Enhancing Circularity in Medical Device Supply Chains by Optimizing EoL Decisions Through Reinforcement Learning: A Multi-Objective Approach

Soufiane El Bechari[1], Oualid Jouini[1][a], Zied Jemai[1,2], Fourat Trabelsi[1] and Robert Heidsieck[3]

[1]*Industrial Engineering Laboratory (LGI) CentraleSupélec, Paris-Saclay University, Gif-sur-Yvette, France*

[2]*OASIS - ENIT, University of Tunis, Elmanar, BP37, Le Belvedere, Tunisia*

[3]*GE HealthCare, 283 Rue de la Minière, 78530 Buc, France*

{*soufiane.el-bechari, oualid.jouini, zied.jemai*}*@centralesupelec.fr, fourat.trabelsi@student-cs.fr,*

Keywords: Multi-Component Systems, Circular Supply Chain, EoL Management, Reinforcement Learning, Multi-Objective Approach.

Abstract: Circular supply chains are becoming essential in the pursuit of sustainability, as they promote the responsible disposal, recycling, and reuse of products at the end of their life cycles. This research, developed in collaboration with GE HealthCare, presents a multi-objective optimization framework that incorporates environmental, economic, and circularity performance in end-of-life (EoL) decision-making. The proposed model leverages historical data on reuse and recycling success rates to capture the operational realities of circular supply chains. By employing Q-learning, this paper aims to develop a decision-support mechanism that optimizes EoL actions for components, thereby enhancing the circularity, reducing carbon footprint, and minimizing economic costs within the circular supply chain.

1 INTRODUCTION

The increasing emphasis on Sustainable and Circular Supply Chain Management (SCSCM) within the healthcare sector reflects a critical acknowledgment of the environmental and economic impacts of medical supply chain practices. This shift is driven by the need to reconcile healthcare operations with sustainability goals, as the sector is responsible for significant waste generation and carbon emissions, contributing to global environmental challenges (D'Alessandro et al., 2024). For instance, the healthcare sector is responsible for around 4.6 % to 5 % of global greenhouse gas (GHG) emissions (Pichler et al., 2019; Romanello et al., 2023), equivalent to 2 billion carbon dioxide equivalent (CO2e). Given the significant impact of the healthcare sector on climate change, there have been a number of policy initiatives aimed at reducing the environmental footprint, most notably through the NHS's "Delivering a Net Zero National Health Service" strategy. the NHS, as one of the world's largest healthcare systems, has set targets to achieve net-zero emissions by 2040 for emissions under its direct control, and by 2045 for those it can influence indirectly, such as those from the supply chain and patient travel (NHS, 2022). Additionally, major healthcare companies like GE HealthCare are aligning their sustainability goals with these broader initiatives. GE HealthCare has set goals to reduce operational GHG emissions (Scope 1 and 2) by 42% and Scope 3 emissions by 25% by 2030, as part of their commitment to reaching net zero by 2050 (GE Healthcare, 2023). The intersection of healthcare and environmental sustainability is becoming increasingly prominent as the global efforts faces the dual challenges of delivering quality healthcare while combating climate change and minimizing waste. In this context, the medical supply chain plays a pivotal role in addressing environmental concerns, especially in efforts to reduce carbon emissions and waste generation (Abaku and Odimarha, 2024). The medical supply chain, essential for delivering healthcare products like pharmaceuticals and medical devices, is a highly intricate system that significantly impacts the environment. Due to the specific nature of medical equipment, significant efforts have been made in design, operations, and supply chain management to maintain

[a] https://orcid.org/0000-0002-9498-165X

El Bechari, S., Jouini, O., Jemai, Z., Trabelsi, F. and Heidsieck, R.
Enhancing Circularity in Medical Device Supply Chains by Optimizing EoL Decisions Through Reinforcement Learning: A Multi-Objective Approach.
DOI: 10.5220/0013122400003893

In *Proceedings of the 14th International Conference on Operations Research and Enterprise Systems (ICORES 2025)*, pages 88-99
ISBN: 978-989-758-732-0; ISSN: 2184-4372

operating conditions and support circular economy principles to reduce environmental impact. But still a need for further innovations and research, particularly in optimizing the supply chain of medical equipment through the integration of circular economy practices to enhance both circularity and sustainability within the healthcare sector.

Unlike traditional linear supply chains, which follow a 'take-make-dispose' approach, circular supply chains aim to "integrate circular economy thinking into supply chain management and its surrounding industrial and natural ecosystems" (Farooque et al., 2019). A "circular economy is a regenerative economic system which necessitates a paradigm shift to replace the 'end of life' concept with reducing, alternatively reusing, recycling, and recovering materials throughout the supply chain, with the aim to promote value maintenance and sustainable development, creating environmental quality, economic development, and social equity, to the benefit of current and future generations" (Kirchherr et al., 2023). This circular supply chains aim to extend the life of products, components and materials through CE strategies such as reuse, recycling, and remanufacturing. These circular actions allow businesses to reduce their dependency on virgin materials, minimize waste, and lower their overall environmental footprint.

One of the primary challenges in implementing circular supply chains is assuring effective returns and managing the EoL's of products by determining whether to reuse, recycle, or dispose of returned used products and components. Each of these decisions carries circular, environmental, and economic implications. For example, reusing components reduces the need for new materials but may be constrained by technical or quality limitations. Recycling can recover valuable materials, but the associated energy and costs may be significant. Finally, disposal results in increased waste and environmental impact, but in some cases, it may be the only viable option if reuse or recycling is not feasible or not the most pertinent and appropriate solution.

Moreover, circular supply chains are often complex and involve multiple components, each with unique environmental and economic profiles. Decision-making in this context requires careful consideration of trade-offs between minimizing environmental impact, reducing costs, and maximizing circularity (i.e., the proportion of products that are successfully reused or recycled). To the best of our knowledge, current models in circular supply chain management often fail to fully integrate these multiple objectives and do not incorporate real-world success rates for reuse and recycling, leading to unrealistic expectations about circularity potential.

This study addresses this challenge by developing a decision-support mechanism utilizing Q-learning, a reinforcement learning (RL) technique, to optimize the management of EoL components in a CSC. The model operates at the component level, enabling real-time decision-making for reuse, recycling, and disposal actions, based on component-specific parameters and performance metrics.

