple supplier operational states and limited repair resources. Numerical examples demonstrate the applicability of the proposed model and highlight the trade-offs between cost efficiency, resilience, and service reliability under uncertainty.

Keywords: Supply chain network design; stochastic programming; supply disruption; continuous-time Markov chains

1 Introduction

Modern supply chains are increasingly exposed to disruptions caused by natural disasters, geopolitical conflicts, transportation interruptions, and operational failures. Recent global events such as the COVID-19 pandemic have demonstrated the vulnerability of supply chain networks and emphasized the importance of resilient supply chain design.

This study develops a two-stage stochastic optimization model for supply chain design under supply and demand uncertainty. A key contribution of the proposed framework is the explicit modelling of unreliable suppliers through continuous-time Markov chains. Suppliers alternate between operational and failed states over time, allowing the model to capture realistic disruption and recovery dynamics.

At the first stage, strategic decisions determine the locations of manufacturing facilities. At the second stage, scenario-dependent operational decisions determine procurement quantities, production levels, inventory holdings, and transportation flows.

Multi-period supply chain network design problem under demand uncertainty was investigated by numerous researchers including Guillen et al. (2005), Correia et al. (2013), Badri et al. (2013), Melo et al. (2012), Wilhelm et al. (2013), and Azaron et al. (2021) without addressing supply uncertainty. The interactions between unreliable suppliers and manufacturing facilities in supply chains were studied by Ahiska et al. (2013), Benyoucef et al. (2013), Ebrahim Nejad et al. (2014), Xanthopoulos et al. (2012), Qi and Lee (2015), Saghafian and Van Oyen (2016), and Wu et al. (2012). However, none of those research works addressed the supply chain design problem. Azaron and Furmans (2025) addressed supply uncertainty with time-dependent lifetime distributions in designing efficient supply chain networks.

The proposed framework in Azaron and Furmans (2025) is further extended in two directions. First, suppliers may operate under multiple operational states such as full capacity, reduced capacity, or complete failure. Second, limited repair resources are incorporated to model realistic supplier recovery processes.

2 Mathematical Model

Consider a supply chain network consisting of suppliers, manufacturing facilities, and markets over a finite planning horizon. Strategic decisions determine whether candidate manufacturing facilities should be established, while operational decisions optimize raw material procurement, production, inventory management, and logistics. The objective is to minimize the expected total cost of the supply chain:

Min z = fixed facility costs + expected procurement costs + production costs + inventory holding costs + transportation costs.

The major decision variables are:

- y_i: binary variable indicating whether manufacturing site i is established,
- x_{ijt}^{ks}: quantity of raw material k to be purchased from supplier i and shipped to manufacturing site j during time interval t under scenario s
- x_{jt}^{ls}: quantity of product l to be produced at manufacturing site j at time t under scenario s,
- I_{jt}^{ks}: inventory level of raw material k at manufacturing site j at time t under scenario s,
- y_{ijt}^{ls}: quantity of finished product l shipped from manufacturing site i to market j at time t under scenario s.

A subset of suppliers is assumed to be unreliable. Each unreliable supplier alternates between operational state (state 0) and failed state (state 1). The time to failure and repair time follow exponential distributions with different parameters. Supplier dynamics are therefore represented using a continuous-time Markov chain.

The transition probabilities between supplier states are obtained from the Chapman-Kolmogorov equations and are used to generate disruption scenarios over the planning horizon. Let $P_{ij}(t)$ denote the probability that a supplier is at state j = 0 or 1 at time t given that it was at state i = 0 or 1 at time 0. The transition probabilities is obtained by solving the following system of linear differential equations derived from the *Chapman–Kolmogorov* forward equations:

$$\begin{bmatrix} P'_{00}(t) & P'_{01}(t) \\ P'_{10}(t) & P'_{11}(t) \end{bmatrix} = \begin{bmatrix} P_{00}(t) & P_{01}(t) \\ P_{10}(t) & P_{11}(t) \end{bmatrix} \times \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

$$\sum_{j=0}^{1} P_{ij}(t) = 1 \quad \forall i = 0,1 \quad P_{00}(0) = P_{11}(0) = 1 \quad P_{01}(0) = P_{10}(0) = 0 \qquad (1)$$

The proposed model includes the following major constraints:

Inventory balance constraints ensure material flow conservation at manufacturing facilities.

Supplier capacity constraints restrict raw material shipments according to supplier operational states under each scenario.

Production capacity constraints limit production quantities according to available manufacturing resources.

Chance constraints guarantee that customer demand is satisfied with a specified probability level.

Facility location constraints ensure that production can occur only at established manufacturing facilities.

