International Conference on Formal Methods
and Foundations of Artificial Intelligence
Eszterházy Károly Catholic University
Eger, Hungary, June 5–7, 2025

Reinforcement Learning for Inventory Management in Reverse Supply Chains of Wood

Nikola Kovačević[a], **Balázs Dávid**[a b]

[a]Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska
nikola.kovacevic@upr.si

[b]InnoRenew CoE
balazs.david@innorenew.eu

Abstract

Traditional supply chain models typically emphasize forward flows, yet reverse supply chains play a critical role in promoting sustainability by recovering and reusing end-of-life waste materials. In this study, we examine inventory management decisions within reverse supply chains by addressing waste generation, collection, and treatment processes under stochastic conditions. Unlike conventional linear disposal methods, such as landfilling, our approach adopts a model in which waste is used as a resource, leading to reduced environmental impact.

A simulation framework will be presented that models dynamic inventory levels in such a reverse supply chain. The inventory management decisions of the model ensure efficient flow and transformation of reusable resources in the system and avoid landfilling. A discrete event simulation model will be formulated for the problem and reinforcement learning techniques will be developed to adaptively optimize inventory management decisions under stochastic resource availability and demand. The efficiency of the proposed approach will be tested on networks generated based on real-world statistical data.

1. Theoretical Background

Modern environmental and economic goals call for a transition from conventional waste management to supply chains that prioritize material recovery and reuse. Recent advancements in reinforcement learning demonstrate significant potential for optimizing these complex systems [5], particularly in managing stochastic demand patterns and dynamic inventory decisions inherent to reverse logistics [3]. Stranieri and Stella [6] demonstrate that deep reinforcement learning algorithms (PPO, A3C) can outperform traditional inventory policies in complex two-echelon recovery networks, suggesting similar approaches could resolve the temporal-spatial mismatches common in waste material flows. However, most existing studies focus on simplified supply chains and do not consider domain-specific issues.

One such challenge is material complexity. Malinverno et al. [4] have developed a material flow analysis framework that systematically maps 110 material flows through 27 value chain processes. Their methodology for addressing data harmonization in cascading material use provides a foundation for modeling reverse logistics networks handling heterogeneous waste streams with various degradation states and contamination levels. Farjana and Ashraf [2] identified 12 key performance indicators for wood waste supply chains, including transportation distance optimization and material circularity index. However, unlike their focus on static waste classification (treated/untreated/engineered), there is a need for dynamic inventory optimization techniques specifically designed for reverse supply chains' unique characteristics - intermittent supply and uncertain quality and quantity of resources.

2. Problem Statement

Building on the identified challenges and insights, we adapt reinforcement learning techniques to address the inventory management questions of reverse supply chains where variable material quality and the uncertain amount of recovered resources limit traditional optimization approaches. These techniques will help decision-making in a discrete event simulation model for the reverse supply chain of waste wood, where three main types of entities (generators, collectors and treatment facilities) manage the resource flows in the system. Generators periodically produce uncertain quantities of waste, which can be then collected by collectors and stored for future use by treatment facilities.

At every level of this network, decisions have to be made about storing or landfilling waste, as well as the potential increase of existing storage capacities to avoid inventory overflow. The goal is to satisfy the uncertain demand in the system, while minimizing both costs and environmental impact.

3. Results

The proposed simulation framework and reinforcement learning approach will be tested on network structures generated based on real-world data from the Slovenian Environment Agency (ARSO) [1], evaluating scenarios with constant and stochastic waste generation rates and demand patterns. Different probability distributions will model key uncertainties, allowing assessment of the system's performance under realistic conditions.

The model performance will be analyzed for both economic and environmental objectives across scenarios. Special focus will be given to reduced landfilling rates and improved resource utilization, validating the effectiveness of the approach to improving circularity. Sensitivity analyses will further examine how critical parameters influence system performance, providing insights for practical implementation in reverse supply chains of wood.

Acknowledgements Balázs Dávid is grateful for the support of the Slovenian Research and Innovation Agency (ARIS) through grant J1-50000.

References

[1] Agencija Republike Slovenije za Okolje (ARSO) / Environmental Agency of the Republic of Slovenia: *Poročila in publikacije o odpadkih / Reports and Publications on Waste*, slovenian, english, Accessed: 2025-04-25, 2023, url: https://www.arso.gov.si/varstvo%20okolja/odpadki/poro%c4%8dila%20in%20publikacije/.

[2] S. H. Farjana, M. Ashraf: *Developing the Conceptual Framework for the Key Performance Indicators for Sustainable Wood Waste Supply Chain*, Environment, Development and Sustainability (Dec. 18, 2023), issn: 1387-585X, 1573-2975, doi: 10.1007/s10668-023-04174-0, url: https://link.springer.com/10.1007/s10668-023-04174-0 (visited on 01/21/2025).

[3] L. Kemmer, H. von Kleist, D. de Rochebouët, N. Tziortziotis, J. Read: *Reinforcement learning for supply chain optimization*, in: European Workshop on Reinforcement Learning 14, Lille, France, 2018.

[4] N. Malinverno, S. Buschor, K. Vogel, F. Schwarze, G. Stadelmann, E. Thürig, B. Nowack, G. Nyström, C. Som: *Resolving Complexity: Material Flow Analysis of a National Wood Flow System Integrating the Versatility of Wood*, Journal of Industrial Ecology 28.6 (Dec. 2024), pp. 1716–1729, issn: 1088-1980, 1530-9290, doi: 10.1111/jiec.13560, (visited on 01/26/2025).

[5] B. Rolf, I. Jackson, M. Müller, S. Lang, T. Reggelin, D. Ivanov: *A Review on Reinforcement Learning Algorithms and Applications in Supply Chain Management*, International Journal of Production Research 61.20 (Oct. 2023), pp. 7151–7179, issn: 0020-7543, 1366-588X, doi: 10.1080/00207543.2022.2140221, (visited on 04/25/2025).

[6] F. Stranieri, F. Stella: *Comparing Deep Reinforcement Learning Algorithms in Two-Echelon Supply Chains*, in: Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Springer Nature Switzerland, 2025, pp. 454–469, isbn: 9783031746406, doi: 10.1007/978-3-031-74640-6_37, url: http://dx.doi.org/10.1007/978-3-031-74640-6_37.