The rest of the paper is organized as follows. Section 2 provides an overview of the state-of-the-art. In Section 3, the problem formulation and modeling are presented, along with a detailed description of the proposed algorithm to solve the problem. Section 4 introduces an industrial case study based on real-world data, including results and analysis. Finally, Section 5 concludes the paper and highlights future perspectives.

2 LITERATURE REVIEW

As discussed in the introduction, the primary contributions of this paper are in the domains of circular supply chain and EoL management, and multi-objective decision-making. Specifically, we focus on identifying the most effective strategies for optimizing reuse, recycling, and disposal actions within complex multi-component circular supply chains. In the following section, we review the key streams of literature in these areas, including circular supply chains, and EoL management. We then position the contributions of this paper within the broader context of existing research, highlighting the novelty of our approach in balancing economic, environmental, and circularity objectives across multiple components.

2.1 Circular Supply Chains and EoL Management

The integration of Circular Economy (CE) principles into Supply Chain Management (SCM) has been widely referred to as Circular Supply Chain Management (CSCM) in the literature (Genovese et al., 2017; Nasir et al., 2017; Farooque et al., 2019). CSCM encompasses various definitions, each emphasizing the role of CE in reshaping supply chain activities (Lahane et al., 2020). According to Farooque et al. (2019), "Circular supply chain management is the integration of circular thinking into the management of the supply chain and its surrounding industrial and natural ecosystems. It systematically restores technical materials and regenerates biological materials toward a zero-waste vision through system-wide innovation in business models and supply chain functions from product/service design to EoL and waste management, involving all stakeholders in a product/service lifecycle, including parts/product manufacturers, service providers, consumers, and users".

The adoption of CE practices within supply chains offers several key benefits (Lahane et al., 2020), including enhanced resource availability (Goyal et al., 2018), improved EoL strategies (de Sousa Jabbour et al., 2019), enriched value propositions (Mishra et al., 2018), reduced waste generation (Herczeg et al., 2018), and improved sustainability (Winkler, 2011).

However, the transition from traditional, linear supply chain models—characterized by the "take-make-dispose" framework—towards circular systems presents several significant challenges. These challenges, as noted by Roy et al. (2022), include a persistent industrial preference for linear models, compounded by feasibility concerns surrounding CE implementation (Tura et al., 2019; Agyemang et al., 2019), and the absence of robust performance measurement systems (Tura et al., 2019). Additionally, higher upfront costs related to circular business models (Vermunt et al., 2019) and complex product designs that hinder the ease and cost-effectiveness of recycling, reuse, or remanufacturing (Khodier et al., 2018; Rosa et al., 2019; Halse and Jæger, 2019) remain significant barriers to wide-scale adoption.

Moreover, the absence of standardized circular economy processes and metrics across industries hinders cross-industry implementation of circular models (Govindan and Hasanagic, 2018; Mangla et al., 2018; Ranta et al., 2018; Bouzon et al., 2018). This gap, combined with the lack of widely accepted metrics and indicators to measure circularity performance (Kravchenko et al., 2019; Bressanelli et al., 2019), limits the scalability and effectiveness of CSCM.

The healthcare sector, in particular, presents unique challenges for implementing circular economy principles due to stringent regulatory standards and the critical importance of hygiene. Research highlights significant barriers to circularity in medical equipment, including perceived safety risks, regulatory complexities, and financial constraints related to medical device design (Hoveling et al., 2024).

Beyond these general barriers, the complexity intensifies when it comes to managing the EoL of multi-component products, such as electronics and medical devices in CSC. In these cases, each component may have its own unique lifecycle and recovery potential for reuse or recycling. Han et al. (2021) highlight the critical need for component-level analysis to determine the optimal EoL strategy, which should consider factors such as regional greenhouse gas (GHG) emissions and market prices for resale. This is particularly relevant in assembly-based products, where entire units are often discarded as waste, even though some components retain value that could be recovered through reuse or recycling (Kinoshita et al., 2016).

The process of disassembling complex products into individual components for reuse, recycling, or disposal presents opportunities to prevent the unnecessary consumption of virgin materials and reduce GHG emissions (Hasegawa et al., 2019) . These actions contribute significantly to the circularity of supply chains by maximizing the lifecycle of each component. However, the disassembly process can be resource-intensive, requiring significant labor and cost, making it essential to anticipate the potential outcomes of reuse, recycling, and disposal before implementation. As Hasegawa et al. (2019) point out, simulating these decisions is crucial to optimize EoL management in multi-component systems, allowing for the efficient allocation of resources and the reduction of environmental impact.

Despite the growing body of research on CSCM, there remains a significant gap in the literature concerning EoL management and its broader impact on the effectiveness of circular supply chains. Specifically, limited attention has been paid to quantitative models that simulate EoL decision-making for multiple components, considering factors such as component-specific reuse and recycling potential, cost, and environmental impact within a multi-objective approach. Addressing this gap, the present study proposes a quantitative model that simulates circular supply chain decisions at the component level, optimizing the actions of reuse, recycling, and disposal. The model developed here provides decision support for EoL scenarios involving multiple components within a single product, offering a more granular understanding of how different EoL strategies affect circularity, cost, and environmental outcomes. By adopting a detailed, component-by-component approach, this study enhances decision-making in circular supply chains, providing a decision-support mechanism for addressing the inherent complexities of multi-component EoL management in CSC.

3 PROBLEM DESCRIPTION AND MODELING

In this paper, the term "product" or "device" refers to spare parts that can either be used in the production of new large-scale equipment or for maintenance services to support an installed base, ensuring the operational continuity of the equipment. We focus on a circular supply chain (CSC) that integrates suppliers, manufacturers, customers (installed base), and reverse logistics for the take-back of defective products. These defective products are typically returned from the field and are not directly repairable or reusable in their entirety.

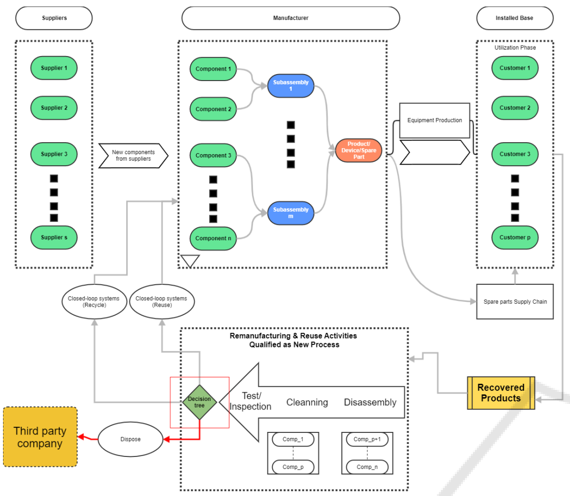

Figure 1: Circular supply chain model and its characteristics.

