Technische Universität München
TUM School of Management

Decision-Making in Dynamic, Stochastic Environments with Large Decision Spaces: Integrated Planning for Spare Parts and Technician Routing

Dai Trong Pham

Vollständiger Abdruck der von der TUM School of Management der Technischen Universität München zur Erlagung des akademischen Grades eines Doktors der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.) genehmigten Dissertation.

Vorsitz:		Prof. Dr. Jingui Xie
Prüfende der Dissertation:	1.	Prof. Dr. Gudrun P. Kiesmüller
	2.	Prof. Dr. Jan Fabian Ehmke

Die Dissertation wurde am 31.01.2024 bei der Technischen Universität München eingereicht und von der TUM School of Management am 15.03.2024 angenommen.

Acknowledgements

Dear reader,

as I stand at this pivotal juncture in my life, marking the completion of my Ph.D., I feel an overwhelming sense of gratitude for the incredible support and love that has been given to me in the past few years. This adventure, sprinkled with its fair share of hurdles and joyful moments, simply would not have been the same without the presence of these amazing people in my life.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. Gudrun Kiesmüller, for her unwavering support, patience, and guidance throughout my Ph.D. journey. She has not only been a dedicated academic mentor but also a source of invaluable life lessons. Her wisdom and expertise have been the guiding lights that illuminated my path to the successful completion of this dissertation.

I am also immensely grateful to all the members of the Center for Digital Transformation at TUM Campus Heilbronn. Your good fellowship and the intellectually stimulating environment you provided made my time in Heilbronn not just productive, but thoroughly delightful. A special thanks to my co-author, Jana, for the discussions and cooperation that greatly enriched our joint work. I would also like to extend my gratitude to my friends and colleagues in the research training group AdONE for the memorable experiences we have shared.

To my family, who mean the world to me: my mother, Thuy; my father, Ho; and my younger sister, Thao. Despite being tens of thousands of kilometers away, you have always been the home I could return to, the ears that would always listen, and the hearts that continuously supported me. Your love and belief in me have been my stronghold.

Finally, to my wife, Kim - words cannot express my gratitude for your presence in my life. During times of struggle and doubt, you were my beacon of hope. Your unwavering support, understanding, and love have been the cornerstones of my strength and perseverance. This accomplishment is as much yours as it is mine.

Thank you for being a part of this journey.

Dai

List of Publications

This work is a cumulative dissertation based on two full papers that I have published as the first author in international, English-language, peer-reviewed journals. Additionally, it includes a working paper where I am a co-author. The list of papers is as follows:

- **Pham, D. T.**, & Kiesmüller, G. P. (2022). *Multiperiod integrated spare parts and tour planning for on-site maintenance activities with stochastic repair requests. Computers & Operations Research*, 148, 105967.
  DOI: https://doi.org/10.1016/j.cor.2022.105967

- **Pham, D. T**., & Kiesmüller, G. P. (2024). *Hybrid value function approximation for solving the technician routing problem with stochastic repair requests. Transportation Science*. Available online. Pending volume and page number.
  DOI: https://doi.org/10.1287/trsc.2022.0434

- Ralf, J., **Pham, D. T.**, & Kiesmüller, G. P. (2024). *Optimal outbound shipment policy for an inventory system with advance demand information*. (Working paper)

Abstract

The dissertation delves into innovative methodologies for dynamic decision-making in stochastic environments characterized by large decision spaces. It specifically aims to improve after-sales field services and shipment consolidation by applying these innovative techniques. The motivation stems from the rapidly evolving economic landscape that demands unprecedented agility and adaptability from businesses. Such a landscape necessitates the ability to make informed decisions over time, even with limited information about future events. Operationally, this involves moving away from the traditional static and deterministic approach to decision-making, which assumes all necessary information is available at decision time. Central to this research is the development of the Hybrid Value Function Approximation (H-VFA), a novel approach that merges a graph-encoding method with a genetic search mechanism and a graph neural network. Methodologically, the transition from conventional tabular methods to a graph neural network, coupled with the handling of the combinatorial decision space through genetic search, enables the H-VFA to effectively navigate the well-known curses of dimensionality. Additionally, this dissertation introduces a learning framework based on the post-decision state concept (Powell, 2022), which excels in finding near-optimal policies in problems with extensive transition spaces. Finally, the comprehensive numerical studies conducted provide valuable insights and have practical implications for industry practitioners.

Contents

List of Abbreviations

DQN	Deep Q-Network
DRL	Deep Reinforcement Learning
GNN	Graph Neural Network
H-VFA	Hybrid Value Function Approximation
MJRKP	Multiple-Job Repair Kit Problem
RKP	Repair Kit Problem
SDVRP	Stochastic Dynamic Vehicle Routing Problem
SJRKP	Single-Job Repair Kit Problem
TRSP	Technician Routing and Scheduling Problem
VFA	Value Function Approximation
VRP	Vehicle Routing Problem

Chapter

1 Introduction

1.1 Motivation

In the rapidly evolving economic landscape, businesses must demonstrate unprecedented agility and adaptability. This involves making informed decisions over time, even when complete information about the future is unavailable. From an operational perspective, it necessitates a paradigm shift from traditional static and deterministic decision-making, which assumes that all relevant information is available at the decision time and remains unchanged over time. Dynamic decision-making involves continuously evaluating and re-evaluating decisions based on evolving circumstances and information availability. In this environment, stochastic elements add further to the complexity.

Consider, for instance, a repair service provider who must decide on technician dispatches and manage the spare parts inventory in their service vehicle without complete knowledge of future repair requests or part failures. Similarly, a transport manager in a multi-echelon supply chain faces the challenge of deciding how to consolidate transportation requests to a lower echelon without foresight into future orders. The difficulty is often further amplified by the combinatorial aspect of decisions at each decision point, which is typical in operations planning problems.

This dissertation aims to thoroughly address the previously described challenges, with a specific focus on the after-sales service sector. The overarching theme of the dissertation revolves around *dynamic* decision-making in *stochastic* environments with *large decision spaces*. Specifically, the first two essays concentrate on the joint planning of technician routing and spare parts management in service vehicles for white goods repair. The final essay builds on the optimization methodologies from the first two essays and shifts focus to a different problem setting, developing near-optimal shipment consolidation policies in a multi-echelon supply chain with advance demand information.

While after-sales support is often considered a major source of revenue and a primary factor distinguishing competitors in the context of durable goods (Kurata and Nam, 2010), the operational complexity in this sector cannot be understated. Service providers face significant logistical challenges, particularly in the complex coordination of spare parts management and technician routing. This is especially pronounced for products with a large number of replaceable components and when dealing with sparsely distributed repair requests across expansive geographical areas. In this setting, uncertainties manifest in two ways. First, despite the rapid modernization of appliances, perfect triangulation of needed spare parts before the first technician visit remains difficult. Second, the exact repair request location and initiation time are also uncertain in this field, largely due to the erratic nature of breakdowns and the widespread presence of white goods. Meanwhile, dynamism is a pertinent characteristic in field service optimization, as decisions have to be re-evaluated daily (i.e., re-planning) with the availability of new information. Finally, the combinatoric nature of the decisions originates from the large number of routing and repair kit stocking options, further exacerbates the issue. Even in deterministic and static environments, choosing a good decision from the set of potential decisions is already an extremely difficult task.

