

OTIMIZAÇÃO DE ESCALONAMENTO EM *FLOW SHOP* PERMUTACIONAL DINÂMICO VIA SIMULAÇÃO HÍBRIDA E APRENDIZADO POR REFORÇO

Felipe Augusto Feichas – feichas2000@unifei.edu.br
Alexandre Ferreira de Pinho – pinho@unifei.edu.br

Palavras-chave: sequenciamento da produção, *flow shop*, simulação baseada em agentes, simulação de eventos discretos, aprendizado por reforço.

1. INTRODUÇÃO

O sequenciamento da produção (SP) é amplamente estudada na literatura e sua solução impacta diretamente o desempenho e a competitividade industrial, refletindo-se no sucesso das empresas (PINEDO, 2022). Entre as classes de sequenciamento, destaca-se o *flow shop* permutacional (PFSP) é um problema típico de otimização do escalonamento, que dada sua complexidade computacional, a resolução exige abordagens cada vez mais sofisticadas.

Nesse contexto, têm surgido soluções híbridas que combinam heurísticas com meta-heurísticas e IA, com simulação para obter as relações complexas entre operações (DEL GALLO *et al.*, 2023). No entanto, grande parte dos estudos permanece restrita a experimentos computacionais estáticos ou determinísticos, com pouca consideração às políticas adaptativas (NEUFELD *et al.*, 2023). Por exemplo, revisões sistemáticas de *flow shops* multiobjetivos revelam lacunas entre simulação e métodos de otimização para cenários dinâmicos, onde incertezas como falhas de máquina e chegada de novos pedidos, o que limita a aplicabilidade em ambientes reais da indústria 4.0 (NEUFELD *et al.*, 2023). Além disso, os paradigmas emergentes da indústria 5.0, onde a integração de tecnologias digitais avançadas e abordagens centradas no ser humano e na sustentabilidade está transformando os sistemas de manufatura (ELBASHEER *et al.*, 2024).

A integração entre simulação baseada em agentes (SBA), simulação de eventos discretos (SED), métodos de otimização e aprendizado por reforço (RL) permite criar ambientes simulados de alta complexidade dos sistemas de manufatura modernos e atender a necessidade de modelar eventos estocásticos em ambientes dinâmicos e incertos (ZHANG *et al.*, 2025). Abordagens de RL com *Q-learning* combinadas com heurísticas construtivas permite aprender políticas que se adaptam a eventos estocásticos, que mostram ganhos sobre heurísticas clássicas em *makespan* (HE *et al.*, 2022).

Dessa forma, o presente trabalho se propõe a integrar simulação híbrida (SED+SBA) e RL para aprender políticas de despacho adaptativas que melhorem múltiplos objetivos: makespan, tempo de fluxo, atraso ponderado, estabilidade do sequenciamento e utilização média e o balanceamento dos recursos.

2. DESENVOLVIMENTO

2.1 Referencial teórico

O problema de SP é NP-difícil devido a restrições operacionais, variabilidade estocástica e múltiplos objetivos (HE *et al.*, 2022). No PFSP, máquinas em sequência fixa processam cada tarefa em todas elas, na mesma ordem, do primeiro ao último estágio (PINEDO, 2022).

As soluções para o problema de *flow shop* têm evoluído ao longo dos anos, iniciado pelo estudo de Johnson (1954), acompanhadas por heurística como NEH (1983). Na década de 1990, destacaram-se as meta-heurísticas inspiradas na natureza, como algoritmos genéticos e inteligência de enxame. Em meados dos anos 2000, IA começaram a ser aplicadas. Além disso, a tendência atual concentra-se em soluções híbridas, que integram heurísticas, meta-heurísticas e IA como RL, integrando com simulações (ZHANG *et al.*, 2025).

O RL modela o escalonamento como um Processo de Decisão Markoviano (MDP), onde um agente aprende políticas ótimas via recompensas cumulativas (SUTTON; BARTO, 2018). Em PFSP, estados incluem filas de *jobs* e *status* de máquinas; ações são regras de despacho; recompensas penalizam *makespan*, atrasos e desequilíbrios.