As shown in Fig. 1, the circular supply chain model consists of various stages including disassembly, reuse, recycling, and disposal. Upon return, defective products are disassembled to extract reusable components. These components are then cleaned, inspected, and tested to determine whether they meet the required specifications for reuse. Components that pass inspection are classified as ”qualified as new” and are reintroduced into the production process for manufacturing new products.

In this circular supply chain, the demand from manufacturer must be satisfied either through the closed-loop system—by reusing or recycling components—or by procuring new materials or components from suppliers. The goal is to maximize the use of reused and recycled materials. However, when the reuse or recycling process cannot meet the required demand, procurement from suppliers is necessary to ensure production continuity. For components that do not meet reuse requirements, two options are considered. If the components are feasible for recycling and the recycling process is legally compliant, the materials are recycled and used in the production of new products, forming a closed-loop system. If recycling is not feasible, either due to design limitations or legal restrictions, the components are disposed of by a third-party company, incurring an additional disposal cost.

Moreover, for components that are non-reusable by design, the same decision-making process applies. If recycling is possible, the materials are processed accordingly. If recycling is not an option, these components are handled by a third-party disposal service, which incurs a disposal cost. As depicted in Fig. 2, a decision tree outlines the methodology for managing each component's EoL on a component-by-component basis.

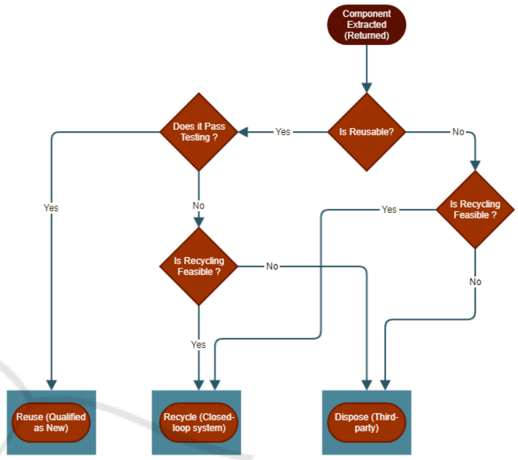

Figure 2: A decision tree for managing components' EoL in a component-by-component methodology.

Thus, for each component within the circular supply chain, a decision must be made between three EoL options: reuse, recycling within a closed-loop system, or disposal by a third-party.

The aim of this research is to develop a decision-making framework for managing returned components in a circular supply chain using a component-by-component methodology, specifically focusing on optimizing these EoL decisions for each component. The complexity of the problem arises from the variability in return flows, the varying success rates for reuse and recycling, and the economic and environmental considerations for each component.

3.1 Agent-Based Modeling of Circular Supply Chain

To apply the reinforcement learning (RL) mechanism to the circular supply chain (CSC) described in this work, it is necessary to formulate the problem as an RL model. As previously discussed, RL models are implemented within an agent-based framework, where each component acts as an independent decision-making entity. The first step in this approach is to model each component and process in the CSC as a multi-agent system. Subsequently, the RL problem is defined within this designed agent-based framework.

A circular supply chain involves various operations—reuse, recycling, and disposal—each of which must be managed efficiently to minimize environmental and economic costs while maximizing circularity. In the real world, each component must autonomously make decisions regarding whether to reuse, recycle, or dispose of itself based on its state and the current system dynamics. These autonomous decisions are key to improving the overall performance of the supply chain by considering carbon footprint reduction, cost minimization, and resource circularity. However, the decisions made by individual components must also be coordinated with the overall system objectives to optimize the global performance of the CSC.

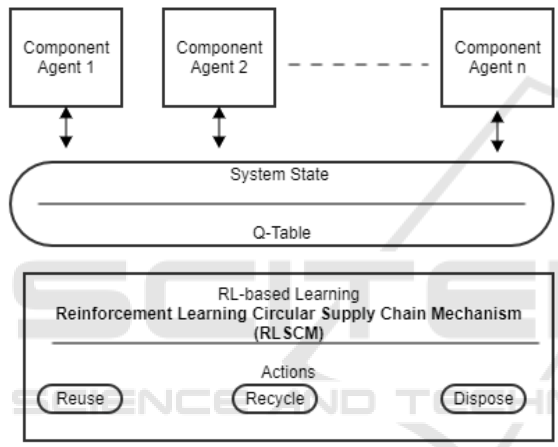

Figure 3: Agent-based framework of circular supply chain EoL management system.

As shown in Fig. 3, the agent-based model treats each component in the CSC as an agent. Each component-agent is responsible for making real-time decisions regarding its EoL treatment—reuse, recycle, or dispose—based on the observed state. These agents interact with the CSC system through a Q-learning-based RL mechanism, where each agent independently learns to optimize its decisions over time based on feedback (rewards) received from the environment.

3.2 RL Modeling of EoL Management Problem in the Circular Supply Chain

In this subsection, we define the characteristics of the reinforcement learning (RL) model used to solve the circular supply chain (CSC) decision-making problem. Key elements of the RL model include the state variable, reward function, value function, and system policy. These components work together within a Q-learning framework to guide agents (components) in deciding between reuse, recycling, and disposal, with the ultimate goal of minimizing the overall carbon footprint, reducing economic costs, and promoting resource circularity.

3.2.1 State Variables

The state of the system at any given time period t is characterized by the state vector:

$$\begin{aligned} S_{i,t} = \{ & x_{i,t}, \\ & \text{reuse\_sr}_i, \\ & \text{recycle\_sr}_i, \\ & \text{reuse\_cf}_i, \\ & \text{recycle\_cf}_i, \\ & \text{dispose\_cf}_i, \\ & \text{reuse\_cost}_i, \\ & \text{recycle\_cost}_i, \\ & \text{dispose\_cost}_i, \\ & \text{weight}_i \}. \end{aligned} \tag{1}$$

The elements of the state variable are detailed as follows:

- **Inventory Level ($x_{i,t}$).** The current quantity of component i in inventory at time t, updated dynamically based on reuse, recycling, or disposal decisions.
- **Reuse Success Rate (reuse_sr$_i$).** The probability that component i can be successfully reused after inspection and cleaning. This rate is based on historical performance.
- **Recycle Success Rate (recycle_sr$_i$).** The likelihood that component i can be recycled if reuse is not possible. This is also derived from past data.
- **Carbon Footprint (CFP).** Environmental impact elements associated with different actions:
  - reuse_cf$_i$ (reuse).
  - recycle_cf$_i$ (recycling).
  - dispose_cf$_i$ (disposal).

