Towards Automated and Robust Production Planning in Remanufacturing: Utilizing Semi Markov Decision Processes

Maurice Engels(✉), Martin Benfer, and Gisela Lanza

Wbk—Institute of Production Science, Karlsruhe Institute of Technology (KIT),
Karlsruhe, Germany
`maurice.engels@kit.edu`

Abstract. Remanufacturing recovers critical components and materials from returned products, enhancing resource efficiency, supply chain resilience, and supporting circularity. To successfully implement remanufacturing operations, production planning must address significant challenges from uncertain return quality, fluctuating volumes, and variable process times, necessitating a stochastic planning approach. This paper introduces a Semi-Markov Decision Process framework that explicitly models stochastic return arrivals and dynamic processing conditions. Policy iteration guides lot-sizing and capacity-allocation decisions within a SimPy simulation. Benchmarking against an EOQ-based, bottleneck-first heuristic demonstrated increased cumulative output and marked the first promising results of the proposed methodology.

Keywords: Remanufacturing · Production planning

1 Production Planning in Remanufacturing Processes

In an era of dwindling resources, pollution and climate change, sustainable production paradigms such as remanufacturing are gaining importance. Remanufacturing is a process that preserves value by converting one or more used products into items that match or exceed the originals [1], thus maximizing the environmental and economic benefits of sustainable production. Implementing remanufacturing processes however is difficult, as return volumes and product quality are uncertain and disassembly process times vary. As a result, determining production batch sizes and assigning capacity is especially challenging when both inputs and outputs are uncertain. Effective production planning is therefore essential to increase profitability in remanufacturing and achieve environmental objectives, such as reducing resource consumption in production.

1.1 Challenges in Remanufacturing Production Planning

Uncertainty lies at the core of remanufacturing, affecting every stage of production planning and scheduling. Unlike traditional manufacturing, with forecastable raw-material supply and product demand, remanufacturing depends on

L. Overmeyer and B.-A. Behrens (eds.), *Production at the Leading Edge of Technology*,
Lecture Notes in Production Engineering, https://doi.org/10.1007/978-3-032-19524-1_9

stochastic return streams whose timing and volume are often unpredictable, leading to either idle capacity or excess inventories [2]. Returned cores also vary widely in condition, therefore processes range from minor refurbishment to full disassembly, which induces large processing time variances and undermines fixed-batch strategies without costly pre-sorting [3]. As a result, planners must design schedules and capacity buffers that hedge against fluctuations in both supply of cores and market demand, effectively balancing service levels with cost efficiency [4].

Operational unpredictability also arises during the disassembly and reprocessing stages. Each returned unit may follow a unique routing through the facility, and variable failure rates can disrupt expected throughput. This transforms material requirements planning into a multi-stream inventory problem, as planners track and manage both recovered parts and new components—each with different quality grades, yields, and lead times [5].

Finally, economic, strategic, and regulatory considerations add another layer of complexity. Remanufacturing cost spanning inspection, disassembly, and reprocessing are highly variable and depend on return quality, necessitating integrated cost-accounting in planning models to maintain profitability across uncertain yield scenarios [6].

1.2 State of the Art in Remanufacturing Production Planning

Early remanufacturing planning extended linear models to include reverse logistics but often under estimated safety stocks by assuming known returns and fixed quality [7]. To address uncertainty, multi-stage stochastic methods (e.g., Stochastic Integer Programming (SIP) have been used to generate scenario trees for returns and demand in capacity allocation [8]. Furthermore, analytical cycle policies offer closed-form schedules based on return rates, capacities, and holding costs, refined via discrete-event simulations [9]. Markovian process-reliability models represent core degradation and rework as probabilistic state transitions, facilitating optimal costreliability trade-offs for remanufacturing cell design [6]. Additionally, Markovian Models have been used to optimize Order Dispatching in Remanufacturing Systems [10].

