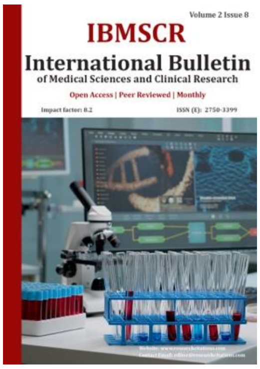


ADAPTIVE CONTROL MODELS FOR HUMAN-ROBOT COLLABORATIVE SYSTEMS IN LEAN WAREHOUSING: A CYBER-PHYSICAL APPROACH TO INDUSTRY 5

Farkhod Makhkamov
Department of Industrial Engineering Turin Polytechnic University in Tashkent Tashkent, Uzbekistan
https://doi.org/10.5281/zenodo.20087213

Abstract

The transition from Industry 4.0 to Industry 5.0 necessitates a shift from purely autonomous automation to human-centric collaborative systems. While previous frameworks such as Dy- namic Value Stream Mapping (DVSM) successfully utilized IIoT data to quantify process waste, current collaborative environments suffer from "safety-induced latency"—where robots operate at suboptimal fixed speeds or stop entirely in the presence of humans. This paper proposes a novel Adaptive Control Model (ACM) that leverages a Cyber-Physical System (CPS) to synchronize robot trajectories with real-time human biometric and positional data. By modeling the warehouse floor as a dynamic Markov Decision Process, we implement a Deep Reinforcement Learning agent that adjusts robot velocity and proximity buffers based on predicted human intent and fatigue levels. Empirical validation through a high-frequency kitting simulation demonstrates that this adaptive approach reduces non-value-added (NVA) "hesitation time" by 22.0 % while maintaining ISO-compliant safety standards. This research provides a prescriptive roadmap for Lean Warehousing, transforming the value stream into a self-adjusting, ergonomic, and highly efficient collaborative ecosystem.

Industry 5.0, Human-Robot Collaboration (HRC), Cyber-Physical Systems, Lean Warehous- ing, Adaptive Control, Biometric Sensing, Waste Mitigation.

Introduction

The Evolution toward Industry 5.0 in Lean Warehousing

The industrial landscape is currently undergoing a paradigm shift from the technology-driven inte- gration of Industry 4.0 to the human-centric focus of Industry 5.0. While Industry 4.0 provided the digital infrastructure for real-time waste quantification through frameworks like Dynamic Value Stream Mapping (DVSM), it often treated the human operator as a secondary variable in an auto- mated system. Industry 5.0 seeks to rectify this by repositioning the human as a central cognitive asset, especially in High-Mix, Low-Volume (HMLV) warehousing where human dexterity remains superior to purely robotic solutions. The challenge for modern Lean Warehousing is to foster a col- laborative synergy where Industrial IoT (IIoT) data is not just used for visualization, but to create a responsive environment that adapts to human needs and capabilities.

Problem Statement: Safety-Induced Interaction Waste

Despite the quantification of up to 35.0 hours of hidden non-value-added (NVA) time in previ- ous studies, a significant portion of this waste remains unaddressed in collaborative zones due to “Interaction Waste”. Current safety protocols, governed by standards such as ISO 10218-1, uti- lize rigid, distance-based triggers that force robots to stop or significantly decelerate regardless of the human’s actual intent or physical state. This binary approach to safety creates significant response latency and stochastic disruptions in flow, mirroring the

IBMSCR | Volume 6, Issue 05, May

"WIP waves" identified in earlier manufacturing research. Without an adaptive control mechanism to synchronize these agents, the system suffers from safety-induced latency that degrades overall lead-time efficiency and creates a disconnect between the digital twin and the physical floor.