According to Dutta (2013), even though top-performing firms prioritize resolving issues during the initial visit, their success rate, known as the first-time fix rate, is only approximately 89%. Furthermore, in over half of the cases requiring a second visit, the need arises due to the unavailability of necessary parts during the initial technician call-out (Dutta, 2013). Upon reviewing customer feedback from an online review platform[1], it is evident that the problem of missing spare parts during call-outs persists across major players in the field. This issue is particularly significant in the household appliance market, which generates 44.1 billion Euros in revenue in Europe alone (Applia Statistical Report 2021-2022).

From the customer's perspective, a fast and inexpensive repair process can increase the willingness to choose repairs over purchasing new products. This is also better for the environment, as extending product lifespans through repair is a crucial aspect of the circular economy (Laitala et al., 2021). Accordingly, manufacturers and authorized repair providers aim to fulfill these customer expectations while balancing costs and maintaining profitability. Unfortunately, the aforementioned logistical complexities of combined spare parts management and technician routing, often under uncertain and rapidly changing operational environments, represent a huge obstacle on the path to achieving this goal.

Similarly, in the trucking sector, the practice of shipment consolidation, which involves merging multiple smaller shipments into a single larger load for dispatch on the same vehicle, serves as an effective logistics strategy to reduce carbon emissions and energy waste (Ülkü, 2012). However, as shown in the third essay of this dissertation, even in a rather stylized setting, the shipment consolidation problem under consideration is also plagued with extensive dynamism and stochasticity, stemming from the fact that decisions made in one shipment cycle have long-lasting consequences, while exact information about future order placements is not known in advance.

In summary, to effectively address dynamic operational problems involving stochasticity and a large decision space spanning multiple periods, it is crucial to adopt a holistic strategy that simultaneously addresses three dimensions. These include the selection

[1] A compilation can be found at `tinyurl.com/cus-compl`

of appropriate modeling techniques and the development of computationally efficient solution approaches.

1.2 Research Questions

This dissertation seeks to develop a deeper understanding and to propose innovative approaches for solving the aforementioned complex planning problems. Specifically, it aims to provide answers to the following research questions:

Research Question 1. *How can a holistic approach be developed to tackle the complexities associated with multiperiod, dynamic, and stochastic problems that feature a combinatorial decision space?*

Research Question 2. *What impact do solution search and solution evaluation have on the success of the methodology?*

Research Question 3. *When does having perfect information about the future add value to the planning tasks?*

Research Question 4. *What impact do the spatial distribution of customer requests, the urgency of these requests, and their arrival rate have on the strategy for joint technician routing and spare parts stocking?*

The subsequent section of this dissertation provides an extensive and detailed examination of the existing literature pertinent to the topics addressed herein. This comprehensive review is carefully constructed to situate this work within the broader spectrum of current research, highlighting specific gaps and opportunities that this dissertation aims to explore. It delves into a range of problem settings, models, and methodologies previously established in this field, offering a critical analysis of their strengths, limitations, and relevance to the objectives of this dissertation. Following the literature review, a concise yet comprehensive summary of the contributions of this dissertation to the current state of research is presented.

1.3 Literature Review

This section presents a comprehensive discussion of the literature. It covers a wide range of topics, including repair kit problem, on-site repair activities, dynamic and stochastic logistical problems over multiple periods, value function approximation, and shipment consolidation. A brief tangent on related work involving the application of reinforcement learning in routing and inventory control is also presented. This is particularly relevant to the methodology employed in this dissertation's second and third essays.

1.3.1 Repair Kit Problem

The spare parts planning aspect in this dissertation originates from the repair kit problem (RKP), which was introduced by Smith, Chambers, and Shlifer (1980) and focuses on the optimal assortment of service parts for field technicians to carry in their repair kits. This research balances the holding costs of these parts against the penalty costs incurred due to spare part shortages during technician visits. Building on this, Graves (1982) addresses a similar issue but includes a minimum service level requirement in the analysis. Mamer and Smith (1982) explore a more complex scenario where the demand for various service parts may be interdependent, and multiple parts of the same kind may be required for a single repair task. Early studies on the RKP often assume that technicians can replenish their repair kits immediately after each job, a scenario termed the single-job repair kit problem (SJRKP).

The concept of technicians performing multiple jobs between restocking their repair kits, or a *tour*, was proposed by Heeremans and Gelders (1995), leading to the introduction of the multiple-job repair kit problem (MJRKP). A notable advancement in MJRKP was made by Teunter (2006), who derived a formula for calculating the job fill rate assuming independent demand for different part types. In this work, Teunter (2006) assumed that required parts are left at the customer's location whether or not the job is completed. Teunter (2006) also proposed two efficient heuristics, the job heuristic and the part heuristic, for quickly solving large-scale problem instances. Bijvank, Koole, and Vis (2010) expanded this research by considering the number of jobs in a tour as a random variable and incorporating scenarios where multiple units of the same part type may be required for a single job. Differing from Teunter (2006), Bijvank, Koole, and Vis (2010) argued that parts should only be removed from the kit if the job is successfully completed and developed a formula for computing the expected job fill rate under these conditions. More recent studies, such as those by Saccani et al. (2017) and Prak et al. (2017), explore optimizing the budget-constrained repair kit and its replenishment frequency and consider scenarios with positive lead times, respectively. Another extension of the repair kit problem has been recently introduced by Rippe and Kiesmüller (2023), focusing on advance demand information. Their study explores a scenario in which appliances are equipped with sensors capable of displaying error codes when malfunctions are detected. These error codes act as imperfect yet proactive indicators of the spare parts needed. The authors propose two well-performing heuristics for addressing this scenario.

Notably, many studies in this field are driven by practical industrial settings. For instance, the research presented in Saccani et al. (2017) is based on an actual case from a global office equipment company. Similarly, Prak et al. (2017) and Neves-Moreira, Veldman, and Teunter (2021) introduce service models that have been tested using data from an equipment manufacturer and a wind turbine manufacturer, respectively.

Despite these significant advancements in addressing the RKP, previous research has often overlooked aspects such as planning routes, allocating customers to trips, and the impact of incomplete repairs on subsequent tours in white goods repair. In situations where external emergency services are available, imposing a penalty cost is an adequate way to model the problem. However, when a technician must revisit a site, failure in repairing the appliance on the first call-out results in additional waiting time and increased workload for the technician in subsequent days, ultimately leading to customer dissatisfaction with the services provided.

1.3.2 Planning On-site Repair Activities

The planning of on-site repair activities, due to its high practical relevance, has garnered significant attention in the operations research and operations management communities. A key focus within this research domain is the technician routing and scheduling problem (TRSP). Castillo-Salazar, Landa-Silva, and Qu (2016) offer a comprehensive review of TRSP.