  These values are used to assess the environmental impact of each action. They are calculated using the Ecoinvent 3.10 database and Brightway Life Cycle Assessment (LCA) software in a parametric approach, connected to a Python program for automatic calculations. Specifically, the model uses the avoided burden 0.100 method for EoL management to assign environmental credits to reuse and recycling actions, based on the research conducted by Nicholson et al. (2009).

- **Costs.** Financial costs of various actions for component i:
  - reuse_cost$_i$.
  - recycle_cost$_i$.
  - dispose_cost$_i$.

  These values are used to assess the financial impact of each action. For the reuse action, a saving equivalent to the component's original price is applied, while recycling savings are based on values from relevant research. Disposal incurs a cost paid to third-party companies for waste management.
- **Weight (weight$_i$).** The physical weight of the component.

3.2.2 Action Set

The actions available for each component are:

$$A = \{a_{reuse}, a_{recycle}, a_{dispose}\}, \tag{2}$$

where a_{reuse}, $a_{recycle}$, and $a_{dispose}$ represent the actions to reuse, recycle, or dispose of a component, respectively.

3.2.3 Transition Dynamics

The transition from state S_t to S_{t+1} is determined by the amount of returned product and the action chosen for each component. The state transition for component i can be expressed as:

$$x_{i,t+1} = x_{i,t} - q_i(a_t) + r_{i,t}, \tag{3}$$

where

- $q_i(a_t)$ is the quantity of component i used in period t, based on the action a_t,
- $r_{i,t}$ is the return quantity of component i in period t.

3.2.4 Reward Function

The reward function $R(S_t, a_t)$ incorporates three components: environmental reward, economic reward, and circularity reward. The total reward is a weighted sum of these three objectives:

$$R(S_t, a_t) = \omega_{env} \cdot R_{env}(S_t, a_t) + \omega_{econ} \cdot R_{econ}(S_t, a_t) + \omega_{circ} \cdot R_{circ}(S_t, a_t), \tag{4}$$

where

- $R_{env}(S_t, a_t)$ is the environmental reward, derived from the CFP of the chosen action,
- $R_{econ}(S_t, a_t)$ is the economic reward, derived from the cost of the chosen action,
- $R_{circ}(S_t, a_t)$ is the circularity reward, based on the contribution of the action to material circularity,
- ω_{env}, ω_{econ}, and ω_{circ} are the respective weights for environmental, economic, and circularity objectives.

The environmental reward $R_{\text{env}}(S_t, a_t)$ minimizes the carbon footprint (CFP) for each action. The effective carbon footprint for reuse, recycle, and dispose actions is determined based on success rates and fallback options.

For *reuse*, the effective carbon footprint is given by:

$$\begin{aligned}\text{effective\_cf} = \text{reuse\_sr}_i \cdot \text{reuse\_cf}_i \quad &+ (1 - \text{reuse\_sr}_i) \cdot \\ &(\text{recycle\_sr}_i \cdot \text{recycle\_cf}_i + (1 - \text{recycle\_sr}_i) \cdot \\ &\text{dispose\_cf}_i)\,.\end{aligned} \tag{5}$$

For *recycle*, the effective carbon footprint is:

$$\begin{aligned}\text{effective\_cf} = (\text{recycle\_sr}_i \cdot \text{recycle\_cf}_i + (1 - \text{recycle\_sr}_i) \cdot \\ \text{dispose\_cf}_i.\end{aligned} \tag{6}$$

For *dispose*, the carbon footprint is simply dispose_cf$_i$.

Thus, the environmental reward is the negative of the effective carbon footprint, multiplied by the component inventory $x_{i,t}$:

$$R_{\text{env}}(S_t, a_t) = -\text{effective\_cf} \cdot x_{i,t}. \tag{7}$$

Similarly, the economic reward $R_{\text{econ}}(S_t, a_t)$ minimizes financial costs. Like the environmental reward, it considers reuse, recycle, and dispose actions, each carrying specific costs.

For *reuse*, the effective cost is calculated as:

$$\begin{aligned}\text{effective\_cost} = \text{reuse\_sr}_i \cdot \text{reuse\_cost}_i \quad &+ (1 - \text{reuse\_sr}_i) \cdot \\ &(\text{recycle\_sr}_i \cdot \text{recycle\_cost}_i + (1 - \text{recycle\_sr}_i) \cdot \\ &\text{dispose\_cost}_i)\,.\end{aligned} \tag{8}$$

For *recycle*, the effective cost is:

$$\begin{aligned}\text{effective\_cost} = \text{recycle\_sr}_i \cdot \text{recycle\_cost}_i + (1 - \text{recycle\_sr}_i) \cdot \\ \text{dispose\_cost}_i.\end{aligned} \tag{9}$$

For *dispose*, the cost is dispose_cost$_i$.

Thus, the economic reward is the negative of the effective cost, multiplied by the component inventory $x_{i,t}$:

$$R_{\text{econ}}(S_t, a_t) = -\text{effective\_cost} \cdot x_{i,t}. \tag{10}$$

The circularity reward $R_{\text{circ}}(S_t, a_t)$ promotes reuse and recycling. The reward is proportional to the amount of inventory successfully reused or recycled.