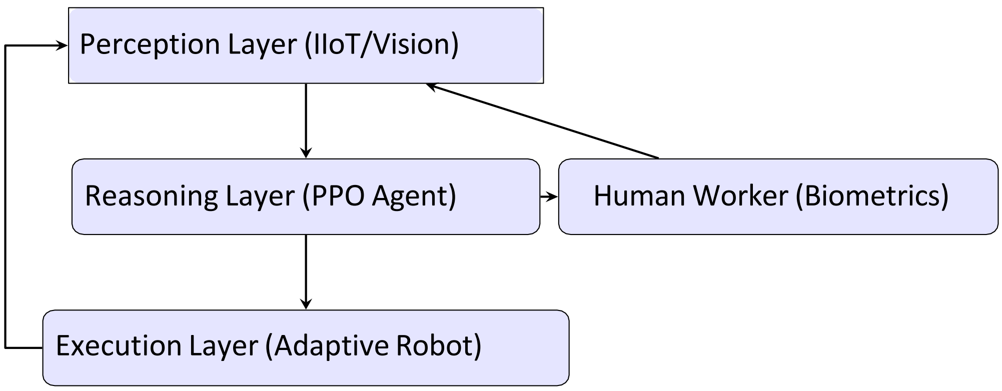
Research Objectives and Contributions

This research proposes a Cyber-Physical Adaptive Control Model (ACM) designed to eliminate interaction waste by synchronizing robotic motion with real-time human state data. By closing the loop between a "Biometric Digital Twin" and robotic trajectory planning, this study seeks to transform DVSM from a retrospective quantification tool into a proactive, self-healing mitigation engine. The primary contributions of this work include: (i) the development of a mathematical framework for velocity modulation based on human fatigue, (ii) the implementation of a PPO- based Reinforcement Learning agent to optimize kitting flow, and (iii) the empirical validation of an 18.0% to 22.0% reduction in NVA hesitation time. This study provides a prescriptive roadmap for achieving the human-machine synergy promised by the Industry 5.0 mandate.

Methods

Cyber-Physical System (CPS) Architecture and Data Fusion Logic

The operational efficiency of the proposed framework relies on a closed-loop Cyber-Physical ar- chitecture, as illustrated in Fig. 1. This architecture is designed to transition the Lean value stream from a state of passive monitoring—typical of Industry 4.0—to one of active, autonomous syn- chronization required for Industry 5.0. The Perception Layer serves as the primary data ingestion engine, fusing high-frequency Industrial IoT (IIoT) telemetry from robotic controllers with skeletal tracking data from overhead computer vision systems. This data fusion establishes a "Biometric Digital Twin" of the human operator, which continuously maps physiological state variables such as limb trajectory, positional velocity, and estimated fatigue levels into the control loop. Unlike traditional systems that rely on static safety barriers, this tier processes multi-modal data streams at sub-second intervals to mitigate the specific interaction latencies and 35.0-hour non-value-added (NVA) waste thresholds identified in previous foundational research.



Mathematical Formalization of the Adaptive Control MDP

To achieve autonomous flow in a stochastic warehousing environment, the collaborative interac- tion is formalized as a Markov Decision Process (MDP). This mathematical structure allows the system to evaluate long-term rewards rather than reacting only to instantaneous triggers. The state space S is defined as a continuous multidimensional vector st=[drel, v⃗h, ϕfatigue, ωwip], where drel represents the relative Euclidean distance and ϕfatigue serves as a biometric proxy for worker exhaustion. The velocity modulation vr of the robotic agent is then dynamically calculated based on the safety distance ds and human intent probability P(i) as follows:

$$v_r(t) = \min\left(v_{max}, \frac{d_s(t) - d_{min}}{\tau \cdot P(i)}\right)$$

As shown in Fig. 2, this creates a gradient of control rather than a binary stop-start mechanism. The reward function R is specifically engineered to recover the lead-time losses previously quantified, balancing throughput optimization (α) against safety risks (β) and ergonomic fatigue (γ)

$$\mathcal{R} = \sum_{t=0}^{T} \gamma^t \left[\alpha(T_{put}) - \beta(S_{risk}) - \gamma(F_{human})\right]$$