Research on the routing and scheduling of technicians for maintenance activities often considers various real-world attributes such as technician skill levels, team formations, and time-dependent rewards. For example, Tsang and Voudouris (1997) explored the scheduling challenges faced by British Telecom, emphasizing the variable proficiency levels of technicians and their impact on task completion times. Similarly, Xu and Chiu (2001) delved into scenarios where service requests vary in priority and are subject to the technicians' skill levels. Tang, Miller-Hooks, and Tomastik (2007) investigated scenarios where the rewards for task completion vary depending on the service period, reflecting customer preferences in choosing the desired repair date. The challenge posed by the French Operations Research Society led to studies by Hashimoto et al. (2011), Kovacs et al. (2012), and Fırat and Hurkens (2012), focusing on the formation of technician teams with diverse skills to accomplish a range of tasks. Meanwhile, Chen, Thomas, and Hewitt (2017) introduced the concept of learning, where technicians enhance their proficiency through repeated task performance.

However, research that concurrently addresses technician routing and spare parts management remains scarce. Pillac, Gueret, and Medaglia (2013) pioneered this approach by integrating consumable resources, such as spare parts, with technician routing. In their model, however, spare part demands at each customer site were known prior to the tour, and technicians had the opportunity to replenish their repair kits by visiting the depot. More recently, Mathlouthi, Gendreau, and Potvin (2021) tackled a similar deterministic problem, adding the challenge of a special spare part that cannot be initially carried by technicians and must be picked up from the depot during the day.

Another closely related area of research is the scheduling and routing of maintenance activities for offshore wind farms. This field, while incorporating unique industry-specific attributes like extended offshore durations and variable weather conditions, shares many parallels with onshore maintenance planning. In a recent study, Neves-Moreira, Veldman, and Teunter (2021) addressed optimizing service operation vessels for offshore wind farm maintenance, modeling the onboard spare parts inventory management problem as an RKP. Their model, however, assumed known wind turbine locations, with the vessel stationed offshore for extended periods and the option for helicopter resupply of spare parts. This dissertation models and solves the multiperiod combined routing and spare parts planning problem, where future requests and spare parts demands are not known in advance.

1.3.3 Multiperiod Dynamic and Stochastic Logistical Problems with Combinatorial Decision Space

The planning problems investigated in this dissertation share similarities with various multiperiod dynamic and stochastic logistical problems, particularly those involving combinatorial decision spaces. In these problems, decision-makers are confronted with a wide array of choices at each decision point, ranging from selecting transportation routes to restocking inventory or scheduling personnel, all undertaken amidst the uncertainties of operations that extend over multiple time periods. The long-term implications of these decisions are intensified by the inherent dynamism and stochasticity in these settings.

In general, there are primarily two modeling strategies used to address these types of problems: two-stage stochastic programming with recourse and sequential decision problems. The former, a method commonly used for modeling uncertainties, often falls short in fully capturing the dynamics over extended periods, as pointed out in Powell (2022). This approach has been applied to a range of multiperiod problems in areas such as location-transportation (Klibi et al., 2010), disaster relief logistics (Alem, Clark, and Moreno (2016); Moreno, Alem, and Ferreira (2016)), and waste management (Gambella, Maggioni, and Vigo, 2019). Conversely, the sequential decision model emphasizes the progressive nature of real-world planning problems over multiple time periods. Notable examples include Chen, Thomas, and Hewitt (2017) in technician scheduling, Avraham and Raviv (2021) in mobile personnel scheduling, and Ulmer, Soeffker, and Mattfeld (2018) and Liu and Luo (2022) in vehicle routing and driver dispatching, respectively. However, many existing studies tend to overlook one or more key aspects: dynamism, stochasticity, or the combinatorial nature of decision-making. For instance, Chen, Thomas, and Hewitt (2017) circumvents stochasticity through point forecasting; Avraham and Raviv (2021) sidestep combinatorial decision space by utilizing tentative routes; Ulmer, Soeffker, and Mattfeld (2018) limit decision points through periods categorization, and Liu and Luo (2022) employ decision decomposition and myopic policies to approximate the cost-to-go function.

Neglecting or oversimplifying any of the three aspects, either in modeling or in developing solution strategies, can undermine the model's adaptability and the quality of the decisions derived from it. This dissertation aims to address this issue by comprehensively and simultaneously considering all three aspects in both the modeling process and the solution methodology.

1.3.4 Value Function Approximation

Value function approximation (VFA), as described in Powell (2007), is a technique for solving sequential decision problems that aim to balance immediate rewards with future rewards. It operates by mapping problem states to their expected values, thereby enabling decision-making processes that consider future impacts. These mapping functions, often referred to as cost-to-go functions, support forward-looking decision-making by selecting options that minimize the sum of immediate and anticipated future costs.

However, implementing VFA in practical, real-world scenarios poses substantial challenges, primarily due to the well-known *curses of dimensionality*. To manage the complexity of the transition space, many methods favor a *model-free* learning approach over a *model-based* strategy. This model-free methodology circumvents the need for explicit computation of state transition probabilities, relying instead on iterative learning through trial and error, often using Monte Carlo simulation for exploration and learning (Gläscher et al., 2010). Conversely, the model-based approach requires an explicit model, including complete state transitions, to derive the optimal policy. Addressing the vast state space often necessitates extensive aggregation into basis functions, each with a unique weight. Solutions typically involve a linear combination of these basis functions with learned weights, integrated into integer programs that are then solved, either accurately or approximately, to make decisions, as seen in the work of Van Heeswijk, Mes, and Schutten (2019) and Heinold, Meisel, and Ulmer (2022). The primary challenge is developing a set of basis functions that effectively capture complex, potentially nonlinear relationships between state representations and future value contributions. A significant drawback of using basis functions is their assumption of well-defined relationships between attributes and state value, potentially limiting the quality of the derived policy, as noted in Ulmer (2017). Moreover, the frequent reliance on integer programs and mathematical solvers for decision extraction limits the size of solvable problems and hampers learning speed. The high-dimensional discrete decision space poses another formidable challenge for advanced deep-learning-based reinforcement learning methods, as highlighted in Hildebrandt, Thomas, and Ulmer (2022).

In modern reinforcement learning approaches, the focus is on directly learning policies or values linked to state-decision pairs. However, these methods face a significant challenge when dealing with high-dimensional discrete decision spaces, as this leads to an exponential increase in the potential decision set. Off-the-shelf reinforcement learning algorithms, such as deep Q-network (DQN) (Mnih et al., 2015), policy gradient (Sutton et al., 1999), and actor-critic (Konda and Tsitsiklis, 1999), are not directly applicable in real-world operational settings as they require significant modifications to handle the combinatorial nature of the decision space.

This work introduces two novel solutions to address the aforementioned challenges in applying VFA: a unique graph-encoding technique, and a genetic search combined with a graph neural network learning method.

1.3.5 Shipment Consolidation

Previous research has identified three practical and easy-to-implement policies for shipment consolidation: time-based, quantity-based, and hybrid policy (Çetinkaya, 2005). In the time-based consolidation policy, outbound shipments are dispatched at regular intervals, each spanning a specific number of time units. For the quantity-based policy, an outbound shipment is initiated each time the total quantity of items for consolidation reaches a predetermined level. In the case of the hybrid policy, outbound shipments are scheduled either at the conclusion of a set time interval or when the specified quantity for consolidation is achieved, depending on which condition is met first.