Despite these advances, remanufacturing models face practical limitations. Large scale stohastic methods such as SIP and SDDIP suffer from scenario explosion as uncertainties in product types, quality grades and lead time variations multiply, making exact solutions computationally infeasible. Stochastic models that treat uncertainties in isolation overlook their interactions, and separate decision support tools cannot reliably guide capacity or investment choices [8]. Furthermore, returns are often treated as independent and identically distributed, quality distributions are only represented as coarse fuzzy classifications, and correlations between the uncertainty factors are frequently not taken into account [11]. These simplifications potentially hide systemic risks, leading to schedule disruptions and higher costs when return patterns differ from expectations. An integrated modeling paradigm is thus required, combining scalable stochastic approaches and simultaneous management of interdependent uncertainties to create resilient and cost-effective remanufacturing systems. Therefore, this paper will present a novel framework that allows to effectively address uncertainties in remanufacturing production planning and accordingly improves the efficiency of remanufacturing.

2 Semi Markov Decision Processes

In this paper, a Semi-Markov Decision Process (SMDP) models and improves remanufacturing production planning. We first review Markov Decision Processes to expose the fixed time-step limitation, then show how SMDPs' yield greater planning flexibility.

2.1 From Markov to Semi Markov Decision Processes

A Markov Decision Process (MDP) is a tuple (S, A, P, R, γ) modeling sequential decisions under uncertainty [12]. Here, S and A are the state and action sets; $P(s' \mid s, a)$ the probability of moving to s' from s via a; $R(s, a, s')$ the immediate reward; and $\gamma \in (0, 1)$ the discount factor. The Markov property ensures that transitions depend only on the current stateaction pair. A stationary policy $\pi\colon S \rightarrow A$ maximizes the expected discounted return

$$V^{\pi}(s) = \mathbb{E}_{\pi}\big[\sum_{t=0}^{\infty} \gamma^{t}\, R(s_t, a_t, s_{t+1}) \mid s_0 = s\big]. \tag{1}$$

An SMDP generalizes the classical MDP by allowing transitions to occur at irregular, state-, and action-dependent time intervals rather than fixed steps. In an MDP, decisions and transitions occur at fixed intervals , without memory of elapsed time (Fig. 1a).

In contrast, an SMDP generalizes this by introducing random sojourn times, during which the system remains in a state-action pair before transitioning (Fig. 1b). This flexibility captures real-world timing variability essential to remanufacturing. MDP solution methods still apply to SMDP [13]. In this paper policy iteration will be used, which at iteration k computes V^{π_k} via the Bellman equations and then updates π_{k+1}, this process can be seen in Fig. 2. Through Policy Iteration the best possible solution can be approximated iteratively.

One advantage of the policy evaluation method shown in Fig. 2 is that if it used to approximate the policy rather than compute the exact policy, which is often infeasible, it can deal with large state spaces [14], making it suitable for complex systems such as remanufacturing production systems.

2.2 Advantages of SMDPs for Decisions in Remanufacturing

Unlike traditional stochastic programming, which requires discretization into a finite scenario tree that grows exponentially with the number of uncertainties and stages [15], SMDPs operate directly in continuous event time. This

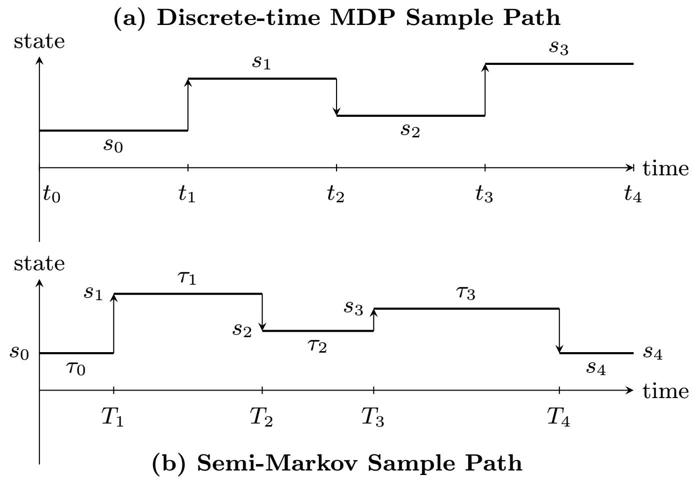

Fig. 1. **a** Discrete-time MDP; **b** semi-Markov decision process [13]