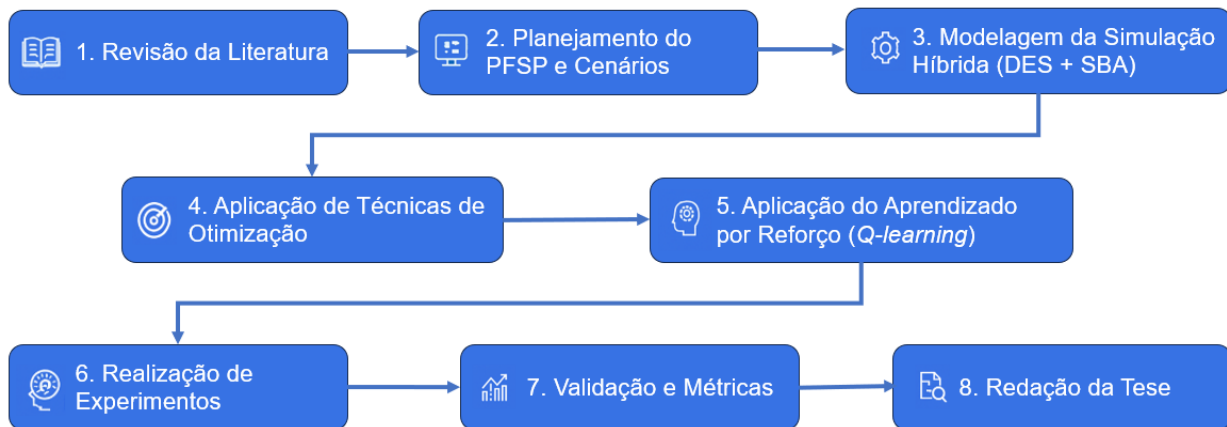
Com a evolução do RL, a integração com simulação tornou-se uma estratégia poderosa para treinar, validar e guiar algoritmos de escalonamento em cenários estocásticos. Esta abordagem de otimização de simulação (SO), combina as capacidades preditivas dos modelos de simulação com algoritmos de otimização para encontrar as estratégias operacionais mais eficazes (GHASEMI *et al.*, 2024). Dois paradigmas se destacam nesse domínio: SED e SBA. Segundo LAW (2015), a SED modela a evolução temporal de um sistema no qual as variáveis de estado mudam apenas quando ocorrem eventos como chegadas, partidas e falhas, e o tempo é controlado por um relógio de simulação que avança para o próximo evento, atualizando o estado. Já a SBA representa entidades (operadores, máquinas, tarefas) como agentes autônomos com estados e regras locais, permitindo capturar comportamentos e interações não triviais do sistema (GRIMM; RAILSBACK, 2019).

A simulação híbrida é a construção de um único modelo em que partes diferentes do sistema são representadas por métodos distintos, cada qual cobrindo um nível de abstração, com o potencial de aumentar o realismo e a capacidade de análise.

2.2 Metodologia

O processo de pesquisa passa por diversas etapas, conforme Figura 1.

Figura 1 - Fluxograma da pesquisa.

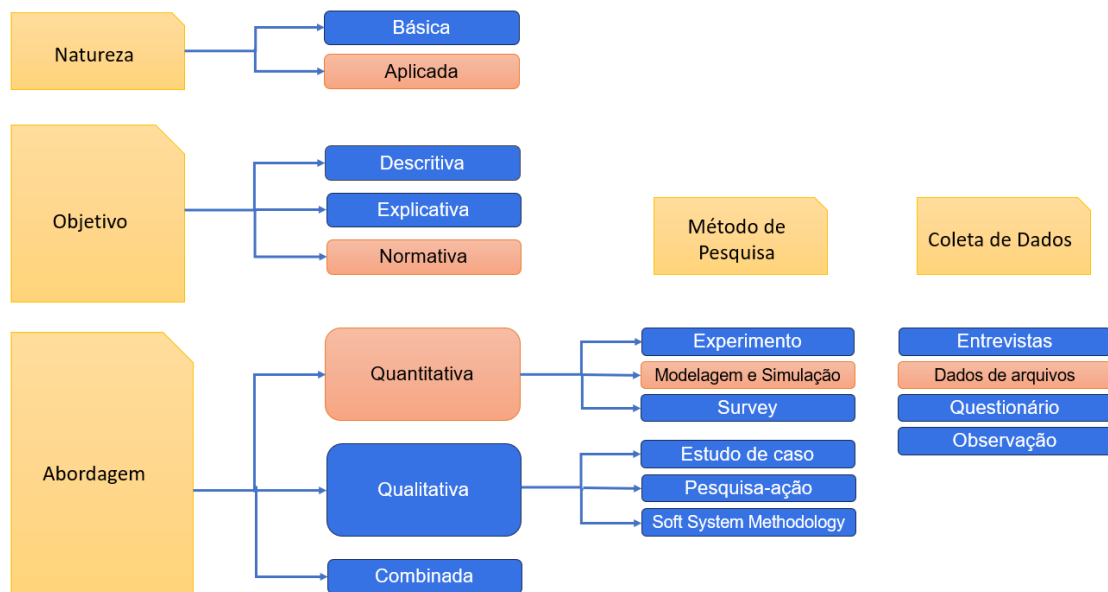


Fonte: Os autores.