Cetinkaya and Lee (2000) and Axsäter (2001) explore stochastic single-echelon inventory systems using policies where orders are placed when inventory falls below a certain level and are filled up to another level (i.e., an (s, S) policy), along with time-based shipment consolidation. In this setting, Axsäter (2001) presents an exact approach for determining the parameters for the shipment and inventory policies. Studies by Çetinkaya and Bookbinder (2003), Chen, Wang, and Xu (2005), and Çetinkaya, Mutlu, and Lee (2006) compare time-based and quantity-based dispatch policies. Analytical expressions for three heuristic policies are provided by Mutlu, Çetinkaya, and Bookbinder (2010) and Wei, Kapuscinski, and Jasin (2021). Joint transport and inventory decisions in single-echelon systems are addressed in works by Cheung and Lee (2002) and Toptal, Çetinkaya, and Lee (2003). In multi-echelon inventory systems, Marklund (2011) offers an exact approach for time-based consolidation setting, while quantity-based consolidation is studied by Kiesmüller and De Kok (2005) and Malmberg and Marklund (2023). Extensions to time-based dispatching can be found in Stenius et al. (2016), Johansson et al. (2020), and Sonntag, Schrotenboer, and Kiesmüller (2023).

While a significant amount of research has been conducted on shipment consolidation in the realm of inventory management, studies that integrate advance demand information are rare. Wang and Toktay (2008) were pioneers in studying a stochastic single-echelon inventory system featuring periodic shipments and flexible deliveries. More recently, Ralfs and Kiesmüller (2022) explored a single-stage inventory system that incorporates advance demand information and flexible deliveries, operating under a time-based shipment consolidation policy, although their approach was based on a heuristic outbound shipment strategy. Notably, research that integrates shipment consolidation with advance demand information and focuses on the derivation of optimal or near-optimal dispatching policies remains unexplored.

This dissertation explores near-optimal outbound shipment policies for single-echelon inventory systems using reinforcement learning, focusing on time-based consolidation with flexible deliveries, and incorporating advance demand information.

1.3.6 Reinforcement Learning in Vehicle Routing and Inventory Control

In recent years, there has been a significant increase in publications applying deep reinforcement learning (DRL) to address operational planning problems. While the number of research applying reinforcement learning to technician routing is limited, there are extensive studies on the vehicle routing problem (VRP), a setting that shares many similarities. In the following, only works that apply reinforcement learning to solve the dynamic and/or stochastic variant of the VRP is discussed. A summary of work using reinforcement learning for the deterministic and static versions is discussed in Raza, Sajid, and Singh (2022).

The stochastic dynamic vehicle routing problem (SDVRP), situated in the realm of dynamic combinatorial optimization within operations research, tackles sequential decision-making while searching vast routing action spaces and evaluating those actions concerning future uncertainties (Hildebrandt, Thomas, and Ulmer, 2022). Most of the work on SDVRP is usually in the context of *online* routing. In this context, customers (or requests) appear dynamically throughout the day, and the decisions typically involve acceptance, assignment, and/or routing. Ulmer et al. (2019) studied the single-vehicle routing problem with stochastic service requests. The authors developed an anticipatory algorithm that employs a roll-out strategy combined with VFA. In a similar context, Joe and Lau (2020) developed an algorithm combining simulated annealing and a fully connected deep neural network to solve the dynamic vehicle routing problem, considering time windows and stochastic requests. Wang et al. (2018) solve the online order dispatching problem using a DQN. The authors also propose a transfer learning method to quickly learn and adapt to changes in location and time. The problem of electric vehicle routing with uncertain energy consumption and stochastic requests was investigated by Basso et al. (2022), where the authors developed a Q-learning-based algorithm. Chen, Ulmer, and Thomas

(2022) also applied a DQN, but in the context of same-day delivery using both vehicles and drones. The recent survey by Hildebrandt, Thomas, and Ulmer (2022) offers a comprehensive discussion in this field. In the same survey, the authors also expressed their desire to develop a general-purpose algorithm capable of both searching and evaluating vast decision spaces.

In the field of inventory control, there has also been a notable increase in the application of DRL methodologies. A recent survey by Boute et al. (2022) provides a thorough overview of DRL in inventory control, including implementation guidelines. Key recent studies in this area include research by Gijsbrechts et al. (2022) on the lost sales problem, Vanvuchelen, Gijsbrechts, and Boute (2020) investigating the joint replenishment problem, Oroojlooyjadid et al. (2022) exploring the supply chain beer game, and De Moor, Gijsbrechts, and Boute (2022) concentrating on perishable inventory management. These studies commonly utilize advanced DRL methods, such as the DQN (Mnih et al., 2013), asynchronous advantage actor-critic (Mnih et al., 2016), and proximal policy optimization (Schulman et al., 2017), all of which depend on a learned "policy network" for decision-making.

1.4 Contributions and Methodologies

Drawing upon the comprehensive literature review presented in Section 1.3, this section is dedicated to elucidating the research gaps in the field and delineating the ways in which this dissertation aims to bridge these gaps. It includes a detailed discussion of the dissertation's contributions across several dimensions: the *novelty of the problem settings* under study, the development and application of *innovative solution methodologies*, and the derivation of *valuable practical insights*.

1.4.1 Contributions in Problem Formulations

The literature review presented in Sections 1.3.1 and 1.3.2 highlights a substantial gap in existing research, particularly in the simultaneous planning of technician routing and repair kit stocking amidst stochastic demands for spare parts. The first and second essays in this dissertation aim to bridge this gap by developing a novel approach to address these challenges. Moreover, scenarios where incomplete tasks require the same technician to revisit rather than relying on external emergency services for completion are also incorporated.

Another unique contribution of this dissertation is the integration of new stochastic requests that emerge within the planning period. This aspect introduces a dynamic element to the problem, necessitating continuous re-planning to effectively integrate these new requests along with the evolving information collected from previous visits, particularly about the remaining missing spare parts.

The first essay of this dissertation presents a comprehensive modeling of the joint planning problem, formulating it as a stochastic mixed-integer program. Building upon this foundation, the second essay delves deeper by addressing a similar problem through the lens of a sequential decision process. Furthermore, the problem is enriched by including the element of random deadlines, adding another layer of complexity to the planning and execution of technician routing and repair tasks and allowing for the derivation of the added value of flexibility. The two essays offer a novel approach to managing routing and spare part planning in maintenance and repair operations, especially under dynamism and stochasticity spanning multiple periods.

The problem discussed in the third essay, revolving around the optimization of shipment consolidation, was previously examined in the study by Ralfs and Kiesmüller (2022). The third essay builds upon that foundation by conceptualizing the problem as a sequential decision process and introducing an innovative methodology. The essay focuses on efficiently learning a near-optimal policy for shipment consolidation by leveraging advance demand information.

1.4.2 Contributions in Solution Methodologies

This dissertation introduces innovative solutions for improving value function approximation, referred to as hybrid value function approximation (H-VFA). The first and second essays stress the importance of simultaneously considering dynamism, stochasticity, and large decision spaces. The first contribution is a novel graph-encoding method that captures both spatial and temporal aspects of operations planning, thereby eliminating the need for manually created basis functions. This state-encoding is based on multi-attribute graphs, where nodes represent individual repair requests (i.e., the request's attributes), and edges reflect the travel distances between them. Additionally, spatial markers, which are artificial nodes, are incorporated into the graph to assist in determining the relative positions of repair requests within the service area.