The proposed two-stage stochastic model is formulated as follows:

$$\text{Min} \quad z = \sum_{i\epsilon B} c_i y_i + \sum_{s\epsilon S} p_s \Big[\sum_{t=2}^{T} \sum_{k\epsilon K} \sum_{i\epsilon A} \sum_{j\epsilon B} a_{ijt}^{k} x_{ijt}^{ks} + \sum_{t=2}^{T} \sum_{l\epsilon L} \sum_{j\epsilon B} b_{jt}^{l} x_{jt}^{ls} + \sum_{t=2}^{T} \sum_{k\epsilon K} \sum_{j\epsilon B} h_{jt}^{k} I_{jt}^{ks} + \sum_{t=2}^{T} \sum_{l\epsilon L} \sum_{i\epsilon B} \sum_{j\epsilon C} q_{ijt}^{l} y_{ijt}^{ls}\Big]$$

s.t.

$$I_{jt}^{ks} = I_{jt-1}^{ks} + \sum_{i\epsilon A} x_{ijt}^{ks} - \sum_{l\epsilon L} c^{kl} x_{jt}^{ls} \ \forall j\epsilon B \ k\epsilon K \ s\epsilon S \ 2 \ll t \ll T$$

$$\sum_{j\epsilon B} x_{ijt}^{ks} \ll s_{it}^{ks} \quad \forall i\epsilon A \ k\epsilon K \ s\epsilon S \ 2 \ll t \ll T$$

$$x_{it}^{ls} = \sum_{j\epsilon C} y_{ijt}^{ls} \quad \forall i\epsilon B \ l\epsilon L \ s\epsilon S \ 2 \ll t \ll T$$

$$\sum_{t=2}^{T}\sum_{l\epsilon L}\sum_{s\epsilon S} x_{it}^{ls} \ll \beta y_i \ \ \forall i\epsilon B$$

$$\sum_{l\epsilon L} r_{jf}^{l} x_{jt}^{ls} \ll \hat{r}_{jf}^{t} \ \ \forall j\epsilon B \ s\epsilon S \ f = 1,2,\dots,F_j \ 2 \ll t \ll T$$

$$\sum_{i\epsilon B} y_{ijt}^{ls} \gg \mu_{jt}^{l} + z_{SL_{jt}^{l}} \sigma_{jt}^{l} \ \ \forall j\epsilon C \ l\epsilon L \ s\epsilon S \ 2 \ll t \ll T$$

$$y_i = \begin{cases} 1 & if\ manufacturing\ site\ is\ built\ in\ location\ i \\ o & otherwise \end{cases}$$

$$x_{ijt}^{ks} \gg 0 \ \ \forall i\epsilon A \ j\epsilon B \ k\epsilon K \ s\epsilon S \ 2 \ll t \ll T$$

$$x_{jt}^{ls} \gg 0 \ \ \forall j\epsilon B \ l\epsilon L \ s\epsilon S \ 2 \ll t \ll T$$

$$I_{jt}^{ks} \gg 0 \ \ \forall j\epsilon B \ k\epsilon K \ s\epsilon S \ 2 \ll t \ll T$$

$$y_{ijt}^{ls} \gg 0 \ \ \forall i\epsilon B \ j\epsilon C \ l\epsilon L \ s\epsilon S \ 2 \ll t \ll T \qquad (2)$$

3 Extensions and Numerical Insights

The proposed framework is extended by allowing suppliers to operate under multiple operational states such as full capacity, reduced capacity, and complete failure. Under this extension, supplier transitions are represented through higher-dimensional continuous-time Markov chains.

The model is also extended by incorporating limited repair resources, where only a finite number of repair facilities are available. These extensions substantially increase the number of disruption scenarios and the computational complexity of the deterministic equivalent optimization model.

Several numerical examples were solved to demonstrate the applicability of the proposed framework. Results show that low-cost unreliable suppliers should be fully utilized whenever operational, while reliable suppliers act as backup sources during disruption periods. Inventory accumulation during low-cost periods was also shown to improve supply chain resilience.

The computational experiments further demonstrate that facilities with lower operational costs may be preferred despite higher construction costs.

4 Conclusions

This study presented a two-stage stochastic optimization framework for supply chain network design under supply disruption and demand uncertainty. A major contribution of the research lies in the explicit modelling of supplier reliability through continuous-time Markov chains.

The framework was further extended to incorporate multiple supplier operational states and limited repair facilities. Numerical examples demonstrated that the proposed approach can generate resilient and cost-efficient supply chain configurations while maintaining required service levels under uncertainty.

Future research may investigate multi-stage stochastic models, AI-assisted optimization approaches, and multi-objective formulations incorporating responsiveness and sustainability.

Acknowledgements

The authors acknowledge the financial support provided by the Institute for Material Handling and Logistics at Karlsruhe Institute of Technology (KIT) and International Excellence Fellowships of KIT.