For *reuse*, the circularity reward is:

$$R_{\text{circ}}(S_t, a_t) = \text{reuse\_sr}_i \cdot x_{i,t} + \text{recycle\_sr}_i \cdot (x_{i,t} - \text{reuse\_sr}_i \cdot x_{i,t}). \quad (11)$$

For *recycle*, the reward reflects the portion recycled:

$$R_{\text{circ}}(S_t, a_t) = \text{recycle\_sr}_i \cdot x_{i,t}. \quad (12)$$

For *dispose*, the circularity reward is zero:

$$R_{\text{circ}}(S_t, a_t) = 0. \quad (13)$$

3.2.5 Value Function and System Policy

The Q-learning algorithm is used to update the action-value function $Q(S_t, a_t)$, which estimates the expected cumulative reward for taking action a_t in state S_t. The update rule is:

$$Q(S_t, a_t) \leftarrow (1-\alpha)Q(S_t, a_t) + \alpha \left[R(S_t, a_t) + \gamma \max_{a'} Q(S_{t+1}, a') \right], \quad (14)$$

where

- α is the learning rate.
- γ is the discount factor for future rewards.
- $\max_{a'} Q(S_{t+1}, a')$ is the maximum expected future reward for the next state S_{t+1}.

3.3 CSC Performance Evaluation

The Q-learning model is simulated to learn the optimal policy for each component. The performance of the circular supply chain is evaluated by tracking key metrics such as total carbon footprint, economic cost, and circularity contribution. These metrics are used to compare the performance of different EoL strategies (reuse, recycle, dispose) and to assess the effectiveness of the Q-learning optimization.

The simulation of the Q-learning model helps in learning the optimal policies for managing the EoL of components. The performance of the circular supply chain is evaluated using the three following key metrics : total carbon footprint, circularity and total economic cost.

3.3.1 Total Carbon Footprint

The total carbon footprint (CFP_{total}) of the circular supply chain is the sum of the carbon footprint generated from the reuse, recycling, disposal of components, and the carbon footprint from virgin material production. It is calculated as:

$$CFP_{\text{total}} = CFP_{\text{csc}} + CFP_{\text{reused}} + CFP_{\text{recycled}} + CFP_{\text{disposed}}, \quad (15)$$

where

$$CFP_{\text{csc}} = CFP_{\text{prod}} \cdot \sum_{t=1}^{T} D_t,$$

$$CFP_{\text{reused}} = \sum_{t=1}^{T} \sum_{i=1}^{n} q_{\text{reused},i,t} \cdot \text{reuse\_cf}_i,$$

$$CFP_{\text{recycled}} = \sum_{t=1}^{T} \sum_{i=1}^{n} q_{\text{recycled},i,t} \cdot \text{recycle\_cf}_i,$$

$$CFP_{\text{disposed}} = \sum_{t=1}^{T} \sum_{i=1}^{n} q_{\text{disposed},i,t} \cdot \text{dispose\_cf}_i,$$

where CFP_{csc} represents the total carbon footprint associated with production, D_t is the product demand at time t, and $q_{\text{reused},i,t}, q_{\text{recycled},i,t}, q_{\text{disposed},i,t}$ represent the quantities reused, recycled, and disposed for component i at time t.

3.3.2 Circularity Contribution

The circularity contribution (CC_{total}) reflects the proportion of materials successfully reused or recycled in the supply chain. It is calculated as:

$$CC_{\text{total}} = \frac{\sum_{t=1}^{T} \sum_{i=1}^{n} \left(q_{\text{reused},i,t} + q_{\text{recycled},i,t} \right)}{\sum_{t=1}^{T} R_t}, \quad (16)$$

where

$q_{\text{reused},i,t}$ is the quantity of component i reused at time t,
$q_{\text{recycled},i,t}$ is the quantity of component i recycled at time t,
R_t is the total product returns at time t.

3.3.3 Total Economic Cost

The total economic cost (C_{total}) is the sum of the production costs and the costs associated with reuse, recycling, and disposal actions. It is given by:

$$C_{\text{total}} = C_{\text{prod}} \cdot \sum_{t=1}^{T} D_t + C_{\text{reused}} + C_{\text{recycled}} + C_{\text{disposed}}, \quad (17)$$

where

$$C_{\text{reused}} = \sum_{t=1}^{T}\sum_{i=1}^{n} q_{\text{reused},i,t} \cdot \text{reuse\_cost}_i,$$

$$C_{\text{recycled}} = \sum_{t=1}^{T}\sum_{i=1}^{n} q_{\text{recycled},i,t} \cdot \text{recycle\_cost}_i,$$

$$C_{\text{disposed}} = \sum_{t=1}^{T}\sum_{i=1}^{n} q_{\text{disposed},i,t} \cdot \text{dispose\_cost}_i,$$

where C_{prod} represents the total production cost, and $q_{\text{reused},i,t}, q_{\text{recycled},i,t}, q_{\text{disposed},i,t}$ represent the quantities reused, recycled, and disposed for component i at time t.

3.4 Proposed Algorithm for Solving RL-Based Circular Supply Chain Optimization Model

In the previous sections, we described and modeled the circular supply chain (CSC) problem in the context of a reinforcement learning (RL) framework. In this part, we propose an algorithm for solving the modeled problem using a Q-learning mechanism, which is a temporal difference method widely used for solving RL problems (Watkins, 1989). The proposed algorithm aims to optimize EoL decision-making by learning the value of Q-functions iteratively. After the learning process, the best action for each component—whether to reuse, recycle, or dispose—is selected as the optimal policy for future decisions.

As shown in Algorithm 1, the system is simulated over multiple episodes and periods, where the Q-values $Q(s,a)$ are learned and updated during the iterations. During each episode, the system tracks the product returns, and for each component in the supply chain, one of the three EoL decisions (Reuse, Recycle, or Dispose) is selected based on the Q-table. In each state, the system selects an action, observes the reward (a function of environmental impact, economic cost, and circularity contribution), and updates $Q(s,a)$ accordingly.

The probability of exploration is a function of the episode number and is reduced as the number of episodes increases, linearly decreasing from 100% exploration (random actions) in the first episodes to 1% in the final episodes. The exploration rate is gradually decreased using a decay factor ε, promoting the balance between exploring new strategies and exploiting the learned Q-values for optimal actions. This ensures that the algorithm explores various EoL strategies in the early phases but converges to the most rewarding policies over time.