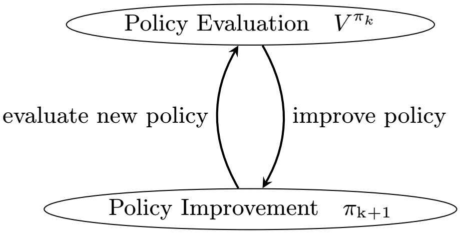

Fig. 2. Policy iteration alternates between evaluating the current policy and improving it by one-step lookahead [14]

allows compact, state-specific policies and continuous cost accrual based on real action durations [13]. This event-driven structure yields substantial planning benefits for remanufacturing. Capacity allocation becomes adaptive: resources can be reassigned immediately when delays or completions change availability. Lot-sizing improves as lots are released precisely when both cores and capacity align, minimizing setup waste and holding costs without introducing batching delays. Inventory is managed through continuous cost tracking, supporting finer-grained trade-offs between batch frequency and setup overhead. Finally, lead-time scheduling reflects the actual sequence of random durations, allowing urgent jobs to preempt ongoing ones and reducing both average and worst-case delays. These features collectively deliver lower inventory levels, increased throughput, and improved delivery reliability compared to discrete-time stochastic models.

3 Production Planning SMDP

After demonstrating the benefits of SMDPs, a simplified SMDP for remanufacturing that focuses on batch-sizing and short-term worker allocation is introduced, although the methodology can be extended to cover multiple planning activities. The remanufacturing production planning SMDP is defined with the following elements. The SMDP's state at any decision epoch is defined by the tuple $s = (I_c, I_f, W, L, d_j)$, which captures the current inventories of cores and finished goods (I_c, I_f), work-in-process levels (W), available labor (L), and the deadline of the next order (d_j). From each state s, the decision maker chooses an action $a = (e_1, \ldots, e_K, q)$, where e_k denotes the number of workers allocated to station k (for $k = 1, \ldots, K$) and q specifies the quantity of finished goods to release. Upon taking action a in state s, the process yields a reward.

$$r = R_{\text{base}} + \begin{cases} B, & t_{\text{delivery}} \leq d_j, \\ 0, & \text{otherwise}, \end{cases} - c_{\text{labor}} \sum_{k=1}^{K} e_k - h_c\, I_c - h_f\, I_f. \tag{2}$$

where R_{base} is the base reward from selling the finished goods, B is an incentive for finishing orders on time and h_c, h_f are holding costs for the cores and finished products respectively.

The Transition-Kernel $Q(s', \tau \mid s, a)$ is a probability distribution over the next state s' and sojourn time τ given the current state s and action a. It encapsulates both the dynamics of the system and the randomness of timing, in regards to core arrivals, new orders and completed orders. Decisions are made at event times T_n, with system dynamics driven by random durations between events—capturing variability in core arrivals, processing times, and resource availability. This SMDP is solved by using policy iteration (see Sect. 2.2).

4 Simulation-Based Production Planning Using an SMDP

The proposed SMDP for remanufacturing production planning is implemented using a simulation approach. This simulation driven approach is visualized in Fig. 3.

The remanufacturing workflow, including the processes inspection, disassembly, cleaning, assembly, and testing, is encoded as a discrete-event model, with hard machine and labor capacity constraints and minimum staffing to ensure feasibility. The Simulation was modeled after a real remanufacturing process of alternators. Core returns (timing, quantity, quality) are generated stochastically; to reduce state-space size, cores are grouped into three quality classes. A core's class determined defect probabilities during disassembly. The SimPy model allows dynamic allocation of workers and batch sizes based on policy decisions. The SMDP-derived production policy is deployed in SimPy to simulate operations and record discounted reward trajectories. These rewards were then fed back into the SMDP for iterative policy evaluation and improvement.

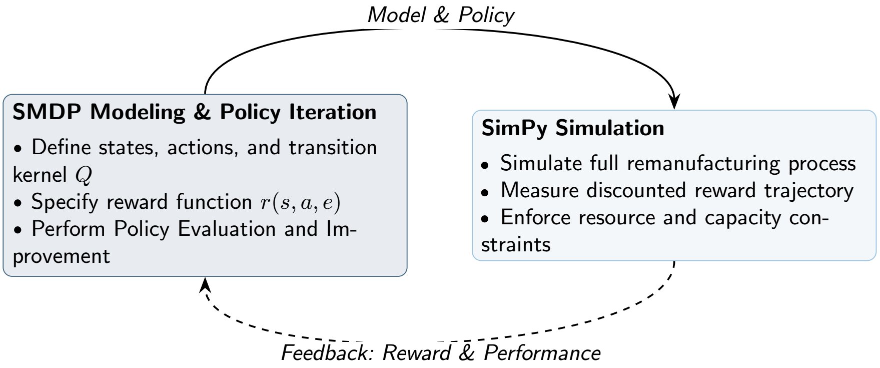

Fig. 3. Feedback loop between SMDP modeling (policy iteration) and SimPy validation for remanufacturing planning