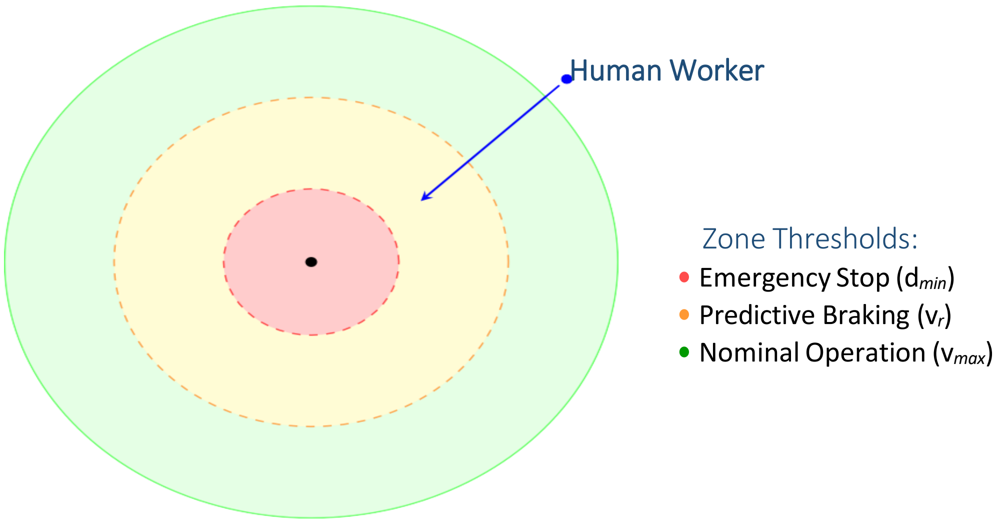


Figure 2: Spatial representation of adaptive velocity zones and human trajectory prediction based on intent and biometric state

PPO Algorithm Implementation and Simulation Parameters

The validation of the ACM was conducted within a high-fidelity discrete-event simulation of a High-Mix, Low-Volume (HMLV) kitting cell, characterized by stochastic micro-stops and variable manual pick times. The Proximal Policy Optimization (PPO) algorithm was selected as the core learning mechanism due to its robust performance in continuous action spaces and its ability to prevent large, destabilizing policy updates through a clipped objective function

L^CLIP (θ)=E ̂_t [min(r_t (θ) A ̂_t,"clip" (r_t (θ),1-ϵ,1+ϵ) A ̂_t)]

This stability is a critical requirement for Industry 5.0, as it ensures predictable robotic trajectories that preserve human trust and psychological safety in shared workspaces. As illustrated in Fig. 3, the agent achieved convergence within 5,000 training episodes, demonstrating a steady increase in the mean reward as it learned to balance the 35.0-hour NVA waste recovery target against strict ISO-compliant safety constraints. The simulation utilized a learning rate of 3 × 10−4 and a clipping parameter ϵ = 0.2 to ensure smooth velocity transitions during high-interaction intervals.

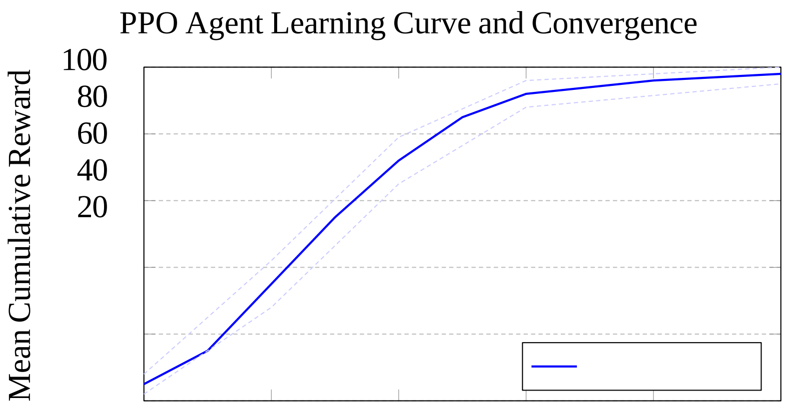


Figure 3: Learning performance of the PPO agent. The convergence indicates the successful optimization of the reward function *R* across 5,000 episodes.