Initialize:
  Q-values $Q(s,a) = 0$ for all states s and actions a
  Set learning rate α, discount factor γ, exploration rate ε, total production carbon footprint c_{fp}^{prod}, number of episodes $n_{episodes}$
  Define action space $A = \{\text{Reuse}, \text{Recycle}, \text{Dispose}\}$
Initialize inventory levels $x_{i,t}$ for all components i
while *episode* $\leq n_{episodes}$ **do**
  for *each period* $t = 1, \ldots, T$ **do**
    for *each component* $i = 1, \ldots, n$ **do**
      if *Random exploration* (ε) **then**
        Select a random action $a \in A$;
      else
        Select action $a = \arg\max_a Q(s,a)$ based on Q-values;
      end
      Observe the next state s' and calculate the rewards;
      Combine rewards using the weighted sum of environmental, economic, and circularity factors:

$$r(s,a) = w_{env} \cdot r_{env}(s,a) + w_{eco} \cdot r_{eco}(s,a) + w_{circ} \cdot r_{circ}(s,a) \qquad (18)$$

      Update the Q-value using the Q-learning update rule:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\Big(r(s,a) + \gamma \max_{a'} Q(s',a') - Q(s,a)\Big) \qquad (19)$$

      Update the inventory level $x_{i,t}$ by adjusting for reused, recycled, and disposed quantities for the component;
    end
  end
  Decrease exploration rate $\varepsilon \leftarrow \max(0.01, \varepsilon \times 0.99)$;
  Increment episode count;
end
Strategic Planning:
After training, for each component i, retrieve the best action $a = \arg\max_a Q(s,a)$ and update inventory levels accordingly. Evaluate system-wide metrics;
return *Optimal Q-values for each component*

Algorithm 1: Q-learning for Circular Supply Chain Optimization.

The reward function during the simulation is computed using a weighted combination of environmental impact, economic costs, and circularity contribution. For each state-action pair, the cumulative reward is calculated, and the Q-function is updated iteratively.

The Q-value update is based on the learning rule:

$$Q(s,a) \leftarrow Q(s,a) + \alpha\,[r(s,a) + \gamma \max_{a'} Q(s',a') - Q(s,a)] \tag{20}$$

where α is the learning rate, $r(s,a)$ is the combined reward, and γ is the discount factor that balances future and immediate rewards.

Initially, the agents (components) have no knowledge of the value of each action in every state. Hence, all Q-values are set to zero for all state-action pairs. The learning rate α determines how much weight to give the most recent reward compared to the existing Q-estimate. A suitable learning rate must be selected to ensure convergence of the algorithm without overfitting to specific samples.

The objective of this model is to maximize circularity and minimize both the total carbon footprint and the economic cost of the circular supply chain. Therefore, the reward function is structured to encourage reuse and recycling while penalizing disposal and excessive carbon emissions. The Q-learning process continues until the Q-values converge for all state-action pairs. After the learning phase, the best action in each state can be retrieved through a greedy search on the Q-table, which then informs the EoL decision for each component in the circular supply chain.

4 RESULTS & ANALYSIS

4.1 Industrial Case Study

In this section, we present an application of the developed model, based on a real industrial case from our partner, GE HealthCare. GE HealthCare (GEHC) is a global leader in the sales and services of medical systems, particularly in the field of medical imaging, with over 4 million systems installed across more than 160 countries. Due to the critical nature of its products (medical devices) and the technological complexity of their components, GE HealthCare implements a circular supply chain strategy through asset recovery and buy-back programs.

In 2023, these initiatives resulted in the recovery of approximately 7,375 units, including imaging systems, ultrasound devices, magnets, and surgical machines, contributing to the reuse of approximately 7.3 million kilograms of materials (GE Healthcare, 2023). The integration of circular economy principles into GE HealthCare's operations and product life cycles is a key approach to managing climate impact, reducing waste, promoting recycling and reuse, and minimizing resource consumption.

In this study, we apply the developed model to a device consisting of 53 components and investigate the impact of EoL management on the efficiency of the circular supply chain (CSC). To demonstrate the benefits of this approach, we evaluate the performance of the proposed Q-learning model across several scenarios, assessing its effectiveness in optimizing CSC outcomes. These scenarios are designed to explore the influence of varying reuse and recycle strategies on CSC circularity, cost reduction, and carbon footprint reduction. Additionally, the performance of these strategies is compared to a baseline linear supply chain scenario, where no circular strategies are implemented.

The goal of the analysis is to reveal the gap between **design circularity**—the ideal circularity based on product design feasibility—and **CSC circularity**—the actual circularity achieved when considering operational constraints in the supply chain.

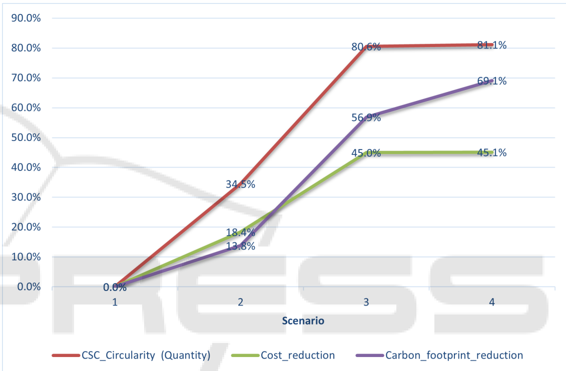

Figure 4: Performance metrics across different scenarios, showing the CSC circularity (by quantity), cost reduction, and carbon footprint reduction.

Table 1 and Figure 4 present the results for four scenarios. In each case, we measure CSC circularity by quantity, along with the total cost reduction and carbon footprint reduction.

4.1.1 Scenario 1: Baseline - No Reuse, No Recycle

Scenario 1 represents the baseline case where no reuse or recycling actions are taken. All returned components are disposed of, resulting in zero circularity. This scenario mimics a traditional linear supply chain where no efforts are made to recover or recycle products at the end of life. As expected, this results in no circularity, no cost savings, and no carbon footprint reduction.

4.1.2 Scenario 2: Partial Reuse Feasibility

In Scenario 2, 41% of the product's components (by design feasibility) are deemed reusable based on product design. However, the actual reuse success

Table 1: Performance metrics for the four scenarios, showing the achieved CSC circularity, cost, and carbon footprint reductions.

Scenario	CSC Circularity (Quantity)	Cost Reduction	Carbon Footprint Reduction
1	0.0%	0.0%	0.0%
2	34.5%	18.4%	13.8%
3	80.6%	45.0%	56.9%
4	81.1%	45.1%	69.1%

rates, drawn from historical data, indicate that not all designed reusable components can actually be reused in the CSC. Despite this, the reuse operations succeed in recovering 34.5% of the product's components, contributing to CSC circularity.