References

1. Ahiska, S. S., Appaji, S. R., King, R. E., and D.P. Wang. 2013 "A Markov Decision Process-Based Policy Characterization Approach for a Stochastic Inventory Control Problem with Unreliable Sourcing." *International Journal of Production Economics* 144: 485-496.
2. Azaron, A., U. Venkatadri, and A. Farhang Doost. 2021. "Designing Profitable and Responsive Supply Chains under Uncertainty." *International Journal of Production Research* 59: 213-225.
3. Azaron A., and K. Furmans. 2025. "Designing Efficient Supply Chains with Unreliable Suppliers." *IFAC PapersOnLine* 59(10): 1059-1064.
4. Badri, H., M. Bashiri, and T. H. Hejazi. 2013. "Integrated Strategic and Tactical Planning in a Supply Chain Network Design with a Heuristic Solution Method." *Computers & Operations Research* 40: 1143-1154.
5. Benyoucef, L., X. Xie, and G. A. Tanonkou. 2013. "Supply Chain Network Design with Unreliable Suppliers: A Lagrangian Relaxation-Based Approach." *International Journal of Production Research* 51: 6435-6454.
6. Correia, I., M. T. Melo, and F. Saldanha-da-Gama. 2013. "Comparing Classical Performance Measure for a Multi-Period, Two-Echelon Supply Chain Network Design Problem with Sizing Decisions." *Computers & Industrial Engineering* 64: 366-380.
7. Ebrahim Nejad, A., I. Niroomand, and O. Kuzgunkaya. 2014. "Responsive Contingency Planning in Supply Risk Management by Considering Congestion Effects." *Omega* 48: 19-35.
8. Guillen, G., F. D. Mele, M. J. Bagajewicz, A. Espuna, and L. Puigjaner. 2004. "Multiobjective Supply Chain Design under Uncertainty." *Chemical Engineering Science* 60: 1535-1553.
9. Melo, M. T., S. Nickel, and F. Saldanha-da-Gama. 2012. "A Tabu Search Heuristic for Redesigning a Multi-Echelon Supply Chain Network over a Planning Horizon." *International Journal of Production Economics* 136: 218-230.
10. Qi, L., and K. Lee. 2015. "Supply Chain Risk Mitigations with Expedited Shipping." *Omega* 57: 98-113.
11. Saghafian, S., and M. P. Van Oyen. 2016. "Compensating for Dynamic Supply Disruptions: Backup Flexibility Design." *Operations Research* 64: 390-405.
12. Xanthopoulos, A., Vlachos, D., and E. Iakovou. 2012. "Optimal Newsvendor Policies for Dual-Sourcing Supply Chains: A Disruption Risk Management Framework." *Computers & Operations Research* 39: 350-357.
13. Wilhelm, W., X. Han, and C. Lee. 2013. "Computational Comparison of Two Formulations for Dynamic Supply Chain Reconfiguration with Capacity Expansion and Contraction." *Computers & Operations Research* 40: 2340-2356.
14. Wu, Y., Dong, M., Fan, T., and S. Liu. 2012. "Performance Evaluation of Supply Chain Networks with Assembly Structure under System Disruptions." *Computers & Operations Research* 39: 3229-3243.

Executable Multi-Agent Path Finding Considering Practical Hardware Constraints

Changhyun Chung
Department of Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology, South Korea

Young Jae Jang
Department of Industrial and Systems Engineering, Korea Advanced Institute of Science and Technology, South Korea

Traditional Multi-Agent Path Finding (MAPF) algorithms model robots as dimensionless points in perfect synchronization, ignoring the physical reality of industrial environments. This causes theoretically optimal paths to result in collisions or deadlocks when deployed on Autonomous Mobile Robots (AMRs) with kinematic constraints and execution delays. This paper proposes the Executable MAPF (E-MAPF) framework, which accounts for the robot's physical swept volume in planning and introduces the Temporal Action Plan Graph (TAPG) for event-driven execution robust to hardware timing uncertainty.

Key words: Multi-Agent Path Finding; Autonomous Mobile Robots; Swept Volume; Conflict-Based Search; Temporal Action Plan Graph

1. Introduction

Industrial Autonomous Mobile Robots (AMRs) operate along predefined trajectories in high-density manufacturing environments (Wurman et al. 2008). Classical MAPF abstracts agents as dimensionless points on discrete graphs, failing to capture swept-volume conflicts during rotation that commonly arise in real deployments (Zhang et al. 2023, Li et al. 2019). Furthermore, conventional MAPF assumes perfect time synchronization, which breaks down under finite acceleration, communication latency, and other hardware-level disturbances (Hönig et al. 2016, Atzmon et al. 2020). As shown in Figure 1, paths deemed collision-free by classical MAPF can cause actual collisions depending on robot orientation.

This paper presents **E-MAPF**, integrating a kinematically accurate collision model into the planning phase via E-CCBS, and introducing TAPG for robust event-driven execution. Validation via AutoMod simulation confirms collision-free navigation under stochastic conditions where traditional methods fail.

2. Problem Formulation

We model the environment as a directed **Driving State Graph** $G_{DS} = (V, E)$, where nodes encode the full kinematic configuration of the agent. Two node types are defined: a *Stop State* (S) representing a static pose (x, y, θ), and a *Move State* (M) representing the continuous transition trajectory between stops (Wen et al. 2022). This enables explicit modeling of the **swept volume**—the robot's polygonal footprint during motion—as the key primitive for collision detection.