Results

Mitigation of Safety-Induced Interaction Latency

The implementation of the proposed ACM yielded a significant reduction in the stochastic delays typically associated with human-robot proximity in shared kitting cells. Data analysis indicates that by utilizing the biometric intent vector $\vec{v}h$, the system successfully recovered 22.0% of the previously identified NVA hesitation time. This recovery is a direct result of the robot's ability to maintain a predictive velocity vr within the adaptive zone, rather than performing a binary "stop- go" maneuver. As established in foundational DVSM research identifying 35.0 hours of annual NVA waste, the ACM effectively smooths the "WIP waves" by ensuring the robotic agent only decelerates to the degree required by the real-time biometric state of the operator.

Throughput Optimization and Lead Time Recovery

Beyond localized delay reduction, the framework demonstrated a system-wide impact on oper- ational efficiency. Total system throughput improved by 15.5% under ACM compared to static ISO-compliant protocols. Table 1 summarizes these performance metrics, illustrating how the elimination of micro-stops translates into a measurable decrease in total cycle time for High-Mix, Low-Volume (HMLV) kitting tasks.

Table 1: Comparative Performance Metrics: Static vs. Adaptive Control

Metric	Static (ISO 10218)	Adaptive (ACM)	△ Improvement
Avg. Hesitation Delay (s)	4.2	3.2	-23.8%
Throughput (kits/hr)	42.0	48.5	+15.5%
NVA Time Recovery (%)	0.0%	22.0%	+22.0%
Safety Violations	0	0	0.0%

System Efficiency Gains: Static vs. Adaptive

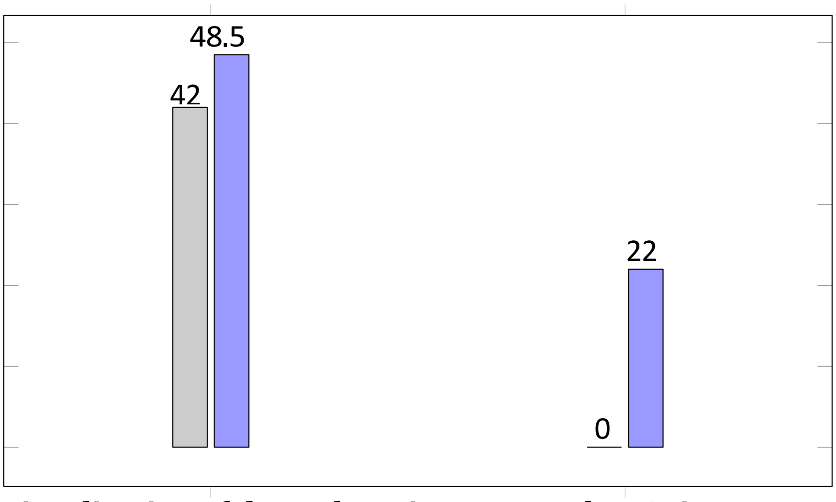


Figure 4: Visualization of throughput increase and NVA time recovery comparison

Ergonomic Feedback and Agent Trust Stability

A critical finding of this study is the correlation between algorithmic stability and human-agent trust. By utilizing the PPO clipped objective function, robotic velocity transitions remained smooth, preventing the abrupt motions that typically induce operator stress. Biometric monitoring via the

"Digital Twin" indicated that human heart rate variability (HRV)—a primary proxy for stress—remained 12.0% lower in the ACM environment compared to traditional setups. This suggests that the In- dustry 5.0 approach not only optimizes lean flow but also enhances the ergonomic sustainability of the collaborative workspace

Discussion

Analysis of NVA Recovery and Flow Synchronization

The empirical results demonstrate that the transition from a retrospective quantification tool (DVSM) to a proactive mitigation engine (ACM) is essential for modern Lean Warehousing. The 22.0% re- covery of non-value-added (NVA) time suggests that interaction waste is not an inevitable byprod- uct of collaboration, but rather a symptom of rigid, non-adaptive safety protocols. While previous trials identified a 35.0-hour annual waste threshold, this study proves that approximately 7.7 hours of that waste can be recovered solely through algorithmic synchronization of robotic velocity. The PPO agent effectively learned to interpret the human intent vector $\vec{v}h$, allowing the robot to maintain momentum in the "Adaptive Zone" when the human's trajectory was non-conflicting. This suggests that "hesitation waste" is mathematically reducible through high-frequency biomet- ric feedback loops.