O processo de pesquisa inicia-se com a revisão da literatura (1), que mapeia lacunas, técnicas e métricas relevantes. Em (2), realiza-se o planejamento do PFSP e dos cenários. Em (3), elabora-se a modelagem da simulação híbrida: o fluxo produtivo em SED e os agentes no SBA. Em (4), aplicam-se heurísticas e meta-heurísticas para comparação (benchmark). Em (5), implementa-se o RL. Em (6), procede-se à execução dos experimentos. Em (7), realizam-se a validação e a mensuração das métricas. Por fim, em (8), ocorre a redação da tese.

A classificação da pesquisa é dada pela Figura 2 e Tabela 1, elaborada seguindo os passos de Turrioni e Mello (2012).

Figura 2 - Classificação da pesquisa.



Fonte: Adaptado de Turrioni e Mello (2012).

Tabela 1. Classificação da pesquisa.

Tipo	Classificação	Justificativa
Natureza	Aplicada	Foco em problemas reais de produção, com uso direto para melhorar o desempenho.
Objetivo	Normativa	Desenvolver e testar novas políticas de sequenciamento, comparando com as usuais.
Abordagem	Quantitativa	Mensuração de variáveis e análise objetiva via métricas e testes estatísticos.
Métodos de pesquisa	Modelagem e simulação	Modelos computacionais para avaliar mudanças quando testar no real é caro/complexo.
Coleta de dados	Dados de arquivo	Uso de logs/benchmarks e dados sintéticos gerados pela simulação.

Fonte: Os autores.

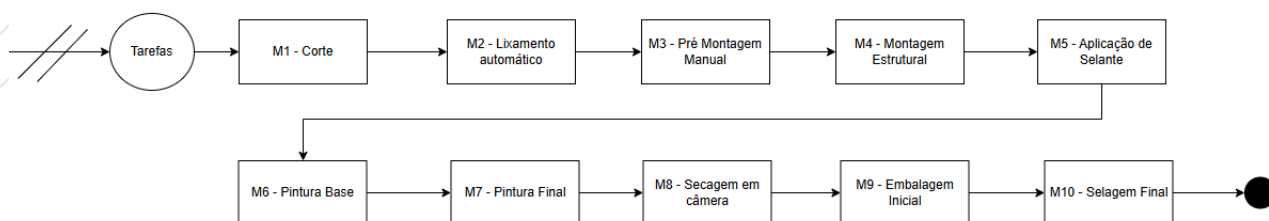
Para este trabalho, foi desenvolvido um cenário de uma fábrica de móveis, baseado em uma linha de produção contínua com dez máquinas, representando as etapas principais da fabricação. A simulação foi construída no *software* AnyLogic. Seguindo o protocolo ODD da Figura 3 e o modelo conceitual pelo IDEF-SIM na Figura 4.

Figura 3. Protocolo ODD resumido.