The E-MAPF problem minimizes total flowtime over collision-free trajectories, where a collision is any intersection of swept volumes in the continuous space-time domain. This captures physical

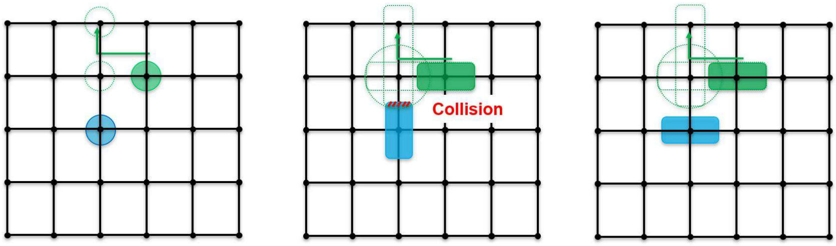


Figure 1. Classical MAPF (left) neglects orientation, causing collisions (middle) or safe passes (right) depending on the robot's swept volume.

conflicts that point-based models miss, such as overlapping footprints during wide turns even when agents occupy different logical nodes.

3. Proposed Methodology

3.1. Phase 1: Planning with E-CCBS

We extend the two-level Conflict-Based Search (CBS) (Sharon et al. 2015) into **Executable Continuous CBS (E-CCBS)**. When a swept-volume conflict is detected, E-CCBS imposes an **Unsafe Interval**—a continuous time range during which an agent is prohibited from entering a contested region (Andreychuk et al. 2021). This conservative constraint accounts for positional uncertainty during movement, guaranteeing plan safety by design. Safe Interval Path Planning (SIPP) (Phillips and Likhachev 2011) serves as the low-level solver.

3.2. Phase 2: Execution via TAPG

Time-stamped plans are fragile under real-world disturbances: a single delayed robot can cause system-wide deadlocks. The **Temporal Action Plan Graph (TAPG)** (Ma et al. 2017) replaces absolute clock times with **precedence constraints**: *Sequence Constraints* ensure each action completes before the next begins, and *Resource Constraints* ensure agents enter shared swept-volume regions only after the preceding agent has physically cleared them. This event-driven policy absorbs stochastic delays without global re-planning.

4. Numerical Experiments

We validated E-MAPF in an AutoMod simulation of a semiconductor fab logistics system with finite acceleration and control latency (Singh et al. 2022). In grid environments from $7{\times}7$ to $15{\times}15$, classical MAPF frequently produced invalid paths due to orientation-dependent collisions, whereas E-CCBS achieved near-100% feasibility. Under varying acceleration profiles, the time-synchronized baseline suffered frequent phantom deadlocks, while TAPG eliminated them entirely by triggering on state-completion events and maintained near-optimal throughput, confirming practical viability for real AMHS deployment.

5. Conclusion

E-MAPF closes the reality gap in industrial multi-robot coordination through kinematically aware planning (E-CCBS) and event-driven execution (TAPG), demonstrating collision-free, deadlock-free operation under stochastic conditions where traditional methods fail. Future work will extend the framework to heterogeneous fleets.

Acknowledgments
This research was supported by the Department of Industrial and Systems Engineering at KAIST.

References

Andreychuk, Anton, Konstantin Yakovlev, Eli Boyarski, Roni Stern. 2021. Improving continuous-time conflict based search. *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35. 11220–11227.

Atzmon, Dor, Roni Stern, Ariel Felner, Glenn Wagner, Roman Barták, Neng-Fa Zhou. 2020. Robust multi-agent path finding and executing. *Journal of Artificial Intelligence Research* **67** 549–579.

Hönig, Wolfgang, TK Kumar, Liron Cohen, Hang Ma, Hong Xu, Nora Ayanian, Sven Koenig. 2016. Multi-agent path finding with kinematic constraints. *Proceedings of the International Conference on Automated Planning and Scheduling*, vol. 26. 477–485.

Li, Jiaoyang, Pavel Surynek, Ariel Felner, Hang Ma, TK Satish Kumar, Sven Koenig. 2019. Multi-agent path finding for large agents. *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33. 7627–7634.

Ma, Hang, Jiaoyang Li, TK Kumar, Sven Koenig. 2017. Lifelong multi-agent path finding for online pickup and delivery tasks. *arXiv preprint arXiv:1705.10868* .

Phillips, Mike, Maxim Likhachev. 2011. Sipp: Safe interval path planning for dynamic environments. *2011 IEEE international conference on robotics and automation*. IEEE, 5628–5635.

Sharon, Guni, Roni Stern, Ariel Felner, Nathan R. Sturtevant. 2015. Conflict-based search for optimal multi-agent pathfinding. *Artificial Intelligence* **219** 40–66. doi:https://doi.org/10.1016/j.artint.2014.11.006. URL https://www.sciencedirect.com/science/article/pii/S0004370214001386.

Singh, Nitish, Quang-Vinh Dang, Alp Akcay, Ivo Adan, Tugce Martagan. 2022. A matheuristic for agv scheduling with battery constraints. *European Journal of Operational Research* **298**(3) 855–873.

Wen, Licheng, Yong Liu, Hongliang Li. 2022. Cl-mapf: Multi-agent path finding for car-like robots with kinematic and spatiotemporal constraints. *Robotics and Autonomous Systems* **150** 103997.