The second contribution is a method that combines genetic search with a graph neural network, facilitating efficient policy learning in large, discrete decision spaces. This method demonstrates high performance and adaptability in various scenarios and maintains robustness against changes in problem parameters and cost structures. These versatile strategies may also be applicable to other multiperiod operations planning tasks that face similar challenges.

The H-VFA is designed to tackle the curse of dimensionality through several strategies. First, this approach avoids the immense state space by moving away from traditional tabular methods and instead utilizes a graph neural network along with graph representation to generalize state values more effectively. Additionally, the H-VFA addresses the complex combinatorial decision space using a genetic search method. Finally, this strategy tackles the vast transition space by implementing systematic state sampling

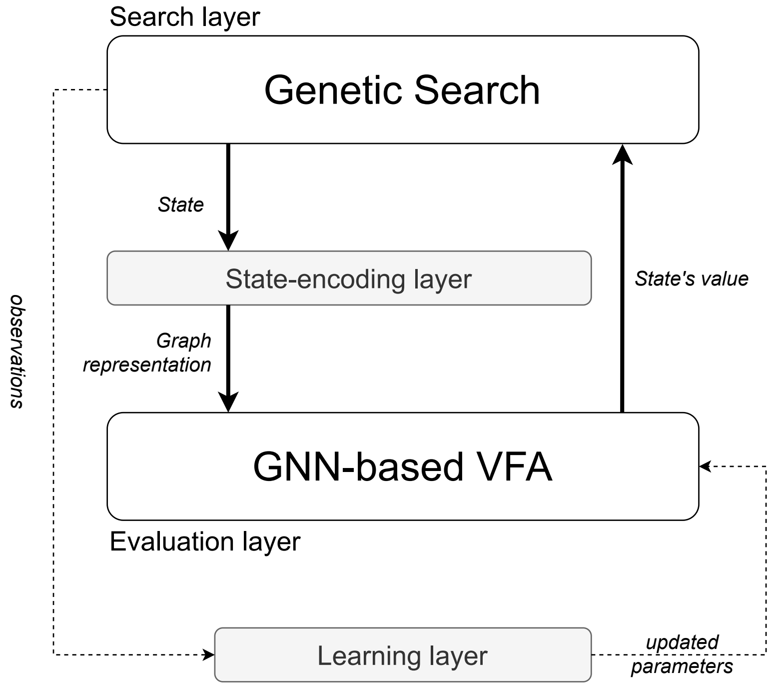

Figure 1.1: High-level overview of the H-VFA

within the genetic search procedure. Figure 1.1 shows the high-level overview of the H-VFA. By developing the H-VFA and testing it against other benchmark policies in various problem and parameter settings, an answer to ***Research Question 1*** is provided. This strategy specifically demonstrates how to develop a comprehensive approach for tackling the complexities inherent in multi-period, dynamic, and stochastic problems, which feature a combinatorial decision space.

The third essay introduces a learning framework based on the concept of post-decision states, as described by Powell (2022). This framework adopts state value approximation as a strategy to navigate the constraints of fixed size that are commonly present in the decision space of policy-based reinforcement learning methods. Unlike policy-based techniques, which are limited by the number of output nodes in the policy network, state value approximation allows for a more general optimization problem for decision-making. However, accurately approximating pre-decision state values often requires extensive forward simulation to determine the expected cost-to-go for decisions. This can be problematic in scenarios with large transition spaces, a common characteristic in real-world problems. To address this, the learning framework in the third essay involves approximating the value of the post-decision state, following the methodology outlined by Powell (2022), thereby providing a partial solution to this challenge. In problems with a relatively small decision space, this modification enables near-instantaneous decision-making since no forward simulation is required to estimate the cost-to-go of making a decision. This speeds up the decision process when the policy is already in place and facilitates rapid roll-out during training to aid post-decision state value approximation.

1.4.3 Contributions in Practical Insights

From the extensive numerical studies performed in the three essays, this dissertation derives several practical insights, which are particularly helpful for industry practitioners who wish to improve their operations and researchers who would like to adopt the method for problems facing similar challenges.

The following insight provides an answer to ***Research Question 2***:

Insight 1. *In the context of jointly planning technician route and spare parts, precise evaluation is more crucial than an exhaustive search for solutions, especially in scenarios that demand intelligent anticipation, such as those with high arrival rates but moderate to low urgency. A basic approach to generating solutions can lead to good performance, provided that the evaluations are accurate. However, achieving an accurate cost-to-go estimation during the training phase necessitates a comprehensive solution search.*

Concurrently, the following insights relate to ***Research Question 3***:

Insight 2. *The significance of possessing information about future requests in the technician routing problem diminishes as the urgency of incoming requests increases. When these requests have distant deadlines, having accurate foresight greatly reduces* *costs by providing more opportunities for efficient consolidation.*

Insight 3. *In the context of time-based shipment consolidation, where flexible deliveries are permitted, having advance demand information can significantly reduce costs. This reduction is achieved by avoiding late delivery fees and emergency shipping expenses through the proactive dispatch of orders. However, the cost-saving impact of advance demand information diminishes with longer lead times. Additionally, the impact of advance demand information on total cost reductions is not linear.*

Meanwhile, the following insights provide answers to ***Research Question 4***:

Insight 4. *Planning in an anticipatory manner is important, especially when customer requests are sparsely distributed, or when there are opportunities to perform spatial and temporal consolidation (e.g., when the incoming requests are not immediately urgent, or when requests are in clusters).*

Insight 5. *In situations where there is a frequent high influx of incoming requests with extended deadlines, it is crucial for anticipatory learning policies to be trained using an accurate data distribution. Moreover, inaccuracies such as underestimating the arrival rate or overestimating the urgency parameter during training can lead to significantly increased costs in the execution phase.*

Insight 6. *As the average daily volume of repair requests rises, it is advisable to increase the inventory of frequently used spare parts on the service vehicle.*

Insight 7. *A high arrival rate coupled with elevated urgency levels results in more missed deadlines and escalated costs due to service delays. The strategy should involve either allocating technicians to areas with manageable arrival rates or exploring methods to enhance flexibility, such as extending due dates.*

1.5 Outline of the Dissertation

This section offers an overview of the dissertation's remaining chapters, providing a clear roadmap for the reader to navigate the subsequent parts of the dissertation.

In Chapter 2, a summary of each essay is provided, along with a detailed description of my individual contributions to each essay. All original publications (the first and second essays) and the working paper (the third essay) have been included in the Appendix of this dissertation.

Chapter 3 marks the conclusion of this dissertation, summarizing the key findings and reflecting on their implications. The limitations of the applied methodologies are also discussed. Additionally, this chapter explores potential avenues for future research, suggesting how upcoming studies could build upon the insights presented in this work to further advance understanding in this area.