Overview (Visão Geral)	
Propósito	Flow shop com 10 máquinas; 20 jobs iniciais; minimizar múltiplos critérios; NEH (sequência inicial) + RL (ajustes dinâmicos).
Entidades	Máquina (M1–M10); Job (J1–J20 + urgentes): a processar nas 10 máquinas.
Variáveis de estado	Máquina: id; state; processingTime; failureRate 5%/h; repairTime (1 a 3h). Job: d; processingTimes[10] ($\pm 20\%$ inicial); priority (1 normal/2 urgente); state.
Escalas	Temporal: minutos; métricas em minutos. Espacial: 1 linha, 10 máquinas em série. Agentes: 10 Máquinas; 20+ Jobs.
Processo e escalonamento	NEH para sequência inicial; eventos: falha 5%/h, reparo 1–3 h, chegada urgente 10%/h; RL adaptativa reordena
Design Concept (Conceitos de Design)	
Princípios básicos	SBA + SED; tempo por eventos; agentes autônomos (Máquina/Job).
Emergência	Makespan e utilização emergem de: NEH+eventos (falha/urgente)+decisões RL
Adaptação	Falha $\rightarrow$ reescalonar; job urgente $\rightarrow$ priorização/reordenação;
Objetivos	Minimizar makespan; maximizar utilização; atender urgentes; comparar estratégias.
Aprendizado	RL por episódios; recompensa: $\downarrow$ makespan, $\uparrow$ utilização; política adaptativa.
Predição	ML para probabilidade de falhas/chegadas; previsão de impacto das ações no makespan.
Sensoriamento	Máquinas leem fila, estado próprio e job atual.
Interação	Job $\leftrightarrow$ Máquina (processo/tempo); Máquina $\leftrightarrow$ Eventos (falha/reparo); IA $\leftrightarrow$ Sistema (ações de despacho).
Estocasticidade	Tempos $\pm 20\%$; falha 5%/h; reparo 1-3h; chegada urgente 10%/h.
Coletivos	Máquinas: desempenho via utilização/makespan. Jobs: conclusão total.
Observação	Makespan, utilização, ociosidade, atraso, estabilidade de sequência.
Details (Detalhes)	
Inicialização	(1) Gerar 20 jobs; (2) NEH $\rightarrow$ sequência inicial; (3) Máquinas em Idle; (4) Fila ordenada por RL; (5) reset de métricas/estados.
Dados de entrada	Tempos base por tipo (mesa/cadeira/armário); falha 5%/h; reparo 1–3 h; chegada urgente 10%/h; 10 máquinas; 20 jobs.
Submodelos	Fila e despacho (NEH/RL/heurísticas); falhas e reparos; chegada urgente; Agente RL (observação/ação/recompensa); Métricas.

Fonte: Os autores.

Figura 4. IDEF-SIM.



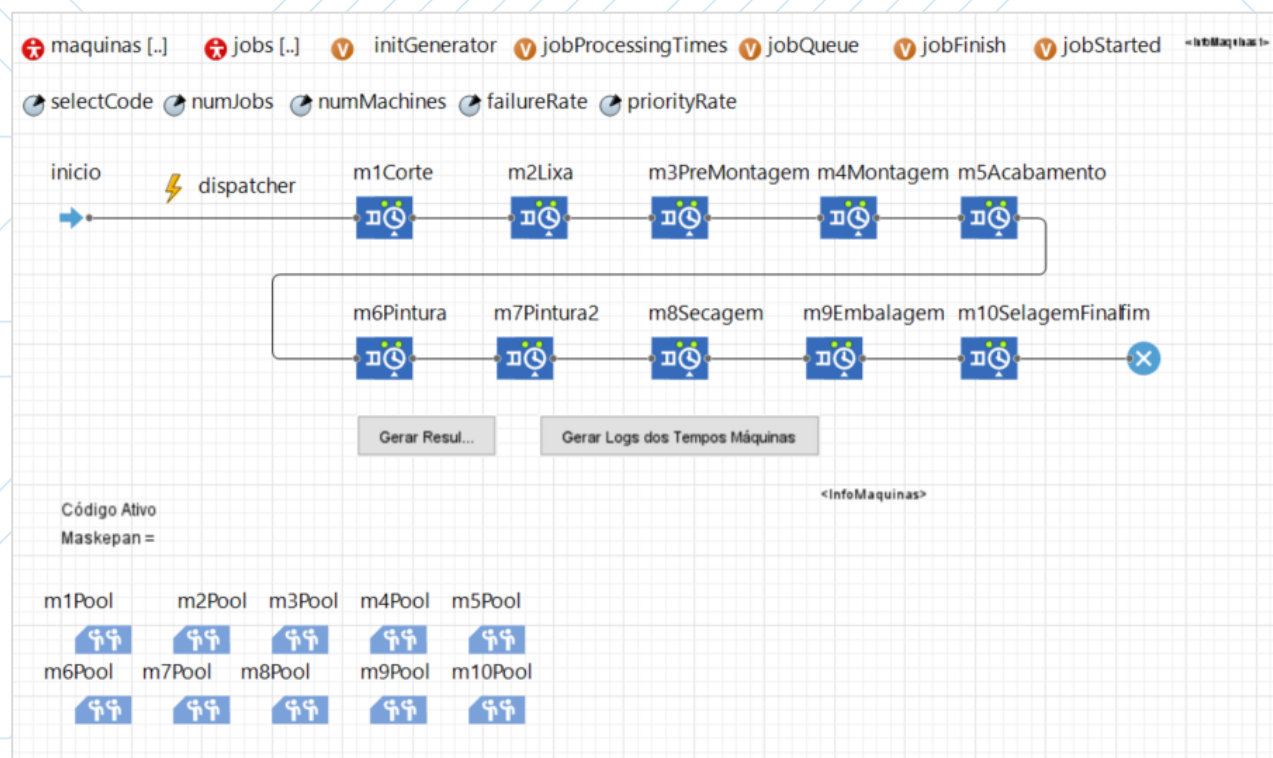
Fonte: Os autores.