Wurman, Peter R, Raffaello D'Andrea, Mick Mountz. 2008. Coordinating hundreds of cooperative, autonomous vehicles in warehouses. *AI magazine* **29**(1) 9–9.

Zhang, Yue, Daniel Harabor, Pierre Le Bodic, Peter J Stuckey. 2023. Efficient multi agent path finding with turn actions. *Proceedings of the International Symposium on Combinatorial Search*, vol. 16. 119–127.

Stochastic Optimization of Patient-to-Room Assignments in Hospital Wards

Justus Arne Schwarz
Chair of Production Management, Faculty of Business, Economics and Real Estate, University of Regensburg], Regensburg, Germany, justus.schwarz@ur.de

Philipp Pelz
Chair of Production Management,Faculty of Business, Economics and Real Estate, University of Regensburg], Regensburg, Germany, philipp.pelz@ur.de

Hospitals face increasing pressure to manage limited bed capacity under growing demand. Assigning patients to rooms is a complex decision problem, driven by stochastic emergency arrivals and lengths of stay (LOS), as well as operational constraints such as room capacity and gender-specific requirements. At the same time, hospitals aim to avoid undesirable actions for patients, including transfers between regular rooms, assignments to the overflow area providing temporary capacity outside regular rooms, and rejections. The literature largely focuses on deterministic models, either assuming full prior knowledge or using rolling-horizon methods that re-optimize whenever a stochastic event occurs. In contrast, we propose a discrete-time Markov Decision Process (MDP) framework that explicitly represents uncertainty over a finite planning horizon. The model supports key operational decisions such as regular room and overflow area allocations, rejections, and transfers to regular rooms. We solve the MDP using backward induction and present first results from a simulation study, comparing its performance with benchmark policies based on the literature and simple heuristic decision rules.

Key words: Patient-to-Room Assignment; Health Care; Markov Decision Process

1. Introduction

This paper studies the operational planning problem of assigning patients to rooms within a hospital ward, commonly referred to as the patient-to-room assignment (PRA) problem. Hospital wards are designated areas within a hospital for patient care based on medical needs. They consist of rooms with varying bed capacities, ranging from single-bed rooms to larger shared rooms, to accommodate admitted patients. Hallways or other areas may be used as overflow areas to manage high patient volumes when demand exceeds regular capacity. Patient to room assignments in hospital wards is a challenging task due to stochastic influences such as emergency patient arrivals and uncertain lengths of stay (LOS). While elective patient admissions are planned in advance, emergency patients arrive randomly but still need bed capacity after their admission. The LOS can also change unexpectedly, for instance, if patients are diagnosed with an infection during their stay, requiring extended treatment. Key decisions include allocating an arrived patient to a regular room, the overflow area, or rejecting them, while ensuring that constraints such as gender separation and limited bed capacity in rooms are respected. Additionally, patients already allocated to the ward may be transferred to a different regular room. The goal is to ensure optimal patient care by minimizing undesired actions, such as transfers between regular rooms, allocations to the overflow area, or rejections.

Most existing approaches to the PRA problem rely on deterministic formulations, many of which can be traced back to the static model introduced by Demeester et al. (2010). While several extensions have been proposed, uncertainty is typically handled implicitly or through rolling-horizon approaches. Stochastic formulations that explicitly capture uncertainty remain scarce, with only few exceptions in the literature (Jiang et al. 2023). Moreover, a systematic comparison of stochastic and deterministic approaches in the context of patient-to-room assignment is missing.

The contribution of this work is twofold:

- **New stochastic model formulation:** We formulate a stochastic model as an Markov Decision Process (MDP), leveraging a carefully designed state space that avoids symmetries inherent in patient-to-room and patient-to-bed allocations. In contrast, most existing approaches relying on deterministic formulations that do not explicitly account for stochastic events. While this assumption simplifies the optimization, it fails to reflect the uncertainty inherent in real hospital operations. A second stream of research considers dynamic approaches, first introduced by Ceschia and Schaerf (2012), where a deterministic model is solved repeatedly in a rolling horizon fashion whenever new information becomes available or a stochastic event occurs (e.g., Ceschia and Schaerf 2016, Lusby et al. 2016, Vancroonenburg et al. 2016, Schäfer et al. 2019). Although this dynamic setting enables some reactivity, decisions rely solely on the current state without accounting for future uncertainty. Stochastic modeling for PRA remains largely unexplored. One exception is the study by Jiang et al. (2023), which formulates the problem as a MDP, albeit under specific assumptions, such as a memoryless LOS distribution, considering only double rooms, excluding elective patients, and without explicit knowledge of which patient is allocated to which room. However, no existing approach offers a stochastic model that jointly accounts for both stochastic influences while maintaining comparability with established formulations in the literature.
- **Performance comparison with existing optimization approaches:** We compare the performance of the proposed stochastic optimization approach with existing dynamic but deterministic models and a simplified stochastic model from the literature. The added value of explicit stochastic modeling in the patient-to-room assignment context is analyzed.

Section 2 provides details regarding the modeling and solution of the proposed MDP and Section 3 describes numerical results and Section 4 provides concluding remarks.