Using the simulation framework illustrated in Fig. 3, production policies can be evaluated in a realistic environment before deployment in real-world settings. By leveraging a SimPy model, it is possible to closely replicates actual remanufacturing processes—capturing resource constraints, stochastic returns, and process dynamics—and thus enhances the validity and practical relevance of the performance metrics generated by the underlying SMDP.

4.1 Preliminary Results

The proposed and implemented SMDP model in the given model and parameters favours smaller lot sizes, which is apparent in Fig. 4, which depicts the lot-size decisions made by the SMDP policy.

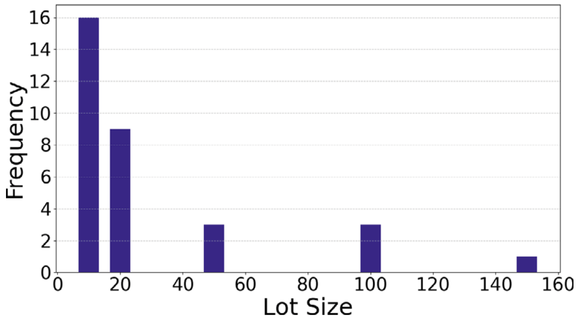

Fig. 4. Chosen lot sizes under the SMDP policy

The favouring of smaller lots is intuitive for remanufacturing, as smaller batches reduce exposure to potential component failure or unusable parts, lowering scrap and recovery costs. Additionally, smaller, more frequent batches accommodate random core returns more effectively, stabilizing inventory and enhancing flexibility. Furthermore, in the presented SMDP relatively high holding costs incentive the use of smaller batches.

To assess the effectiveness of the SMDP-based policy, it was compared against a heuristic using the classical Economic Order Quantity (EOQ) model,

$$Q^* = \sqrt{\frac{2\,D\,S}{h}} \tag{3}$$

where: D is the demand rate (units per period), S is the setup cost per batch, h is the holding cost per unit per period and Q is the lot-size calculated. This model minimizes total setup and holding costs. For capacity allocation, a bottleneck-first approach was applied: at fixed two-hour intervals, available extra capacity was preferentially assigned to the production process with the highest utilization. This combined heuristic balances responds to workload changes, providing a suitable benchmark for evaluating the SMDP policy. Figure 5 presents the results obtained from implementing this heuristic.

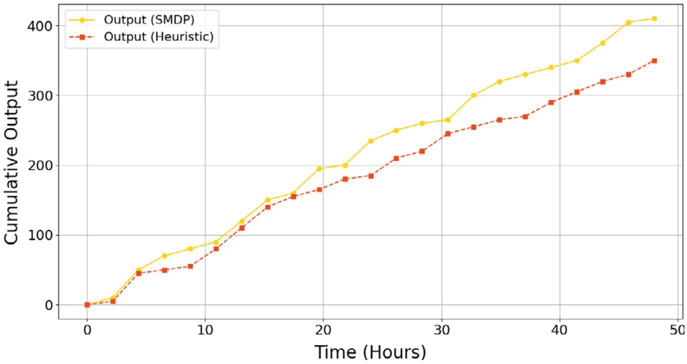

Fig. 5. Comparison finished products heuristic versus SMDP

Figure 5 illustrates the cumulative output over a 48-h simulation horizon. The SMDP policy clearly delivers a higher total of finished products than the benchmark heuristic, demonstrating its superior ability to navigate the complexities inherent in remanufacturing production planning.

5 Conclusion and Discussion

This paper introduces an SMDP framework for remanufacturing planning that integrates stochastic core returns, resource constraints, and process dynamics through policy iteration and a SimPy simulation. The resulting decision-support tool outperforms a classical heuristic in cumulative output while managing uncertainty effectively.