3. RESULTADOS E DISCUSSÃO

Este trabalho encontra-se em desenvolvimento. A pesquisa está na fase de construção da simulação Figura 5. Foi implementada a camada de SED, que representa o fluxo de sequenciamento, os *jobs* são gerados em uma fonte (Início), aguardam em fila e são processados por um recurso máquina (*Service*), deixando o sistema ao final (Fim). Para elevar o realismo e incorporar comportamento autônomo, as entidades foram modeladas como agentes, compondo a camada de SBA. A Figura 6 ilustra cada máquina como um agente com estados e comportamentos próprios, o diagrama de estados permite representar eventos estocásticos e políticas de manutenção, viabilizando um ambiente de produção dinâmico. A Figura 7 representa cada tarefa (*job*) como agente que transita por estados ao longo do sistema. Modelar o *job* como agente abre espaço para regras de prioridade dinâmicas e comportamentos específicos por tarefa em etapas subsequentes.

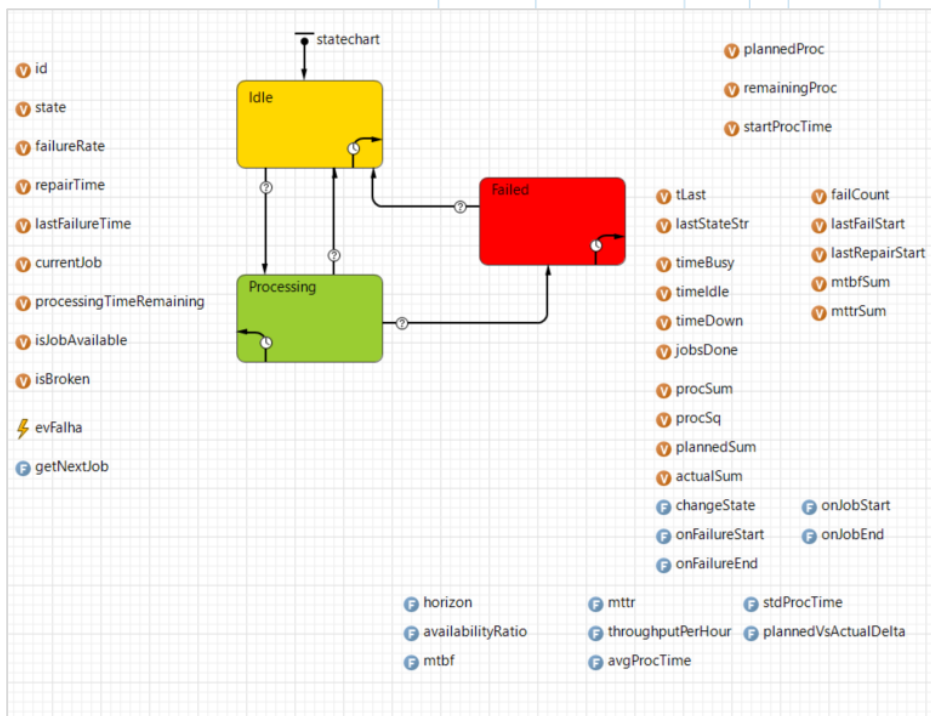
Nos resultados parciais, o RL foi integrado, utilizando a heurística NEH como política inicial para acelerar a exploração do espaço de sequências. No momento, estão em curso experimentos e a parametrização, com foco em ajuste de parâmetros e calibração dos componentes estocásticos do modelo.

Figura 5. Simulação Anylogic.