2. Modeling and Solution Approach

The PRA problem is modeled as a discrete-time MDP, capturing the dynamics of a hospital ward over a finite planning horizon. The finite time horizon is divided into macro periods (days) and micro periods (within-day), allowing the model to account for decisions and stochastic events such as patient arrivals and discharges (modeled via stochastic LOS) that may occur. The course of a day and its micro periods are illustrated in Figure 1. The planning horizon is finite, as at any decision point only information on currently scheduled elective patients is available. It is implemented in a rolling manner, such that after each day the entire planning horizon is shifted forward by one day and the optimization problem is re-solved.

For initial results on small instances, backward induction is used to compute optimal decisions. The planning horizon is set to 14 days, for which elective patient arrivals are known. To obtain realistic terminal state values, the optimization is extended to 28 days. For days 14 to 28, the probability distributions of emergency arrivals are augmented by the distributions of elective patients, reflecting the lack of deterministic information beyond the initial horizon.

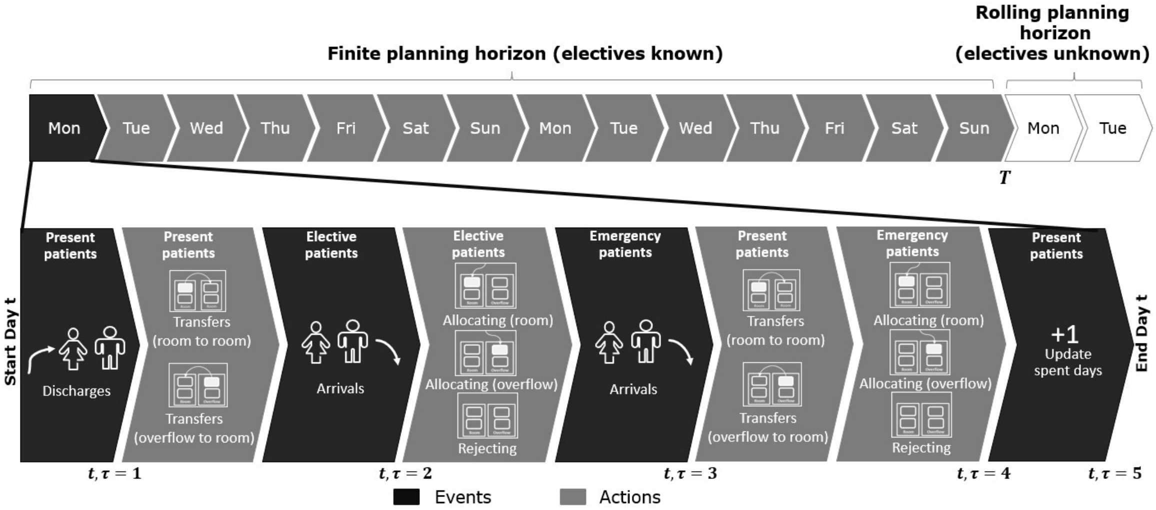


Figure 1. Order of events and actions in the micro periods

3. Numerical Results

The numerical study evaluates the decisions derived from the Markov Decision Process (MDP) with a discrete-event simulation over a planning horizon of multiple months. The decisions are benchmarked against decisions obtained from a deterministic MIP that is resolved whenever a stochastic event occurs, as well as from a simplified MDP based on Jiang et al. (2023), for the same realizations. In addition, two heuristic baseline policies are considered: a no-transfer policy, where patients remain in their initially assigned rooms throughout their stay, and an always-transfer-to-allocate policy, where transfers are executed whenever feasible to enable the allocation of additional patients. The study further investigates the sensitivity of results with respect to key parameters, such as the penalty associated with patient transfers. The results indicate that the MDP-based approach outperforms both the literature-based benchmark policies and the heuristic policies.

4. Conclusions

This work introduces a novel stochastic formulation of the patient-to-room assignment problem as a Markov Decision Process, explicitly capturing uncertainty in hospital operations. By designing a state space that avoids allocation symmetries, the model enables computationally efficient representation as well as a more realistic approach compared to existing deterministic approaches. The results demonstrate that explicitly incorporating stochastic elements can improve decision quality over dynamic but deterministic rolling-horizon methods and a simplified stochastic approach. Future work is needed to solve larger instances that remain numerically intractable with the proposed approach.

References

Ceschia, Sara, Andrea Schaerf. 2012. Modeling and solving the dynamic patient admission scheduling problem under uncertainty. *Artificial Intelligence in Medicine* **56**(3) 199–205.

Ceschia, Sara, Andrea Schaerf. 2016. Dynamic patient admission scheduling with operating room constraints, flexible horizons, and patient delays. *Journal of Scheduling* **19**(4) 377–389.

Demeester, Peter, Wouter Souffriau, Patrick De Causmaecker, Greet Vanden Berghe. 2010. A hybrid tabu search algorithm for automatically assigning patients to beds. *Artificial Intelligence in Medicine* **48**(1) 61–70.

Jiang, Yanping, Feifei Yang, Zhenpeng Tang, Quan-Lin Li. 2023. Admission control of hospitalization with patient gender by using markov decision process. *International Transactions in Operational Research* **30**(1) 70–98.