However, three main limitations remain. The model's finite-horizon, short term planning scope limits long-term planning and risks state-space explosion, though hierarchical or multi-timescale generalized formulations could mitigate this. Second, comprehensive benchmarking against industry-standard heuristics is needed to validate practical performance and uncover potential gaps. Third, incorporating finer quality grading, dynamic pricing, or correlated uncertainties makes explicit transition probabilities infeasible; sampling-based methods like Monte-Carlo Simulation offer scalable alternatives that bolster planning robustness. Addressing these areas will enhance the presented planning framework.

Acknowledgments. This research has received funding by the German Federal Ministry of Education and Research (BMBF)—funding number 02L19C250 and from the European Union's Horizon Europe research and innovation programme under Grant Agreement 101138930—Digitally-enhanced multi-level solution for smart human-centric remanufacturing (rEUman Project).

Competing Interests. The author(s) has no competing interests to declare that are relevant to the content of this manuscript.

References

1. DIN SPEC 91472:2023-06: Remanufacturing (Reman)—Qualitätsklassifizierung für Zirkuläre Prozesse (2023)
2. Lanza, G., Deml, B., Matthiesen, S., Martin, M., Brützel, O., Hörsting, R.: The vision of the circular factory for the perpetual innovative product. Automatisierungstechnik **72**(9), 774–788 (2024). https://doi.org/10.1515/auto-2024-0012
3. Zhang, W., Shi, J., Zhang, S., Chen, M.: Scenario-based robust remanufacturing scheduling problem using improved biogeography-based optimization algorithm. IEEE Trans. Syst. Man Cybernet. Syst. **53**, 3414–3427 (2023). https://doi.org/10.1109/TSMC.2022.3225443
4. Jing, Y., Wang, X., Li, W., Deng, L.: Robust production planning with remanufacturing and heterogeneous demands under return and demand uncertainty. Int. J. Appl. Math. Stat. **41**, 183–196 (2013). https://doi.org/10.1007/s13243-025-00149-8
5. Hoffmann, M., Krini, A., Mueller, A., Knorn, S.: Remanufacturing production planning and control: conceptual framework for requirement definition. J. Remanufact. (2025)
6. Stamer, F., Sauer, J.: Optimizing quality and cost in remanufacturing under uncertainty. Prod. Eng. (2024). https://doi.org/10.1007/s11740-024-01314-x
7. Guide, V.D., Jayaraman, V., Srivastava, R.: Production planning and control for remanufacturing: a state-of-the-art survey (1999). https://doi.org/10.46254/an12.20220224
8. Quezada, F., Gicquel, C., Kedad-Sidhoum, S.: A stochastic dual dynamic integer programming based approach for remanufacturing planning under uncertainty. Int. J. Prod. Res. **61**, 5992–6012 (2022). https://doi.org/10.1080/00207543.2022.2120924
9. Polotski, V., Kenné, J., Gharbi, A.: Set-up and production planning in hybrid manufacturing–remanufacturing systems with large returns. Int. J. Prod. Res. **55**, 3766–3787 (2017). https://doi.org/10.1080/00207543.2017.1293863

10. Wurster, M., Michel, M., May, M.C., et al.: Modelling and condition-based control of a flexible and hybrid disassembly system with manual and autonomous workstations using reinforcement learning. J Intell Manuf **33**, 575–591 (2022). https://doi.org/10.1007/s10845-021-01863-3
11. Ropi, N.M., Hishamuddin, H., Wahab, D.A., Saibani, N.: Optimisation models of remanufacturing uncertainties in closed loop supply chains—a review. IEEE Access 1 (2021). https://doi.org/10.1109/ACCESS.2021.3132096
12. Garcia, F. and Rachelson, E.: Markov decision processes. In: Sigaud, O., Buffet, O. (eds.) Markov Decision Processes in Artificial Intelligence (2013). https://doi.org/10.1002/9781118557426.ch1
13. Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley (1994). https://doi.org/10.1002/9780470316887
14. Rachelson, E., Quesnel, G., Garcia, F., Fabiani, P.: A Simulation-Based Approach for Solving Generalized Semi-Markov Decision Processes, pp. 583–587 (2008). https://doi.org/10.3233/978-1-58603-891-5-583
15. Shapiro A., Dentcheva, D., Ruszczyński, A.: Lectures on Stochastic Programming: Modeling and Theory (2009)
16. Birge, J.R., Louveaux F.: Introduction to stochastic programming, Springer (1997)