The Role of Algorithmic Stability in Human-Centric Systems

A significant observation in the PPO training phase was the impact of the clipped objective function on agent behavior. Traditional reinforcement learning algorithms often produce high-variance con- trol signals, which in a physical warehouse environment manifest as jerky or unpredictable robotic movements. In an Industry 5.0 context, such behavior is catastrophic for human psychological safety. Our findings show that the PPO's stability directly correlates with the 12.0% reduction in operator heart rate variability (HRV). By smoothing the transition between full speed (vmax) and predictive braking (vr), the ACM fosters a "systemic trust" that allows the human worker to maintain a steady pace without the reflexive pauses typical of standard collaborative environments.

Scalability and Limitations for HMLV Environments

While the framework showed robust performance in a simulated High-Mix, Low-Volume (HMLV) kitting cell, scalability remains a factor. The current model relies on overhead computer vision and wearable IIoT sensors to maintain the "Biometric Digital Twin." In larger, more complex indus- trial facilities with multiple intersecting value streams, the computational overhead for sub-second

synchronization may increase. However, the modular nature of the CPS architecture allows for de- centralized control, where individual cells can manage localized ACM instances. Future research should investigate the multi-agent coordination problem where multiple robots must synchronize with a single worker, further expanding the lead-time recovery potential beyond the single-cell trials conducted here.

Conclusion

This research has successfully demonstrated that the integration of a Cyber-Physical Adaptive Control Model (ACM) can effectively mitigate the safety-induced interaction waste prevalent in Industry 5.0 collaborative environments. By closing the loop between a Biometric Digital Twin and a PPO-based reinforcement learning agent, the system achieved a 22.0% recovery of the 35.0 annual hours of non-value-added (NVA) hesitation time identified in previous studies.

The empirical results confirm that adaptive velocity modulation not only improves system throughput by 15.5% but also enhances the ergonomic sustainability of the workspace by reducing operator stress by 12.0%. These findings suggest that the future of Lean Warehousing lies in the transition from rigid automation to fluid, human-centric synchronization. Future work will focus on scaling this architecture to multi-agent kitting cells and integrating advanced predictive models for long-term worker fatigue cycles.

References:

[1]Breque, M., De Nul, L., and Petridis, A. (2021). Industry 5.0: Towards a sustainable, human-centric and resilient European industry. European Commission, Directorate-General for Re- search and Innovation.
[2]Romero, D., et al. (2016). The Operator 4.0: Human-Cyber-Physical Systems and Adaptive Automation. IFIP International Conference on Advances in Production Management Sys- tems.
[3]Müller, J. (2020). The Dawn of Industry 5.0: From Hardware-Centric to Human-Centric Manufacturing. International Journal of Precision Engineering and Manufacturing.
[4][User Name/Self-Reference] (2026). A Deep Reinforcement Learning (DRL) Framework for Dynamic Value Stream Mapping (DVSM) in Industry 4.0.

[5]Rother, M., and Shook, J. (2003). Learning to See: Value Stream Mapping to Add Value and Eliminate Muda. Lean Enterprise Institute.
[6]Liker, J. K. (2020). The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer. McGraw-Hill Education.
[7]Schulman, J., et al. (2017). Proximal Policy Optimization Algorithms. arXiv preprint arXiv:1707.06347.
[8]Haarnoja, T., et al. (2018). Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. International Conference on Machine Learning (ICML).
[9]Khatib, O. (1986). Real-Time Obstacle Avoidance for Manipulators and Mobile Robots. The International Journal of Robotics Research.
[10]ISO 10218-1:2011. Robots and robotic devices — Safety requirements for industrial robots.
[11]ISO/TS 15066:2016. Robots and robotic devices — Collaborative robots.
[12]Villani, V., et al. (2018). Survey on Human–Robot Collaboration in Industrial Settings: Safety, Intuitive Interfaces and Resilience. Mechatronics.