Chapter

2 Summary of Essays and Author Contribution

Publication 1: Pham, D. T., & Kiesmüller, G. P. (2022). *Multiperiod integrated spare parts and tour planning for on-site maintenance activities with stochastic repair requests. Computers & Operations Research*, 148, 105967.
DOI: https://doi.org/10.1016/j.cor.2022.105967

- ***Problem:*** The paper addresses the challenge faced by home-attended maintenance service providers in jointly optimizing spare parts planning and routing a single technician for on-site maintenance activities at geographically distributed customers. This includes dealing with stochastic spare parts demand, customer requests, and the necessity for return visits.

- ***Methods and main results:*** The problem is formulated as a stochastic mixed-integer program, which captures various real-world characteristics. An anticipatory solution approach is proposed, utilizing a look-ahead technique to incorporate stochastic information into the planning process. This method is compared with two myopic approaches. The paper also investigates the value of having perfect information about yet-to-be-realized requests and spare part demand.

- ***Practical insights***: The findings suggest there are benefits from integrating repair kit planning and technician tour scheduling, notably in reducing penalty costs for late service. The numerical study demonstrates that this anticipatory approach can offer cost reductions and improved service levels, particularly in scenarios where customers are sparsely distributed.

My contributions to this paper encompass several areas: First, I participated in modeling the problem. Then, I developed and proposed methods to address this problem. I was primarily responsible for programming and implementing the proposed method, along with other benchmark methods for comparison. Additionally, I played a significant role in analyzing the results obtained from the numerical experiments. Lastly, I was responsible for drafting the manuscript of this paper and participating in subsequent revisions.

Publication 2: Pham, D. T., & Kiesmüller, G. P. (2024). *Hybrid value function approximation for solving the technician routing problem with stochastic repair requests. Transportation Science*. Available online. Pending volume and page number.
DOI: https://doi.org/10.1287/trsc.2022.0434

- ***Problem:*** The paper focuses on the complex planning problem of routing technicians and stocking spare parts for servicing geographically distributed repair tasks. This problem is characterized by operational uncertainties such as unpredictable future repair requests and spare parts needed for replacing malfunctioned components. In addition, new requests are not immediately due but rather have randomly sampled due dates.

- ***Methods and main results:*** The problem is modeled as a sequential decision problem, with decisions made daily about the technician's route and the spare parts inventory. Exact methods are intractable due to high-dimensional state, decision, and transition spaces. To address this, two novel algorithmic techniques are introduced: a hybrid value function approximation method combining a genetic search with a graph neural network, and a unique state-encoding method using multi-attribute graphs and spatial markers. These methods facilitate efficient learning without the need for instance-specific hyperparameter tuning. The numerical study shows that the hybrid learning technique surpasses other benchmark policies and adapts effectively to environmental changes.

- ***Practical insights:*** The study reveals several key managerial insights for optimizing on-site repair tasks. The value of having information about future repair requests is shown to diminish as the urgency of these requests increases; however, when requests have distant deadlines, possessing accurate information can significantly reduce operational costs by facilitating more efficient consolidation of tasks. Training anticipatory learning policies using accurate data distributions is crucial, particularly under conditions of frequent, high-volume requests with extended deadlines. Any miscalculations in estimating arrival rates or urgency during the training phase can lead to increased costs in execution. Additionally, as the average daily volume of repair requests rises, it becomes advisable to increase the inventory of frequently used spare parts on service vehicles. Finally, in scenarios where there is a high arrival rate of requests coupled with elevated urgency, it is essential to adopt strategies that either allocate technicians to areas with manageable arrival rates or enhance operational flexibility, for instance, by extending due dates, to avoid missed deadlines and escalating costs.

My responsibilities and contributions to this paper are as follows. I participated in conceptualizing and modeling the problem. Subsequently, I was responsible for developing and proposing methods to tackle the problem. Furthermore, I proposed various competitive benchmark methods for comparison. I also took the lead in analyzing the results from the numerical experiments, ensuring a comprehensive understanding of the outcomes. Finally, I was responsible for drafting the manuscript and actively participated in the revision process that followed.

Working Paper: Ralf, J., Pham, D. T., & Kiesmüller, G. P. (2023). *Optimal outbound shipment policy for an inventory system with advance demand information*. (Working paper)

- ***Problem:*** This paper focuses on a single-echelon inventory system that fulfills stochastic orders from a production facility under a time-based shipment consolidation scheme. The key objective is to determine the optimal outbound shipment quantities while considering the costs associated with early-delivery, late-delivery, and shipping.

- ***Methods and main results:*** The problem is modeled as a Markov decision process. A deep reinforcement learning algorithm that approximates the value of the post-decision state was developed. The algorithm's effectiveness is validated through comparison with value iteration, revealing an impressively low average optimality gap of 0.08%. The study also finds an easy-to-implement policy, which follows a multi-threshold structure, is competitive. Additionally, simpler heuristic policies are proposed and found to be reasonable in specific cost settings, though they are less effective than the multi-threshold policy.

- ***Practical insights:*** A key finding is that the value of advance demand information does not increase linearly but decreases as the demand lead time increases. This implies that while advance demand information is beneficial if flexible deliveries are allowed, its utility diminishes over a longer lead time. Furthermore, it is recommended that transportation capacity be planned around the mean demand level occurring between two shipments.

As a co-author, my responsibilities and contributions to this paper are as follows: I was involved in modeling the problem as a Markov decision process. Subsequently, I was responsible for developing the DRL approach to tackle the problem. I also took part in programming the main methods and other benchmark policies. I participated in analyzing the results from the numerical experiments. Finally, I contributed to drafting the manuscript, particularly in the sections on modeling and solution approach.

Chapter

3 Summary and Outlook

In summary, this dissertation has advanced the understanding of decision-making in dynamic, stochastic environments with large decision spaces. These advancements were achieved by leveraging value function approximation methodology, combined with other techniques such as graph-encoding, graph neural networks, and genetic search. While focused on the integrated planning of spare parts and technician routing, the developed methodologies could potentially be applied to other problems facing similar challenges. Computational experiments demonstrate that these novel approaches are capable of producing effective decisions in complex planning scenarios. The practical insights derived from this dissertation are beneficial for industry practitioners and researchers, offering useful applications and guiding future research in this field.

This dissertation, while yielding promising results, also encounters certain limitations that necessitate further exploration and refinement. A notable challenge is the extensive training time. To address this, we suggest exploring more rigorous hyperparameter tuning and shifting towards distributed and asynchronous learning models, as proposed by Mnih et al. (2016). Additionally, our efforts to estimate problem parameters accurately could benefit from a more thorough empirical analysis and expanded data collection within the repair industry. This would better align our study with real-world planning scenarios. An exciting future direction would be to investigate dynamic route adaptations in real-time, responding to changes in repair kit compositions and new service requests as they arise during the day. This approach could significantly enhance the responsiveness and efficiency of field service operations. Furthermore, the demonstrated effectiveness of post-decision state learning, when combined with deep neural networks and fast roll-out evaluation in managing large state and transition spaces, calls for continued development to further explore and improve this relatively under-explored methodology.

Bibliography

Alem, D., A. Clark, and A. Moreno (2016). “Stochastic network models for logistics planning in disaster relief”. *European Journal of Operational Research* 255(1), pp. 187–206.