However, this scenario highlights the gap between **design circularity**—which indicates 41% reuse feasibility—and the achieved **CSC circularity** of 34.5%. The operational constraints and quality issues in the supply chain cause a drop in the actual circularity achieved. Similarly, there is a modest 18.4% cost reduction and a 13.8% reduction in carbon footprint, showing that even with partial reuse, significant savings can be realized.

4.1.3 Scenario 3: Advanced Reuse and Recycling

In Scenario 3, 83% of the product's components are designed to be reusable or recyclable. Furthermore, recycling is introduced for components where the process is feasible, with most recyclable components having a success rate of 100%. However, for one critical and rare component, ID 16, the recycling success rate is only 70%, reflecting operational difficulties.

This scenario achieves a substantial 80.6% CSC circularity by quantity. This again highlights the gap between **design circularity** and **CSC circularity**, driven by the limitations of recycling for certain components. The impact on cost and environmental performance is notable, with a 45.0% cost reduction and a 56.9% reduction in carbon footprint. This demonstrates that recycling, even with operational constraints, provides substantial benefits in circular supply chain management.

4.1.4 Scenario 4: Improved Recycling for Component ID 16

In Scenario 4, the only difference from Scenario 3 is that the recycling success rate for Component ID 16 is increased to 100%. This small improvement results in a noticeable rise in both circularity and environmental performance. CSC circularity by quantity increases to 81.1%, and the carbon footprint reduction improves to 69.1%. This scenario demonstrates how addressing operational constraints, even for a single component, can significantly improve circularity and reduce environmental impact.

4.1.5 The Gap Between Design Circularity and CSC Circularity

The scenarios reveal an important insight: the gap between **design circularity** and **CSC circularity**. While product design plays a crucial role in determining circularity potential, the actual circularity achieved in the CSC is constrained by operational and quality issues. These constraints, represented by success rates for reuse and recycling, prevent the full realization of circularity potential in the supply chain.

In Scenario 3, for example, the design circularity suggests that 83% of the product's components can be reused or recycled, but the CSC circularity is only 80.6%. Scenario 4 closes this gap slightly, but only by addressing operational constraints on recycling success rates. This illustrates that optimizing the circular supply chain involves not only design improvements but also a focus on addressing real-world operational challenges that hinder circularity.

4.2 Limitation

One limitation of this study is the circularity metric used to assess the performance of the circular supply chain (CSC). In this analysis, the metric for CSC circularity is based on the quantity of components reused and recycled compared to what has been returned. This approach assumes that reuse and recycling contribute equally to circularity, which is not always the case. In practice, the reuse process typically does not require virgin materials, while recycling may sometimes necessitate the addition of virgin materials to meet product specifications and requirements.

This simplification overlooks the varying resource-saving potentials of reuse and recycling. To provide a more accurate assessment of circularity, future research should incorporate more comprehensive circularity metrics that capture the true resource-saving benefits of reuse and recycling actions.

5 CONCLUSIONS & PERSPECTIVES

This research presents a decision-making mechanism for circular supply chains focusing on EoL management that integrates environmental, economic, and circularity performance through Q-learning. By leveraging historical success rates for reuse and recycling actions, the model reflects the operational realities of EoL management, contrasting the idealized design circularity with actual circular supply chain (CSC) performance. Our findings demonstrate that even when components are designed for reuse or recycling, operational constraints, such as quality issues, can significantly impact the realized circularity, carbon footprint reduction, and cost savings.

Through scenario analysis, we showed the trade-offs between various EoL strategies and the sensitivity of circularity and sustainability outcomes to component-specific success rates. The model addresses the decision-making challenge of whether to reuse, recycle, or dispose of returned components in a circular supply chain, providing a robust framework for managing EoL in a sustainable and cost-effective manner.

However, the cost modeling in this study assumes simplified scenarios. Future work should refine these cost assumptions to better reflect real-world complexities and explore other optimization techniques, such as multi-objective evolutionary algorithms (e.g., NSGA-III or genetic algorithms), to better explore the solution space and identify trade-offs between different objectives. Additionally, while the model currently evaluates environmental impact primarily through carbon footprint reduction, future studies should incorporate a broader range of impact categories to provide a more comprehensive environmental analysis.

Moreover, the potential for integrating more comprehensive circularity metrics to capture the resource-saving benefits of circular supply chains should be explored to further enhance decision-making in circular supply chain management.

REFERENCES

Abaku, E. A. and Odimarha, A. C. (2024). Sustainable supply chain management in the medical industry: a theoretical and practical examination. *International Medical Science Research Journal*, 4(3):319–340.

Agyemang, M., Kusi-Sarpong, S., Khan, S. A., Mani, V., Rehman, S. T., and Kusi-Sarpong, H. (2019). Drivers and barriers to circular economy implementation: An explorative study in pakistan's automobile industry. *Management Decision*, 57(4):971–994.

Bouzon, M., Govindan, K., and Rodriguez, C. M. T. (2018). Evaluating barriers for reverse logistics implementation under a multiple stakeholders' perspective analysis using grey decision making approach. *Resources, conservation and recycling*, 128:315–335.

Bressanelli, G., Perona, M., and Saccani, N. (2019). Challenges in supply chain redesign for the circular economy: a literature review and a multiple case study. *International Journal of Production Research*, 57(23):7395–7422.

de Sousa Jabbour, A. B. L., Luiz, J. V. R., Luiz, O. R., Jabbour, C. J. C., Ndubisi, N. O., de Oliveira, J. H. C., and Junior, F. H. (2019). Circular economy business models and operations management. *Journal of cleaner production*, 235:1525–1539.

D'Alessandro, C., Szopik-Depczyńska, K., Tarczyńska-Łuniewska, M., Silvestri, C., and Ioppolo, G. (2024). Exploring circular economy practices in the healthcare sector: A systematic review and bibliometric analysis. *Sustainability*, 16(1):401.

Farooque, M., Zhang, A., Thürer, M., Qu, T., and Huisingh, D. (2019). Circular supply chain management: A definition and structured literature review. *Journal of cleaner production*, 228:882–900.

GE Healthcare (2023). Sustainability Report. https://www.gehealthcare.com/-/jssmedia/gehc/us/files/about-us/sustainability/reports/ge-healthcare-sustainability-report-2023.pdf?rev=-1. Accessed: 2024-09-24.