Lusby, Richard Martin, Martin Schwierz, Troels Martin Range, Jesper Larsen. 2016. An adaptive large neighborhood search procedure applied to the dynamic patient admission scheduling problem. *Artificial Intelligence in Medicine* **74** 21–31.

Schäfer, Fabian, Manuel Walther, Alexander Hübner, Heinrich Kuhn. 2019. Operational patient-bed assignment problem in large hospital settings including overflow and uncertainty management. *Flexible Services and Manufacturing Journal* **31**(4) 1012–1041.

Vancroonenburg, Wim, Patrick De Causmaecker, Greet Vanden Berghe. 2016. A study of decision support models for online patient-to-room assignment planning. *Annals of Operations Research* **239**(1) 253–271.

Elements of a branch-price-and-cut algorithm to solve a task scheduling and assignment problem

Stefan Helber

Institut für Produktionswirtschaft, Leibniz Universität Hannover, Hannover, Germany stefan.helber@prod.uni-hannover.de

In this presentation, we consider the healthcare-inspired problem to assign a set of tasks for multiple periods to multiple persons with different and overlapping skill sets under deterministic conditions. For each task, there is a certain degree of temporal flexibility modeled by a time window of the start times. In addition, we assume that there is some degree of flexibility with respect to the working hours, overtime etc. The problem contains elements of a vehicle routing problem and turns out be hard to solve using standard MIP solvers. We present elements of a branch-price-and-cut algorithm implemented within the SCIP framework to solve that problem. In particular, we comment on the solution of the pricing problem as resource-constrained shortest path problem and the usage of the SCIP framework to implement the branch-and-price approach as well as the C++ Boost library to solve the pricing problem via a labeling algorithm.

Key words: Task assignment and scheduling, Branch-and-price, SCIP

1. Integrated multi-period task assignment and scheduling

We consider a problem setting that can be found in healthcare service systems, as well as in other service systems. The problem is to find schedules in which heterogeneous tasks requiring a specific qualification have to be assigned to heterogeneously qualified persons. Qualification patterns for the personnel are assumed to be overlapping, as shown in Figure 1. Consequently, the assignment problem cannot be cleanly separated by personnel groups and must be solved as an integrated whole.

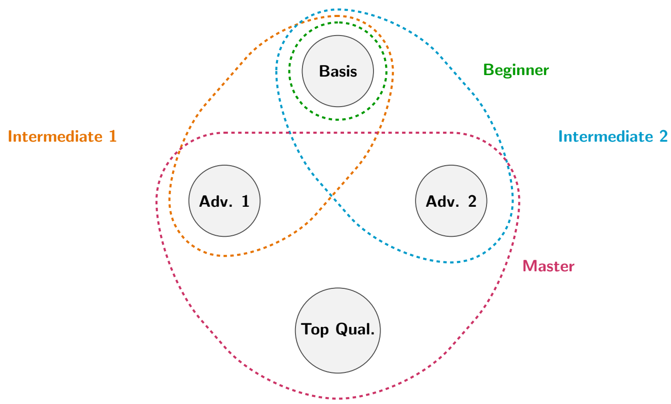


Figure 1. Personnel categories with overlapping qualifications.

Every task has a specific duration and a defined *task time window* specifying the earliest and latest possible start times. A core requirement is that each task is scheduled exactly once, meaning it is assigned to exactly one qualified person within a single work time window.

To manage the temporal flexibility of employee shifts, the model introduces *work time windows*. These represent periods during which a person can actively work, and they are assigned to the week in which they begin. This allows for shifts that cross calendar days, such as night shifts.

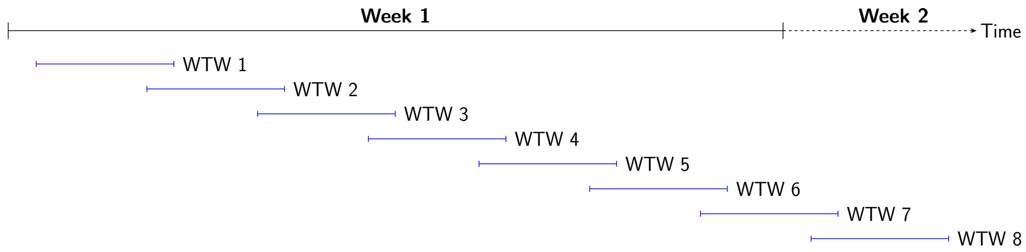


Figure 2. Work time windows and their assignment to weeks.

Within each work time window, time is modeled using discrete moments (e.g., full hours or half hours). An employee's work phase is defined by selecting exactly one begin moment and one end moment. The actual assignment of tasks to an employee can be viewed as finding a feasible route or path through a graph of tasks. The employee moves forward in time, completing tasks in sequence without any temporal overlap, as shown in Figure 3.