Applia Statistical Report 2021-2022 (2022). Applia. URL: `https://statreport2022.applia-europe.eu/` (visited on 10/20/2023).

Avraham, E. and T. Raviv (2021). “The steady-state mobile personnel booking problem”. *Transportation Research Part B: Methodological* 154, pp. 266–288.

Axsäter, S. (2001). “A note on stock replenishment and shipment scheduling for vendor-managed inventory systems”. *Management Science* 47(9), pp. 1306–1310.

Basso, R. et al. (2022). “Dynamic stochastic electric vehicle routing with safe reinforcement learning”. *Transportation research part E: logistics and transportation review* 157, p. 102496.

Bijvank, M., G. Koole, and I. F. Vis (2010). “Optimising a general repair kit problem with a service constraint”. *European Journal of Operational Research* 204(1), pp. 76–85.

Boute, R. N. et al. (2022). “Deep reinforcement learning for inventory control: A roadmap”. *European Journal of Operational Research* 298(2), pp. 401–412.

Castillo-Salazar, J. A., D. Landa-Silva, and R. Qu (2016). “Workforce scheduling and routing problems: literature survey and computational study”. *Annals of Operations Research* 239(1), pp. 39–67.

Çetinkaya, S. (2005). "Coordination of inventory and shipment consolidation decisions: A review of premises, models, and justification". *Applications of supply chain management and e-commerce research*, pp. 3–51.

Çetinkaya, S. and J. H. Bookbinder (2003). "Stochastic models for the dispatch of consolidated shipments". *Transportation Research Part B: Methodological* 37(8), pp. 747–768.

Cetinkaya, S. and C.-Y. Lee (2000). "Stock replenishment and shipment scheduling for vendor-managed inventory systems". *Management Science* 46(2), pp. 217–232.

Çetinkaya, S., F. Mutlu, and C.-Y. Lee (2006). "A comparison of outbound dispatch policies for integrated inventory and transportation decisions". *European Journal of Operational Research* 171(3), pp. 1094–1112.

Chen, F. Y., T. Wang, and T. Z. Xu (2005). "Integrated inventory replenishment and temporal shipment consolidation: A comparison of quantity-based and time-based models". *Annals of Operations Research* 135, pp. 197–210.

Chen, X., B. W. Thomas, and M. Hewitt (2017). "Multi-period technician scheduling with experience-based service times and stochastic customers". *Computers & Operations Research* 82, pp. 1–14.

Chen, X., M. W. Ulmer, and B. W. Thomas (2022). "Deep Q-learning for same-day delivery with vehicles and drones". *European Journal of Operational Research* 298(3), pp. 939–952.

Cheung, K. L. and H. L. Lee (2002). "The inventory benefit of shipment coordination and stock rebalancing in a supply chain". *Management science* 48(2), pp. 300–306.

De Moor, B. J., J. Gijsbrechts, and R. N. Boute (2022). "Reward shaping to improve the performance of deep reinforcement learning in perishable inventory management". *European Journal of Operational Research* 301(2), pp. 535–545.

Dutta, S. (2013). *Field Service 2013: Workforce Management Guide*. Industry Research Report. Previously available at http://www.aberdeen.com/Aberdeen-Library/8325/RA-field-service-workforce.aspx. As of December 27, 2023, the document requires the creation of an account for access at https://www.aberdeen.com/research-insights. Aberdeen Group.

Fırat, M. and C. A. Hurkens (2012). "An improved MIP-based approach for a multi-skill workforce scheduling problem". *Journal of Scheduling* 15(3), pp. 363–380.

Gambella, C., F. Maggioni, and D. Vigo (2019). "A stochastic programming model for a tactical solid waste management problem". *European Journal of Operational Research* 273(2), pp. 684–694.

Gijsbrechts, J. et al. (2022). "Can deep reinforcement learning improve inventory management? Performance on lost sales, dual-sourcing, and multi-echelon problems". *Manufacturing & Service Operations Management* 24(3), pp. 1349–1368.

Gläscher, J. et al. (2010). "States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning". *Neuron* 66(4), pp. 585–595.

Graves, S. C. (1982). "Note—A Multiple-Item Inventory Model with a Job Completion Criterion". *Management Science* 28(11), pp. 1334–1337.

Hashimoto, H. et al. (2011). "A GRASP-based approach for technicians and interventions scheduling for telecommunications". *Annals of Operations Research* 183(1), pp. 143–161.

Heeremans, D and L. Gelders (1995). "Multiple period repair kit problem with a job completion criterion: A case study". *European Journal of Operational Research* 81(2), pp. 239–248.

Heinold, A., F. Meisel, and M. W. Ulmer (2022). "Primal-Dual Value Function Approximation for Stochastic Dynamic Intermodal Transportation with Eco-Labels". *Transportation Science*.

Hildebrandt, F. D., B. W. Thomas, and M. W. Ulmer (2022). "Opportunities for reinforcement learning in stochastic dynamic vehicle routing". *Computers & Operations Research*, p. 106071.

Joe, W. and H. C. Lau (2020). "Deep reinforcement learning approach to solve dynamic vehicle routing problem with stochastic customers". *Proceedings of the international Conference on Automated Planning and Scheduling*. Vol. 30, pp. 394–402.

Johansson, L. et al. (2020). "Controlling distribution inventory systems with shipment consolidation and compound Poisson demand". *European Journal of Operational Research* 280(1), pp. 90–101.

Kiesmüller, G. P. and A. De Kok (2005). *A multi-item multi-echelon inventory system with quantity-based order consolidation*. Technische Universiteit Eindhoven.

Klibi, W. et al. (2010). "The stochastic multiperiod location transportation problem". *Transportation Science* 44(2), pp. 221–237.

Konda, V. and J. Tsitsiklis (1999). "Actor-critic algorithms". *Advances in Neural Information Processing Systems* 12.

Kovacs, A. A. et al. (2012). "Adaptive large neighborhood search for service technician routing and scheduling problems". *Journal of Scheduling* 15(5), pp. 579–600.

Kurata, H. and S.-H. Nam (2010). "After-sales service competition in a supply chain: Optimization of customer satisfaction level or profit or both?" *International Journal of Production Economics* 127(1), pp. 136–146.

Laitala, K. et al. (2021). "Increasing repair of household appliances, mobile phones and clothing: Experiences from consumers and the repair industry". *Journal of Cleaner Production* 282, p. 125349.

Liu, S. and Z. Luo (2022). "On-Demand Delivery from Stores: Dynamic Dispatching and Routing with Random Demand". *Manufacturing & Service Operations Management*.

Malmberg, F. and J. Marklund (2023). "Evaluation and control of inventory distribution systems with quantity based shipment consolidation". *Naval Research Logistics (NRL)* 70(2), pp. 205–227.

Mamer, J. W. and S. A. Smith (1982). "Optimizing field repair kits based on job completion rate". *Management Science* 28(11), pp. 1328–1333.

Marklund, J. (2011). "Inventory control in divergent supply chains with time-based dispatching and shipment consolidation". *Naval Research Logistics (NRL)* 58(1), pp. 59–71.

Mathlouthi, I., M. Gendreau, and J.-Y. Potvin (2021). "A metaheuristic based on Tabu search for solving a technician routing and scheduling problem". *Computers & Operations Research* 125, p. 105079.