Genovese, A., Acquaye, A. A., Figueroa, A., and Koh, S. L. (2017). Sustainable supply chain management and the transition towards a circular economy: Evidence and some applications. *Omega*, 66:344–357.

Govindan, K. and Hasanagic, M. (2018). A systematic review on drivers, barriers, and practices towards circular economy: a supply chain perspective. *International Journal of Production Research*, 56(1-2):278–311.

Goyal, S., Esposito, M., and Kapoor, A. (2018). Circular economy business models in developing economies: lessons from india on reduce, recycle, and reuse paradigms. *Thunderbird International Business Review*, 60(5):729–740.

Halse, L. L. and Jæger, B. (2019). Operationalizing industry 4.0: Understanding barriers of industry 4.0 and circular economy. In *Advances in Production Management Systems. Towards Smart Production Management Systems: IFIP WG 5.7 International Conference, APMS 2019, Austin, TX, USA, September 1–5, 2019, Proceedings, Part II*. Springer.

Han, J., Ijuin, H., Kinoshita, Y., Yamada, T., Yamada, S., and Inoue, M. (2021). Sustainability assessment of reuse and recycling management options for end-of-life computers-korean and japanese case study analysis. *Recycling*, 6(3):55.

Hasegawa, S., Kinoshita, Y., Yamada, T., and Bracke, S. (2019). Life cycle option selection of disassembly parts for material-based co2 saving rate and recovery cost: Analysis of different market value and labor cost for reused parts in german and japanese cases. *International Journal of Production Economics*, 213:229–242.

Herczeg, G., Akkerman, R., and Hauschild, M. Z. (2018). Supply chain collaboration in industrial symbiosis networks. *Journal of cleaner production*, 171:1058–1067.

Hoveling, T., Nijdam, A. S., Monincx, M., Faludi, J., and Bakker, C. (2024). Circular economy for medical devices: Barriers, opportunities and best practices from a design perspective. *Resources, Conservation and Recycling*, 208:107719.

Khodier, A., Williams, K., and Dallison, N. (2018). Challenges around automotive shredder residue production and disposal. *Waste Management*, 73:566–573.

Kinoshita, Y., Yamada, T., Gupta, S. M., Ishigaki, A., and Inoue, M. (2016). Disassembly parts selection and analysis for recycling rate and cost by goal programming. *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, 10(3):JAMDSM0052–JAMDSM0052.

Kirchherr, J., Yang, N.-H. N., Schulze-Spüntrup, F., Heerink, M. J., and Hartley, K. (2023). Conceptualizing the circular economy (revisited): an analysis of 221 definitions. *Resources, Conservation and Recycling*, 194:107001.

Kravchenko, M., Pigosso, D. C., and McAloone, T. C. (2019). Towards the ex-ante sustainability screening of circular economy initiatives in manufacturing companies: Consolidation of leading sustainability-related performance indicators. *Journal of Cleaner Production*, 241:118318.

Lahane, S., Kant, R., and Shankar, R. (2020). Circular supply chain management: A state-of-art review and future opportunities. *Journal of Cleaner Production*, 258:120859.

Mangla, S. K., Luthra, S., Mishra, N., Singh, A., Rana, N. P., Dora, M., and Dwivedi, Y. (2018). Barriers to effective circular supply chain management in a developing country context. *Production Planning & Control*, 29(6):551–569.

Mishra, J. L., Hopkinson, P. G., and Tidridge, G. (2018). Value creation from circular economy-led closed loop supply chains: a case study of fast-moving consumer goods. *Production Planning & Control*, 29(6):509–521.

Nasir, M. H. A., Genovese, A., Acquaye, A. A., Koh, S., and Yamoah, F. (2017). Comparing linear and circular supply chains: A case study from the construction industry. *International Journal of Production Economics*, 183:443–457.

NHS (2022). Delivering a 'net zero' national health service. https://www.england.nhs.uk/greenernhs/wp-content/uploads/sites/51/2022/07/B1728-delivering-a-net-zero-nhs-july-2022.pdf. Accessed: 2024-09-24.

Nicholson, A. L., Olivetti, E. A., Gregory, J. R., Field, F. R., and Kirchain, R. E. (2009). End-of-life lca allocation methods: Open loop recycling impacts on robustness of material selection decisions. In *2009 IEEE international symposium on sustainable systems and technology*, pages 1–6. IEEE.

Pichler, P.-P., Jaccard, I. S., Weisz, U., and Weisz, H. (2019). International comparison of health care carbon footprints. *Environmental research letters*, 14(6):064004.

Ranta, V., Aarikka-Stenroos, L., Ritala, P., and Mäkinen, S. J. (2018). Exploring institutional drivers and barriers of the circular economy: A cross-regional comparison of china, the us, and europe. *Resources, Conservation and Recycling*, 135:70–82.

Romanello, M., Di Napoli, C., Green, C., Kennard, H., Lampard, P., Scamman, D., Walawender, M., Ali, Z., Ameli, N., Ayeb-Karlsson, S., et al. (2023). The 2023 report of the lancet countdown on health and climate change: the imperative for a health-centred response in a world facing irreversible harms. *The Lancet*, 402(10419):2346–2394.

Rosa, P., Sassanelli, C., and Terzi, S. (2019). Circular business models versus circular benefits: An assessment in the waste from electrical and electronic equipments sector. *Journal of cleaner production*, 231:940–952.

Roy, T., Garza-Reyes, J. A., Kumar, V., Kumar, A., and Agrawal, R. (2022). Redesigning traditional linear supply chains into circular supply chains–a study into its challenges. *Sustainable Production and Consumption*, 31:113–126.

Tura, N., Hanski, J., Ahola, T., Ståhle, M., Piiparinen, S., and Valkokari, P. (2019). Unlocking circular business: A framework of barriers and drivers. *Journal of cleaner production*, 212:90–98.

Vermunt, D. A., Negro, S. O., Verweij, P. A., Kuppens, D. V., and Hekkert, M. P. (2019). Exploring barriers to implementing different circular business models. *Journal of cleaner production*, 222:891–902.

Watkins, C. J. C. H. (1989). Learning from delayed rewards.

Winkler, H. (2011). Closed-loop production systems—a sustainable supply chain approach. *CIRP Journal of Manufacturing Science and Technology*, 4(3):243–246.