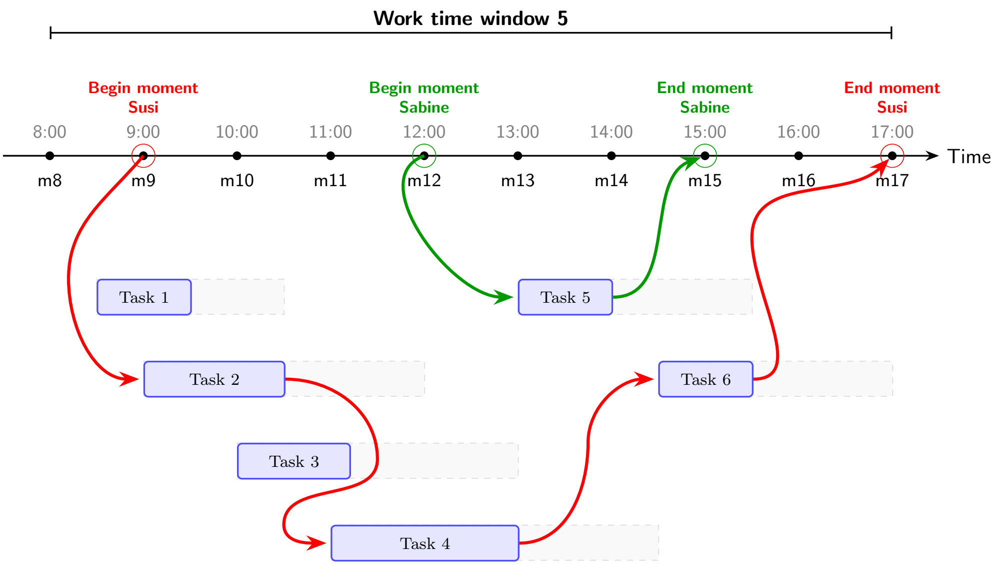


Figure 3. Finding a feasible schedule as a path through assignable tasks over discrete moments.

The goal is to determine which person performs which tasks and at what time, while minimizing the overall personnel costs. These costs consist of the baseline costs for activating a person during a work time window, as well as penalties for both daily and weekly overtime.

2. Dantzig-Wolfe decomposition

While solving the integrated problem directly can be computationally prohibitive, the mathematical formulation is tailored for a Dantzig-Wolfe decomposition. The problem's structure allows it to be broken down into a master problem and multiple smaller subproblems, e.g., one for each combination of person and planning time window. The master problem is an LP generating dual values for the constraints, while the subproblems are IPs generating new schedules for each combination of person and time window. The latter problem can be solved as a resource-constrained shortest-path problem, see Figure 4. Its solutions, if they have negative reduced costs, become new variables of the master problem in a column-generation approach serving to determine a lower bound on objective function values in the nodes of a branch-and-bound search tree.

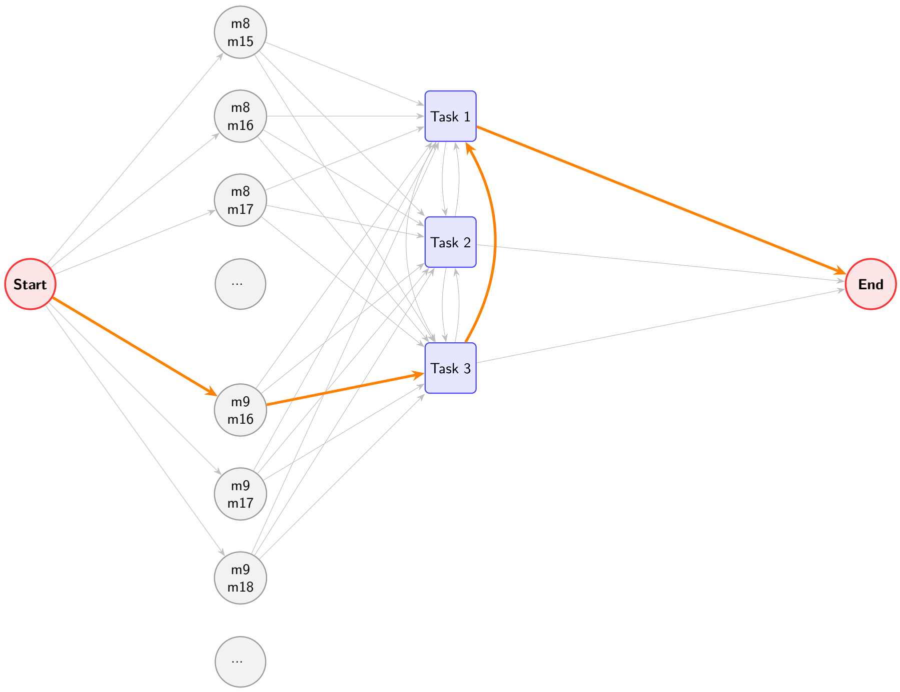


Figure 4. Modeling the subproblem as a resource-constrained shortest path problem.

3. Implementation aspects

We use the SCIP framework to implement the Branch-price-and-cut algorithm and the C++ Boost Library framework to model resource-constrained shortest path problems. We report some first numerical results related to the performance of the method by comparing it to results obtained by using the Gurobi MIP solver for the monolithic model.