Mnih, V. et al. (2013). "Playing atari with deep reinforcement learning". *arXiv preprint arXiv:1312.5602*.

Mnih, V. et al. (2015). "Human-level control through deep reinforcement learning". *Nature* 518(7540), pp. 529–533.

Mnih, V. et al. (2016). "Asynchronous methods for deep reinforcement learning". *International Conference on Machine Learning*. PMLR, pp. 1928–1937.

Moreno, A., D. Alem, and D. Ferreira (2016). "Heuristic approaches for the multi-period location-transportation problem with reuse of vehicles in emergency logistics". *Computers & Operations Research* 69, pp. 79–96.

Mutlu, F., S. i. I. Çetinkaya, and J. H. Bookbinder (2010). "An analytical model for computing the optimal time-and-quantity-based policy for consolidated shipments". *Iie Transactions* 42(5), pp. 367–377.

Neves-Moreira, F., J. Veldman, and R. H. Teunter (2021). "Service operation vessels for offshore wind farm maintenance: Optimal stock levels". *Renewable and Sustainable Energy Reviews* 146, p. 111158.

Oroojlooyjadid, A. et al. (2022). "A deep q-network for the beer game: Deep reinforcement learning for inventory optimization". *Manufacturing & Service Operations Management* 24(1), pp. 285–304.

Pillac, V., C. Gueret, and A. L. Medaglia (2013). "A parallel matheuristic for the technician routing and scheduling problem". *Optimization Letters* 7(7), pp. 1525–1535.

Powell, W. B. (2007). *Approximate Dynamic Programming: Solving the curses of dimensionality*. Vol. 703. John Wiley & Sons.

Powell, W. (2022). *Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions*. Wiley. ISBN: 9781119815037. URL: `https://books.google.de/books?id=y8V6EAAAQBAJ`.

Prak, D. et al. (2017). "The Repair Kit Problem with positive replenishment lead times and fixed ordering costs". *European Journal of Operational Research* 261(3), pp. 893–902.

Ralfs, J. and G. P. Kiesmüller (2022). "Inventory management with advance demand information and flexible shipment consolidation". *OR Spectrum* 44(4), pp. 1009–1044.

Raza, S. M., M. Sajid, and J. Singh (2022). "Vehicle routing problem using reinforcement learning: Recent advancements". *Advanced Machine Intelligence and Signal Processing*. Springer, pp. 269–280.

Rippe, C. and G. P. Kiesmüller (2023). "The repair kit problem with imperfect advance demand information". *European Journal of Operational Research* 304(2), pp. 558–576.

Saccani, N. et al. (2017). "Improving spare parts management for field services: a model and a case study for the repair kit problem". *IMA Journal of Management Mathematics* 28(2), pp. 185–204.

Schulman, J. et al. (2017). "Proximal policy optimization algorithms". *arXiv preprint arXiv:1707.06347*.

Smith, S. A., J. C. Chambers, and E. Shlifer (1980). "Note—optimal inventories based on job completion rate for repairs requiring multiple items". *Management Science* 26(8), pp. 849–854.

Sonntag, D. R., A. H. Schrotenboer, and G. P. Kiesmüller (2023). "Stochastic inventory routing with time-based shipment consolidation". *European Journal of Operational Research* 306(3), pp. 1186–1201.

Stenius, O. et al. (2016). "Exact analysis of divergent inventory systems with time-based shipment consolidation and compound poisson demand". *Operations Research* 64(4), pp. 906–921.

Sutton, R. S. et al. (1999). “Policy gradient methods for reinforcement learning with function approximation”. *Advances in Neural Information Processing Systems* 12.

Tang, H., E. Miller-Hooks, and R. Tomastik (2007). “Scheduling technicians for planned maintenance of geographically distributed equipment”. *Transportation Research Part E: Logistics and Transportation Review* 43(5), pp. 591–609.

Teunter, R. H. (2006). “The multiple-job repair kit problem”. *European Journal of Operational Research* 175(2), pp. 1103–1116.

Toptal, A., S. Çetinkaya, and C.-Y. Lee (2003). “The buyer-vendor coordination problem: modeling inbound and outbound cargo capacity and costs”. *IIE transactions* 35(11), pp. 987–1002.

Tsang, E. and C. Voudouris (1997). “Fast local search and guided local search and their application to British Telecom's workforce scheduling problem”. *Operations Research Letters* 20(3), pp. 119–127.

Ülkü, M. A. (2012). “Dare to care: Shipment consolidation reduces not only costs, but also environmental damage”. *International Journal of Production Economics* 139(2), pp. 438–446.

Ulmer, M. W., N. Soeffker, and D. C. Mattfeld (2018). “Value function approximation for dynamic multi-period vehicle routing”. *European Journal of Operational Research* 269(3), pp. 883–899.

Ulmer, M. W. et al. (2019). “Offline–online approximate dynamic programming for dynamic vehicle routing with stochastic requests”. *Transportation Science* 53(1), pp. 185–202.

Ulmer, M. W. (2017). *Approximate dynamic programming for dynamic vehicle routing*. Vol. 1. Springer.

Van Heeswijk, W. J., M. R. Mes, and J. M. Schutten (2019). “The delivery dispatching problem with time windows for urban consolidation centers”. *Transportation Science* 53(1), pp. 203–221.

Vanvuchelen, N., J. Gijsbrechts, and R. Boute (2020). "Use of proximal policy optimization for the joint replenishment problem". *Computers in Industry* 119, p. 103239.

Wang, T. and B. L. Toktay (2008). "Inventory management with advance demand information and flexible delivery". *Management Science* 54(4), pp. 716–732.

Wang, Z. et al. (2018). "Deep reinforcement learning with knowledge transfer for online rides order dispatching". *2018 IEEE International Conference on Data Mining (ICDM)*. IEEE, pp. 617–626.

Wei, L., R. Kapuscinski, and S. Jasin (2021). "Shipping consolidation across two warehouses with delivery deadline and expedited options for e-commerce and omni-channel retailers". *Manufacturing & Service Operations Management* 23(6), pp. 1634–1650.

Xu, J. and S. Y. Chiu (2001). "Effective heuristic procedures for a field technician scheduling problem". *Journal of Heuristics* 7(5), pp. 495–509.

Appendix

Appendix 1: Essay 1

Pham, D. T., & Kiesmüller, G. P. (2022). *Multiperiod integrated spare parts and tour planning for on-site maintenance activities with stochastic repair requests. Computers & Operations Research*, 148, 105967.
DOI: https://doi.org/10.1016/j.cor.2022.105967

Appendix 2: Essay 2

Pham, D. T., & Kiesmüller, G. P. (2024). *Hybrid value function approximation for solving the technician routing problem with stochastic repair requests. Transportation Science*. Available online. Pending volume and page number.
DOI: https://doi.org/10.1287/trsc.2022.0434

Appendix 3: Essay 3

Ralf, J., Pham, D. T., & Kiesmüller, G. P. (2024). *Optimal outbound shipment policy for an inventory system with advance demand information*. (Working paper)

Note: for copyright reasons, the original publications are not included in the online version of this dissertation.