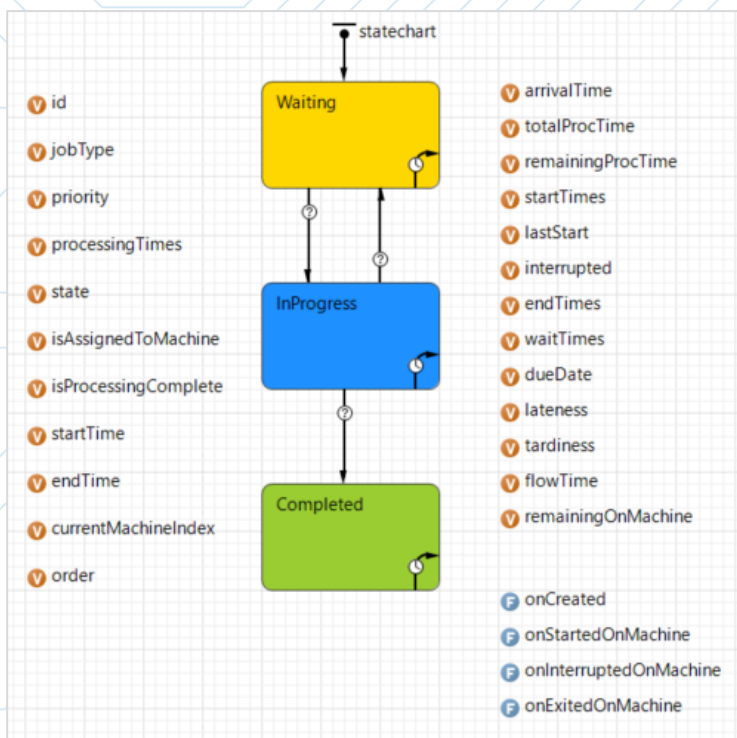
Fonte: Os autores.

Figura 6. Agente Máquina.



Fonte: Os autores.

Figura 7. Agente Job.



Fonte: Os autores.

4. CONSIDERAÇÕES FINAIS

Este trabalho está em desenvolvimento, mas apresentou uma abordagem para o PFSP que integra simulação híbrida e RL com NEH. O ambiente de simulação representou eventos estocásticos, permitindo comparar políticas de despacho sob múltiplos critérios, citados. Os experimentos em cenários representativos indicaram que a política aprendida pelo agente de RL tende mostrar ótimos resultados. A pesquisa reforça o potencial da SO para apoiar decisões em ambientes industriais.

AGRADECIMENTOS

Agradeço à Fundação de Amparo à Pesquisa do Estado de Minas Gerais pelo apoio (FAPEMIG), e ao Instituto de Engenharia de Produção e Gestão (IEPG).

REFERÊNCIAS

- DEL GALLO, M. *et al.* Artificial Intelligence to Solve Production Scheduling Problems in Real Industrial Settings: Systematic Literature Review. **Electronics**, 2023.
- ELBASHEER, M. *et al.* Flexible Symbiosis for Simulation Optimization in Production Scheduling: A Design Strategy for Adaptive Decision Support in Industry 5.0. **Journal of Manufacturing and Materials Processin**, 2024.
- GHASEMI, A. *et al.* Simulation optimization applied to production scheduling in the era of Industry 4.0: a review and future roadmap. **Journal of Industrial Information Integration**, v.39, 2024.
- GRIMM, V.; RAILSBACK, S. F. **Agent-Based and Individual-Based Modeling: A Practical Introduction**. 2. ed. Princeton: Princeton University Press, 2019.
- HE, Z. *et al.* Improved Q-learning algorithm for solving permutation flow shop scheduling problems. **IET Collaborative Intelligent Manufacturing**, v.4, n.1, 2022.
- JOHNSON, S. M. Optimal two- and three-stage production schedules with setup times included. **Naval Research Logistics Quarterly**, v.1, n.1, 1954.
- LAW, A. M. **Simulation Modeling and Analysis**. 5. ed. New York: McGraw-Hill, 2015.
- NAWAZ, M.; ENSCORE, E.; HAM, I. **A heuristic algorithm for the m-machine, n-job flow-shop sequencing problem**. Omega, v.11, n.1, 1983.
- NEUFELD, S.; SCHULZ, S.; BUSCHER, U. A systematic review of multi-objective hybrid flow shop scheduling. **European Journal of Operational Research**, v.309, n.1, 2023.

TURRIONI, B.; MELLO, P. **Metodologia de pesquisa em engenharia de produção**. Itajubá: Universidade Federal de Itajubá, 2012.

PINEDO, M. Scheduling: theory, algorithms and systems. **Englewood Cliffs: Prentice-Hall**, 2022.

SUTTON, S.; BARTO, G. **Reinforcement Learning: An Introduction**. 2. ed. MIT Press, 2018.

ZHANG, W.; BAO, X.; HAO, X.; GEN, M.; Metaheuristics for multi-objective scheduling problems in industry 4.0 and 5.0: a state-of-the-arts survey. **Front. Ind. Eng.** 2025.